Positron Emission Tomography and Improved Survival in Patients With Lung Cancer

The Will Rogers Phenomenon Revisited

Karen G. Chee, MD; Danh V. Nguyen, PhD; Monica Brown, PhD; David R. Gandara, MD; Ted Wun, MD; Primo N. Lara Jr, MD

Background: The Will Rogers phenomenon occurs when newer technology allows for more sensitive detection of tumor spread, resulting in stage migration and an apparent improvement in patient survival. We investigated whether use of highly sensitive positron emission tomography (PET) scanning in non–small cell lung cancer has had this effect.

Methods: We performed a retrospective analysis involving 12,395 patients with non–small cell lung cancer in the pre-PET (1994-1998) and PET (1999-2004) periods. Interperiod differences in staging procedures, clinical variables, and patient survival were evaluated.

Results: There was a 5.4% decline in the number of patients with stage III disease and an 8.4% increase in the number of patients with stage IV disease in the PET period, corresponding with an increase in PET use from 6.3% to 20.1% (P < .001). The PET period predicted better survival with a hazard ratio (HR) of 0.95 (95% confidence interval [CI], 0.91-0.99) (P = .02). Use of PET was independently associated with better survival in patients with stage III (HR, 0.77; 95% CI, 0.69-0.85) and stage IV (HR, 0.64; 95% CI, 0.58-0.70) disease, but not those with stage I or II disease.

Conclusion: These data support the notion that stage migration is responsible at least in part for an apparent improvement in survival for patients with stage III and IV non–small cell lung cancer in the PET scan era.

Arch Intern Med. 2008;168(14):1541-1549

Stage migration in oncology describes a change in cancer stage distribution caused by advancements in technology that allow for more sensitive detection of tumor spread and can result in an apparent improvement in patient survival. Stage migration is one form of what is widely referred to as the Will Rogers phenomenon, which occurs if moving an element from one set to another raises the average values of both sets. Rogers' original quip—"When the Okies left Oklahoma and moved to California, they raised the average intelligence level in both states"—humorously illustrates this phenomenon.

In a seminal report, patients with lung cancer treated in 1977 appeared to have higher survival rates compared with those treated between 1953 and 1964, apparently as a result of stage migration from the use of newer diagnostic imaging procedures.1 When patients were classified according to symptom stages that would be unaltered by new diagnostic procedures, the 2 cohorts had similar survival rates.

In the latter part of the 20th century, fluorodeoxyglucose F 18–positron emission tomography (FDG-PET) emerged as a standard diagnostic imaging modality for the staging of lung cancer, offering enhanced sensitivity in detecting otherwise occult tumor spread.2,3 Since its approval for reimbursement for lung cancer staging by Medicare in 1998, PET scanning has seen a tremendous expansion in its use and is now widely available in many communities.4

We hypothesized that FDG-PET scan usage in non–small cell lung cancer (NSCLC) staging is associated with a shift in stage distribution, particularly from stage III to IV, and that this shift would be characterized by an apparent improvement in survival in stage III and IV cohorts. We explored this hypothesis in a large NSCLC cohort from the Sacramento region of the population-based California Cancer Registry (CCR).

METHODS

For the purposes of this analysis, the 10-year study period from 1994 to 2004 was divided into 2: 1994 to 1998 and 1999 to 2004, henceforth known as the pre-PET and PET periods, respectively. The year 1999 was selected as the beginning of the PET period because the inflection point for increasing PET use in the da-
We defined NSCLC using International Classification of Diseases for Oncology, second and third editions, site codes C34.0-C34.9 and histology codes 8004, 8010, 8012, 8013, 8020-8022, 8031-8033, 8046, 8070-8076, 8082-8084, 8123, 8140, 8145, 8230, 8249-8253, 8255, 8260, 8310, 8430, 8550, 8560, 8569, 8574, and 8720.6,7

Race/ethnicity, based on information obtained from medical records, was derived from patient self-identification, assumptions based on appearance, or inferences based on the race/ethnicity of the patient’s parents, birthplace, surname, or maiden name. Hispanic ethnicity was based on information from the medical record and computerized comparisons to the 1980 US Census list of Hispanic surnames. Patients identified as Hispanic on their medical record as well as patients identified as white, black, or of unknown race but who had a Hispanic surname were classified as Hispanic. Patients for whom race/ethnicity, age, or sex were unknown were excluded from the analyses.

