



## Quality Gaps and Comparative Effectiveness in Lung Cancer Staging and Diagnosis

David E. Ost, MD, MPH, FCCP; Jianguo Niu, PhD; Linda S. Elting, DrPH; Thomas A. Buchholz, MD; and Sharon H. Giordano, MD

**Background:** Guidelines recommend mediastinal lymph node sampling as the first invasive test in patients with suspected lung cancer with mediastinal lymphadenopathy without distant metastases, but there are no comparative effectiveness studies on how test sequencing affects outcomes. The objective was to compare practice patterns and outcomes of diagnostic strategies in patients with lung cancer.

**Methods:** The study included a retrospective cohort of 15,316 patients with lung cancer with regional spread without distant metastases in the Surveillance, Epidemiology, and End Results or Texas Cancer Registry Medicare-linked databases. If the first invasive test involved mediastinal sampling, patients were classified as receiving guideline-consistent care; otherwise, they were classified as receiving guideline-inconsistent care. We used propensity matching to compare the number of tests performed and multivariate logistic regression to compare the frequency of complications.

**Results:** Twenty-one percent of patients had guideline-consistent diagnostic evaluations. Among patients with non-small cell lung cancer, 44% never had mediastinal sampling. Patients who had guideline-consistent care required fewer tests than those with guideline-inconsistent care ( $P < .0001$ ), including thoracotomies (49% vs 80%,  $P < .001$ ) and CT image-guided biopsies (9% vs 63%,  $P < .001$ ), although they had more transbronchial needle aspirations (37% vs 4%,  $P < .001$ ). The consequence was that patients with guideline-consistent care had fewer pneumothoraxes (4.8% vs 25.6%,  $P < .0001$ ), chest tubes (0.7% vs 4.9%,  $P < .001$ ), hemorrhages (5.4% vs 10.6%,  $P < .001$ ), and respiratory failure events (5.3% vs 10.5%,  $P < .001$ ).

**Conclusions:** Guideline-consistent care with mediastinal sampling first resulted in fewer tests and complications. We found three quality gaps: failure to sample the mediastinum first, failure to sample the mediastinum at all in patients with non-small cell lung cancer, and overuse of thoracotomy.

*CHEST* 2014; 145(2):331–345

**Abbreviations:** EBUS = endobronchial ultrasound; NSCLC = non-small cell lung cancer; SEER = Surveillance, Epidemiology, and End Results; TBNA = transbronchial needle aspiration; TCR = Texas Cancer Registry

In patients with suspected lung cancer without distant metastases, assessment of the mediastinal lymph nodes is important because the status of the lymph nodes will help the physician to determine whether the disease is surgically resectable.<sup>1</sup> Because of the limited accuracy of both CT and PET scanning, current evidence-based guidelines recommend that patients with mediastinal adenopathy by CT or PET scan undergo lymph node sampling to ensure accurate staging.<sup>1–4</sup>

However, significant discordance may exist between what is recommended in evidence-based guidelines and what is actually done in practice. Previous studies

of patients with non-small cell lung cancer (NSCLC) found that mediastinoscopy is infrequently performed, and even then, lymph nodes are biopsied in < 50% of cases.<sup>5,6</sup> Alternative methods of mediastinal lymph node sampling, such as transbronchial needle aspiration (TBNA), have also been underused partly because of inadequate fellowship training.<sup>7–10</sup>

Although these studies demonstrate that mediastinal sampling techniques have been underused, an equally important question is how mediastinal sampling techniques are used in practice. Multiple evidence-based guidelines recommend mediastinal lymph node sampling as the first invasive diagnostic procedure in patients

with suspected lung cancer with mediastinal adenopathy without distant metastases because the procedure can be used for both diagnosis and staging.<sup>2-4,11-16</sup> However, to our knowledge, only one single-center comparative effectiveness study has evaluated how test sequencing affects outcomes.<sup>17</sup>

The goal of the present study was to compare practice patterns and outcomes of diagnostic and staging strategies in patients with lung cancer with mediastinal lymph node involvement without distant metastasis. We hypothesized that peripheral lung mass biopsy often occurs prior to sampling of the mediastinal lymph nodes, contrary to guidelines. We further hypothesized guideline-inconsistent care would result in unnecessary procedures and more complications.