Socioeconomic status and rural/urban designation were assigned at the 2000 US Census block group level and based on address at the time of initial diagnosis as reported in the medical record. Quintile of socioeconomic status was derived from a principle component analysis using data from the 2000 US Census.8 Rural/urban designation was defined by Rural-Urban Commuting Area codes developed and categorized by the University of California, Davis, and the CCR.

Stage at diagnosis was collected from the patient’s medical record and coded according to the AJCC Cancer Staging Manual, 6th edition.9 Text fields pertaining to radiographs, scoping procedures, and any additional remarks were extracted from the CCR data management system. Chest radiograph, computed tomographic scan, magnetic resonance imaging, PET scan, mediastinoscopy, bronchoscopy, thoracoscopy, and radiofrequency ablation were coded as “done” or “not done” based on the presence or lack of keywords or phrases within the extracted text fields. Cases that could not be assigned values based on keyword coding were manually reviewed. This study was reviewed and approved by the institutional review boards at the University of California, Davis, and the CCR.

Differences in the distribution of computed tomographic and PET scan usage, age at diagnosis, sex, race, disease stage, treatment received (surgery, radiation, or chemotherapy), and Rural-Urban Commuting Area (urban vs rural) were evaluated using the t-test for continuous variables and χ² testing for proportions. Survival curves were estimated using the Kaplan-Meier method10 for the pre-PET and PET periods. The log-rank statistic was used for the comparison of overall survival distributions.11

Cox regression analysis12 was used to obtain hazard ratios (HRs) for death for PET periods and use of PET, adjusted for age, sex, race, urban vs rural residence, and treatment received (surgery, radiation, or chemotherapy), allowing for different hazards for each stage. Departure from the proportional hazard assumption was checked using graphical methods based on Schoenfeld residuals vs time and testing procedure-based inclusion of time-dependent covariates. Only radiation and chemotherapy showed departure from the Cox model assumption. Thus, we considered an extended Cox model that allows for different hazards by radiation and chemotherapy (stratified Cox model). However, the conclusions and the estimates of HRs (model coefficients) did not change to a significant degree. For this reason, we have chosen to present the simpler model. Logistic regression analysis was used to evaluate the odds ratio of stage IV (relative to stage III) for the pre-PET vs PET periods adjusted for potential confounders. The Hosmer-Lemeshow goodness-of-fit test and diagnostic plots were used to assess model fit and were found to be adequate. Analyses were performed using SAS statistical software, version 9.1 (SAS Inc, Cary, North Carolina), and the statistical software R (http://www.r-project.org/). All P values are for 2-sided tests.

To assess for potential confounders and improve balance in the covariates between patients who underwent PET (PET-yes group or cases) and those who did not (PET-no group or controls), matched pairs were obtained to evaluate survival differences for each stage by matching with the propensity score (PS).13,14 The PS is the probability of receiving a PET scan given the covariates and was obtained for each stage based on logistic regression. Matched pairs (PET-yes and PET-no groups) within each stage were then obtained by matching cases to controls on the PS using a ‘greedy match’ technique.15 Survival for the PET-yes and PET-no groups was estimated using the Kaplan-Meier method. After matching, covariate balance between the PET-yes and PET-no groups was compared using paired t test for continuous covariates and the McNemar test for categorical covariates. For covariates that were not balanced after matching on PS, Kaplan-Meier analysis was based on further stratification of those covariates (eg, chemotherapy). In addition, we also considered covariate/confounder adjustment by including the PS, together with covariates that could not be balanced, in a survival analysis using the Cox regression model.

In the data analysis, 334 patients (2.7%) were excluded because of missing follow-up time. This extremely small number would likely have negligible effects, if any, on the subsequent conclusion. Although there is growing use of multiple imputation methods for missing predictors/covariates, the theory on imputing survival time is lacking, and we chose not to impute missing follow-up time for this analysis.

![Figure 1](https://jamanetwork.com/10/20/2023)
In survival analysis, power is related to the number of events (here, lung cancer deaths) and not the number of subjects. Because the follow-up period in our study was reasonably long for lung cancer, the event was observed for 86.9% of subjects (ie, 10,771 deaths). In regression modeling, studies suggest that at least 10 events per covariate are needed, and for our data there were more than 600 events per covariate. Therefore, the Cox model has sufficient power. For the Kaplan-Meier analysis, we did not do an a priori power analysis (ie, prestudy power).
Changes in the percentages of patients with stage I through IV disease from 1994 to 2004 are depicted in Figure 1. Use of PET scans increased during this period. The increase in the percentage of patients with stage IV disease during this period coincided with a decrease in the percentage of patients with stage III disease. Overall, there was a 5.4% decline in stage III disease and an 8.4% increase in stage IV disease from the pre-PET to the PET period. There was no substantial change in the percentage of patients with stage I or II disease (Table 1). In addition, PET imaging use significantly increased from the pre-PET period to the PET period (6.5% vs 20.1%, respectively) ($P < .001$). Logistic regression analysis was used to evaluate the odds of stage IV (relative to stage III) disease, adjusted for potential confounders (age, sex, race, disease stage, urban residence, and treatment with surgery, radiation, and chemotherapy). For the PET period, the odds ratio for stage IV disease compared with stage III disease was 1.53 (95% CI, 1.40-1.70; $P < .001$). Thus, the adjusted odds for a diagnosis of stage IV disease was significantly higher during the PET period compared with the pre-PET period.

SURVIVAL BY DISEASE STAGE AND PET PERIOD

Unadjusted Analysis
Overall survival significantly improved during the PET period compared with the pre-PET period for stage III ($P < .001$) and IV ($P < .001$) disease. In contrast, survival for patients with stage I and II disease was unchanged. Figure 2 shows the Kaplan-Meier survival curves during the pre-PET and PET periods for each disease stage. The 2-year survival rate for patients with stage III disease during the pre-PET period was 18% (95% CI, 16%-20%) compared with 22% (95% CI, 21%-24%) during the PET period. For patients with stage IV disease,
the 2-year survival rate in the pre-PET period was 6% (95% CI, 5%-7%), increasing to 8% in the PET period (95% CI, 7%-9%).

**Adjusted Analysis**

Cox proportional hazards regression analysis was performed to evaluate survival differences between PET periods and PET use, controlling for age, sex, race, disease stage, urban residence, and treatment received. Table 3 provides estimates of adjusted HRs from the model. The PET period was found to be a significant predictor for better survival, with a reduced HR for death (0.95; 95% CI, 0.91-0.99) \((P = .02)\). There was a significant difference in survival for patients with stage III and IV disease in the pre-PET period vs the PET period \((P < .001)\) (Figure 2). In addition, PET use was significantly associated with a reduced HR for death in stage III (0.77; 95% CI, 0.69-0.85) and IV (0.64; 95% CI, 0.58-0.70) disease, but not for stage I and II disease. Figures 3 and 4 provide Kaplan-Meier estimates for patients who underwent PET vs those who did not undergo PET for each disease stage during the combined pre-PET and PET periods as well as the PET period alone. Patients with stage III and IV disease who received PET had improved survival rates (Figures 3 and 4) \((P < .001)\). Similar results were obtained for the pre-PET period alone (results not shown).

**ADJUSTED SURVIVAL BASED ON PS**

Propensity score matching is an alternative to the Cox multivariate adjustment for confounders. Figure 5 shows Kaplan-Meier survival curves for disease stages I through IV based on PS-matched pairs (PET-yes and PET-no groups) for each stage. A significant improvement in survival was observed in the PET-yes group for disease stages III and IV \((P < .001)\). For patients with stage III disease, receipt of surgical, chemotherapy, and radiation treatment was significantly different between the PET-yes and PET-no groups before PS matching. Furthermore, PS matching eliminated differences in receipt of surgical and radiation treatment. Survival differences because of PET use, stratified by chemotherapy treatment (using the PS-matched pairs), were also evaluated. Improved survival was again observed for patients who did not receive chemotherapy and were in the PET-yes group \((P < .001)\), as well as in those who received chemotherapy \((P = .002)\). For patients with stage IV disease, PS matching eliminated differences between surgical treatment and chemotherapy use, although there were some differences with respect to radiation use between the PET-yes and PET-no groups. Similarly, improved survival owing to PET use was observed among those who received radiation \((P < .001)\) and those who did not \((P < .001)\). Additional covariate/confounder adjustment was performed by including PS with covariates that could not be balanced (ie, chemotherapy use in stage III and radiation use in stage IV) in a survival analysis using the Cox regression model for each stage. As in the conventional multivariate modeling, results based on PS also indicate that PET use was strongly associated with a reduced risk of death among patients with stage III \((HR, 0.72; 95\% CI, 0.65-0.81)\) \((P < .001)\) and IV \((HR, 0.65; 95\% CI, 0.59-0.71)\) \((P < .001)\) disease, although not among patients with stage I \((HR, 0.95; 95\% CI, 0.83-1.09)\) \((P = .45)\) and II \((HR, 0.84; 95\% CI, 0.66-1.07)\) \((P = .16)\) disease. Furthermore, these HR estimates from the Cox regression models with PS added as a covariate for each stage are very similar to estimates based on the conventional Cox multivariate analysis using data from all stages.