## MATERIALS AND METHODS

### Data Source

We performed a retrospective cohort analysis of two datasets: the National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) database and the Texas Cancer Registry (TCR). The registry data have been linked to Medicare claims and 2000 US Census data. We compared the registries and analyzed practice patterns and outcomes. This study was approved by institutional review board 4, and a waiver of informed consent was obtained.

### Study Participants

The cohort comprised patients with lung cancer with regional spread to the hilar or mediastinal lymph nodes without distant metastases. The algorithms and search results are shown in Figure 1 (see e-Table 1 for additional details). For patients entered into SEER prior to 2004 and for all patients in the TCR, American

Joint Committee on Cancer nodal staging was not recorded; therefore, it was not possible to further stratify patients into N1 vs N2 vs N3 status. For patients in SEER from 2004 or later, precise TNM staging could be obtained.

### Diagnostic and Staging Strategy

The type and sequencing of invasive tests used for diagnosis and staging were determined by Current Procedural Terminology and *International Classification of Diseases, Ninth Edition*, codes (e-Table 1). Invasive tests were defined as CT image-guided needle biopsy, bronchoscopy, endoscopy with ultrasound-guided needle aspiration, mediastinoscopy, or thoracotomy. Only tests done within the 6 months preceding the initiation of treatment were considered. Patients were placed into groups based on their diagnostic testing sequence: (1) evaluation consistent with guidelines, some form of mediastinal sampling done first; (2) evaluation inconsistent with guidelines, NSCLC present, mediastinal sampling performed on the second or later biopsy; (3) evaluation inconsistent with guidelines, NSCLC present, mediastinal sampling never done; and (4) evaluation inconsistent with guidelines, small cell lung cancer. Mediastinal sampling procedures were defined as bronchoscopy with TBNA or endobronchial ultrasound (EBUS)-guided TBNA, endoscopy with ultrasound-guided needle aspiration, mediastinoscopy, thoracoscopy, or thoracotomy with mediastinal lymph node sampling (see e-Appendix 1 for details on categories and criteria).

### Outcomes

The primary outcome was whether the evaluation strategy was consistent with guidelines. Secondary outcomes were whether mediastinal lymph node sampling was ever performed prior to treatment in patients with NSCLC, complications related to the diagnostic evaluation, and the number of invasive diagnostic tests performed. We used a methodology similar to that previously published to identify complications, including pneumothorax, hemorrhage, and respiratory failure.<sup>18</sup> For thoracotomy, any hemorrhage or respiratory failure occurring within 14 days of surgery was considered a complication. For all other procedures, complications were only counted if they occurred up to 1 day after the procedure. We conducted a subset analysis of patients in SEER from 2004 and later to assess the impact of T and N stage on practice patterns. We also conducted an exploratory analysis to assess the relationship among diagnostic practice patterns, subsequent treatment modalities used, and survival.

### Statistical Analysis

Characteristics of patients and outcomes were compared using  $\chi^2$  test for categorical variables; *t* tests for continuous, normally distributed variables; and Wilcoxon rank sum test for nonnormally distributed variables. We used multivariate logistic regression to analyze factors associated with complications due to diagnostic testing. We decided a priori that variables significantly associated with outcomes at the 0.2 level in univariate analysis would be considered candidate variables for multivariate analysis. Backward selection was used to retain only variables with a level of significance  $< .05$ . The number of invasive tests performed was not normally distributed, so we used propensity scores to match patients who had guideline-consistent care with mediastinal sampling first with counterparts who had mediastinal sampling performed second or later. The conditional probability to have guideline-consistent care was estimated by logistic regression analysis incorporating the following variables: age, sex, race, year of diagnosis, Charlson comorbidity index, T stage, geographic region, and cancer type. All statistical analyses were performed at a significance level of .05. All

Manuscript received July 11, 2013; revision accepted September 9, 2013.

**Affiliations:** From the Department of Pulmonary Medicine (Dr Ost), Department of Health Services Research (Drs Niu, Elting, and Giordano), and Department of Radiation Oncology (Dr Buchholz), The University of Texas MD Anderson Cancer Center, Houston, TX.

Portions of these data were presented at the American Thoracic Society International Conference, May 17-22, 2013, Philadelphia, PA.