**Table 3. Clinical Variables Predictive of Survival: Cox Multivariate Regression Analysis**

<table>
<thead>
<tr>
<th>Clinical Variable</th>
<th>Hazard Ratio (95% CI)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis</td>
<td>1.00 (1.00-1.01)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Male sex</td>
<td>1.21 (1.17-1.26)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Disease stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>1.83 (1.61-2.09)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>II</td>
<td>2.52 (2.33-2.72)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>III</td>
<td>3.80 (3.51-4.12)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgical treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lobectomy/partial pneumonectomy</td>
<td>0.29 (0.27-0.31)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Pneumonectomy</td>
<td>0.34 (0.29-0.39)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Lung procedure, not otherwise specified</td>
<td>0.41 (0.29-0.58)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Chemotherapy received</td>
<td>0.54 (0.52-0.56)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Radiation received</td>
<td>0.78 (0.74-0.81)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Post-PET period</td>
<td>0.95 (0.91-0.99)</td>
<td>.02</td>
</tr>
<tr>
<td>PET use by disease stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>0.89 (0.78-1.02)</td>
<td>.10</td>
</tr>
<tr>
<td>II</td>
<td>0.82 (0.65-1.04)</td>
<td>.10</td>
</tr>
<tr>
<td>III</td>
<td>0.77 (0.69-0.85)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>IV</td>
<td>0.64 (0.58-0.70)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

**COMMENT**

Lung cancer will be diagnosed in more than 215 000 Americans in 2008 and remains the leading cause of cancer-related death among men and women. The most common subtype is NSCLC, which accounts for approximately 85% of all cases. Unfortunately, most patients are diagnosed as having locally advanced (stage III) or metastatic (stage IV) disease. In patients with stage IV NSCLC, treatment goals are palliative. Platinum-based chemotherapy, which modestly improves survival and quality of life in these patients, is often offered as the primary treatment.

From 1994 to 2004, no significant improvement in survival was seen among patients with stage IV NSCLC following a series of large randomized controlled clinical trials. It was not until 2005 when a novel agent (bevacizumab) was found to improve survival outcomes when added to a platinum-based chemotherapy backbone. It is interesting to note that, from 1994 to 2004, clinical trials testing platinum-based combinations have reported numerical improvements in median survival time even though none of these trials showed superiority of either arm. For example, the median sur-
vival time of patients with NSCLC treated with the control regimen of carboplatin and paclitaxel in phase III trials ranged from 8 to 9 months in the mid- to late-1990s, later rising to 9.9 to 10.5 months in the early part of the 21st century. We wondered whether this change was in part owing to the Will Rogers' phenomenon, brought about by increased use of PET scanning.

In a population-based registry study of all NSCLC cases diagnosed in the Sacramento region, we found that the PET scan era (1999-2004) was strongly associated with improved survival compared with the pre-PET era (1994-1998). We also found that PET scan usage was an independent predictor of improved survival, even after adjustment for potentially confounding variables such as sex, age, and treatment received. It appears that many more patients are being diagnosed as having stage IV disease, with a concomitant decrease in the proportion of patients with stage III disease. It is likely that FDG-PET scan use has “up-staged” patients who might otherwise be diagnosed as having stage III disease to stage IV by detecting distant metastatic disease not visualized by other routine imaging studies. It is not surprising that it is the stage III population that is most affected by stage migration (into stage IV) because of PET scans. After all, patients with stage III disease, who have bulkier primary tumors or mediastinal nodal involvement, have a much higher chance of occult distant metastases than those with stage I or II disease. These metastatic deposits are more readily detected by PET scans than by traditional imaging studies. It is also more likely that patients with stage III disease will have distant metastases that are larger than 10 mm in diameter, a threshold size that increases the likelihood of detection with PET scanning. In contrast, earlier stage tumors are more likely to harbor smaller, truly occult metastases below the limits of detection on PET scans. Nevertheless, stage migration could also have occurred in patients with stage I and II disease but may not have