**Funding/Support:** This work was supported in part by Comparative Effectiveness Research on Cancer in Texas, a multiuniversity consortium funded by the Cancer Prevention and Research Institute of Texas [Grant RP101207 by 2P30 CA016672]. Dr Giordano is also supported by a grant from the American Cancer Society [RSG-09-149-01-CPHPS]. The collection of cancer incidence used in this study was supported by the Texas Department of State Health Services and Cancer Prevention Research Institute of Texas, as part of the statewide cancer reporting program, and the Centers for Disease Control and Prevention National Program of Cancer Registries Cooperative Agreement #5U58/DP000824-05.

**Correspondence to:** David E. Ost, MD, MPH, FCCP, Department of Pulmonary Medicine, Unit 1462, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030; e-mail: dost@mdanderson.org

© 2014 American College of Chest Physicians. Reproduction of this article is prohibited without written permission from the American College of Chest Physicians. See online for more details.

DOI: 10.1378/chest.13-1599

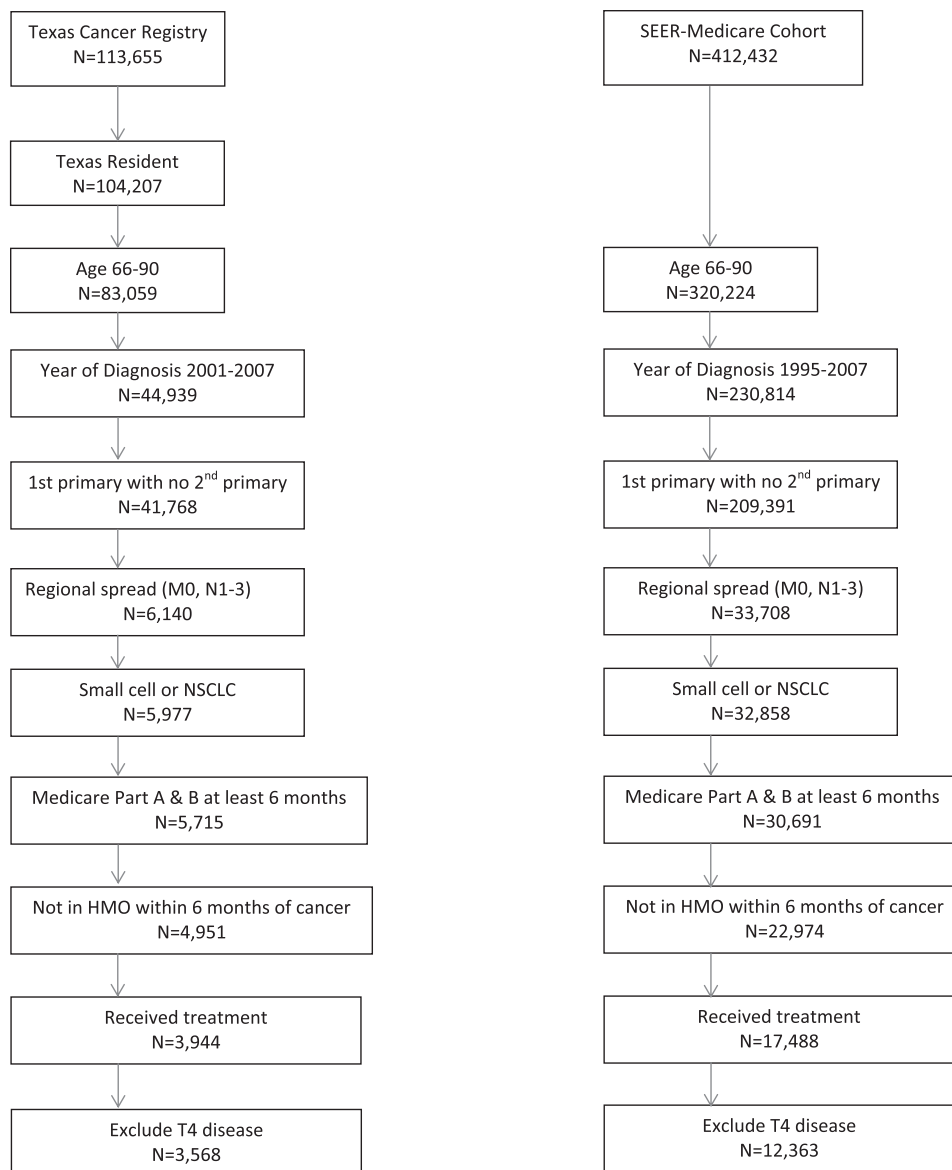


FIGURE 1. Study cohort selection results: SEER and Texas Cancer Registry 1995 to 2007. HMO = health maintenance organization; NSCLC = non-small cell lung cancer; SEER = Surveillance, Epidemiology, and End Results.

data were analyzed with SAS version 9.2 (SAS Institute Inc) statistical software.