Figure 3. Kaplan-Meier survival curves by positron emission tomography (PET) use (PET-yes and PET-no groups) based on all data (1994-2004) for patients with stage I (A), stage II (B), stage III (C), and stage IV non–small cell lung cancer (D).
been detected by this analysis because of relatively small cohort sizes.

Limitations of our study include its regional nature (mitigated by its relative completeness), lack of detailed treatment information, and lack of prognostic information, such as weight loss and functional status. In addition, there is a bias as to which patients were selected to receive a PET scan. Patients who were not considered for further PET testing may have had a poor functional status or comorbidities, conditions that could have affected survival outcomes. The strengths of this analysis include its large sample size, representing all lung cancer cases diagnosed within the region from 1994 to 2004 and its ethnic diversity, enhancing its potential to be generalizable to the entire population. Our Cox model provides external validation for the data set since the clinical variables analyzed were consistent with known prognostic variables in NSCLC (eg, stage and treatment). Finally, our use of a stratified analysis as well as matching based on PS further supported the results of the Cox multivariate model to account for potential confounders.

Our findings demonstrate that the Will Rogers phenomenon, originally described in relation to cancer staging in 1985, remains an important issue to consider when interpreting the results of clinical studies, particularly uncontrolled single-arm trials. Most important, this phenomenon is not limited to oncology and has been described in other biomedical fields. For example, more sensitive biomarkers of myocardial injury, such as troponin levels, may redefine patients with myocardial infarction and subsequently influence outcomes for this subset.27 The phenomenon also has been described in the context of health care economics. Sicker patients were found to preferentially migrate from indemnity to managed care plans, resulting in apparent shifts in resource use.28

In conclusion, our data support the hypothesis that enhanced survival for patients with stage III and IV NSCLC

Figure 4. Kaplan-Meier survival curves by positron emission tomography (PET) use (PET-yes and PET-no groups) based on 1999 to 2004 data for patients with stage I (A), stage II (B), stage III (C), and stage IV non–small cell lung cancer (D).
in the PET scan era is in part because of stage migration from III to IV. This migration is caused by the increased sensitivity of the PET scan and is not necessarily due solely to enhancements in oncological care.

Accepted for Publication: January 24, 2008.

Correspondence: Primo N. Lara Jr, MD, UC Davis Cancer Center, 4501 X St, Sacramento, CA 95817 (primo.lara@ucdmc.ucdavis.edu).

Author Contributions: Drs Chee, Wun, and Lara had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Chee, Wun, and Lara. Acquisition of data: Chee, Nguyen, and Brown. Analysis and interpretation of data: Nguyen, Brown, Gandara, Wun, and Lara. Drafting of the manuscript: Chee, Nguyen, Wun, and Lara. Critical revision of the manuscript for important intellectual content: Nguyen, Brown, Gandara, and Wun. Statistical analysis: Nguyen. Obtained funding: Lara. Administrative, technical, and material support: Chee, Brown, and Lara. Study supervision: Gandara, Wun, and Lara.

Financial Disclosure: None reported.

Funding/Support: This study was supported by grant UL1 RR024146 from the National Center for Research Resources. The collection of cancer incidence data used in this study was supported by the California Department of Health Services as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; contracts N01-PC-35136 (Northern California Cancer Center), N01-PC-35139 (University of Southern California), and N02-PC-15105 (Public Health Institute) from the Surveillance, Epidemiology and End Results Program, National Cancer Institute; and agreement U55/CCR921930-02 (Public Health Institute) from the National Program of Cancer Registries, Centers for Disease Control and Prevention.

Figure 5. Kaplan-Meier survival curves by positron emission tomography (PET) use (PET-yes vs PET-no groups; matched pairs) based on propensity score matching for patients with stage I (541 matched pairs) (A), stage II (134 matched pairs) (B), stage III (496 matched pairs) (C), and stage IV (549 matched pairs) (D) non–small cell lung cancer.