## RESULTS

### *SEER and TCR Medicare Cohorts*

A total of 12,363 patients from the SEER and 3,658 patients from the TCR Medicare datasets met the inclusion criteria (Fig 1). We compared the characteristics of the patients in SEER and TCR, practice patterns, and lung cancer types (e-Table 2). Because of the large sample size, *P* values were significant, but there was little absolute difference between groups.

For subsequent analysis, therefore, we combined the two registries and controlled for geographic region.

Patient characteristics for the combined cohort are shown in Table 1. Of the 15,931 eligible patients, 615 (4%) had no Medicare data, indicating that any diagnostic testing was performed. The remaining 15,316 patients (96%) had Medicare data, and this group comprised the final study cohort.

### *Practice Patterns and Consistency With Guidelines*

Only 21% of patients had an evaluation consistent with guidelines, with mediastinal sampling done first (Fig 2). Of all patients with NSCLC, 44% never had

**Table 1—Patient Characteristics**

| Variable                   | Evaluation Consistent With Guidelines, Mediastinal Sampling Done First | Evaluation Inconsistent With Guidelines             |  |                                | No Evaluation Recorded <sup>a</sup> | Total          |
|----------------------------|--|---|--|--------------------------------|-------------------------------------|----------------|
|                            |  | NSCLC Present, Mediastinal Sampling Second or Later | NSCLC Present, Mediastinal Sampling Never Done | Small Cell Lung Cancer Present |                                     |                |
| All patients               | 3,155 (100)  | 4,659 (100)   | 5,881 (100)                                    | 1,621 (100)                    | 615 (100)                           | 15,931 (100)   |
| Age, y                     |  |   |  |                                |                                     |                |
| 66-70                      | 1,074 (34.04)  | 1,527 (32.78)                                       | 1,549 (26.34)                                  | 571 (35.23)                    | 196 (31.87)                         | 4,917 (30.86)  |
| 71-75                      | 995 (31.54)  | 1,566 (33.61)                                       | 1,697 (28.86)                                  | 486 (29.98)                    | 169 (27.48)                         | 4,913 (30.84)  |
| 76-80                      | 761 (24.12)  | 1,046 (22.45)                                       | 1,521 (25.86)                                  | 360 (22.21)                    | 143 (23.25)                         | 3,831 (24.05)  |
| > 80                       | 325 (10.3)   | 520 (11.16)   | 1,114 (18.94)                                  | 204 (12.59)                    | 107 (17.4)                          | 2,270 (14.25)  |
| Sex                        |  |   |  |                                |                                     |                |
| Female                     | 1,530 (48.49)  | 2,146 (46.06)                                       | 2,637 (44.84)                                  | 854 (52.68)                    | 254 (41.3)                          | 7,421 (46.58)  |
| Male                       | 1,625 (51.51)  | 2,513 (53.94)                                       | 3,244 (55.16)                                  | 767 (47.32)                    | 361 (58.7)                          | 8,510 (53.42)  |
| Race                       |  |   |  |                                |                                     |                |
| Non-Hispanic white         | 2,745 (87.01)  | 3,989 (85.62)                                       | 4,826 (82.06)                                  | 1,409 (86.92)                  | 490 (79.68)                         | 13,459 (84.48) |
| Hispanic                   | 129 (4.09)   | 223 (4.79)  | 289 (4.91)                                     | 70 (4.32)                      | 34 (5.53)                           | 745 (4.68)     |
| Non-Hispanic black         | 199 (6.31)   | 247 (5.3)   | 548 (9.32)                                     | 99 (6.11)                      | 66 (10.73)                          | 1,159 (7.28)   |
| Non-Hispanic other         | 82 (2.6)   | 200 (4.29)  | 218 (3.71)                                     | 43 (2.65)                      | 25 (4.07)                           | 568 (3.57)     |
| Urban/rural                |  |   |  |                                |                                     |                |
| Big metropolitan           | 1,763 (55.88)  | 2,557 (54.88)                                       | 3,092 (52.58)                                  | 799 (49.29)                    | 320 (52.03)                         | 8,531 (53.55)  |
| Metropolitan               | 861 (27.29)  | 1,266 (27.17)                                       | 1,625 (27.63)                                  | 481 (29.67)                    | 181 (29.43)                         | 4,414 (27.71)  |
| Urban                      | 196 (6.21)   | 306 (6.57)  | 431 (7.33)                                     | 132 (8.14)                     | 39 (6.34)                           | 1,104 (6.93)   |
| Less urban                 | 279 (8.84)   | 433 (9.29)  | 611 (10.39)                                    | 171 (10.55)                    | 64 (10.41)                          | 1,558 (9.78)   |
| Rural                      | 56 (1.78)  | 97 (2.08)   | 122 (2.07)                                     | 38 (2.34)                      | 11 (1.79)                           | 324 (2.03)     |
| Year of diagnosis          |  |   |  |                                |                                     |                |
| 1995                       | 105 (3.33)   | 181 (3.89)  | 135 (2.3)                                      | 45 (2.78)                      | 21 (3.42)                           | 487 (3.06)     |
| 1996                       | 88 (2.79)  | 183 (3.93)  | 150 (2.55)                                     | 58 (3.58)                      | 21 (3.42)                           | 500 (3.14)     |
| 1997                       | 92 (2.92)  | 178 (3.82)  | 131 (2.23)                                     | 51 (3.15)                      | 18 (2.93)                           | 470 (2.95)     |
| 1998                       | 89 (2.82)  | 165 (3.54)  | 152 (2.59)                                     | 38 (2.34)                      | 15 (2.44)                           | 459 (2.88)     |
| 1999                       | 86 (2.73)  | 131 (2.81)  | 107 (1.82)                                     | 48 (2.96)                      | 24 (3.9)                            | 396 (2.49)     |
| 2000                       | 207 (6.56)   | 349 (7.49)  | 338 (5.75)                                     | 107 (6.6)                      | 40 (6.5)                            | 1,041 (6.53)   |
| 2001                       | 335 (10.62)  | 483 (10.37)   | 636 (10.81)                                    | 184 (11.35)                    | 63 (10.24)                          | 1,701 (10.68)  |
| 2002                       | 302 (9.57)   | 516 (11.08)   | 614 (10.44)                                    | 191 (11.78)                    | 57 (9.27)                           | 1,680 (10.55)  |
| 2003                       | 377 (11.95)  | 501 (10.75)   | 716 (12.18)                                    | 189 (11.66)                    | 69 (11.22)                          | 1,852 (11.63)  |
| 2004                       | 401 (12.71)  | 497 (10.67)   | 746 (12.69)                                    | 197 (12.15)                    | 82 (13.33)                          | 1,923 (12.07)  |
| 2005                       | 344 (10.9)   | 540 (11.59)   | 742 (12.62)                                    | 201 (12.4)                     | 65 (10.57)                          | 1,892 (11.88)  |
| 2006                       | 357 (11.32)  | 484 (10.39)   | 712 (12.11)                                    | 155 (9.56)                     | 65 (10.57)                          | 1,773 (11.13)  |
| 2007                       | 372 (11.79)  | 451 (9.68)  | 702 (11.94)                                    | 157 (9.69)                     | 75 (12.2)                           | 1,757 (11.03)  |
| SEER/TCR region            |  |   |  |                                |                                     |                |
| Atlanta and rural Georgia  | 68 (2.16)  | 164 (3.52)  | 218 (3.71)                                     | 44 (2.71)                      | 14 (2.28)                           | 508 (3.19)     |
| California                 | 640 (20.29)  | 1,011 (21.7)  | 1,122 (19.08)                                  | 309 (19.06)                    | 134 (21.79)                         | 3,216 (20.19)  |
| Connecticut                | 239 (7.58)   | 392 (8.41)  | 327 (5.56)                                     | 108 (6.66)                     | 53 (8.62)                           | 1,119 (7.02)   |
| Detroit                    | 353 (11.19)  | 387 (8.31)  | 477 (8.11)                                     | 115 (7.09)                     | 42 (6.83)                           | 1,374 (8.63)   |
| Hawaii                     | 24 (0.76)  | 68 (1.46)   | 89 (1.51)                                      | 13 (0.8)                       | 15 (2.44)                           | 209 (1.31)     |
| Iowa                       | 175 (5.55)   | 337 (7.23)  | 408 (6.94)                                     | 146 (9.01)                     | 33 (5.37)                           | 1,099 (6.9)    |
| Kentucky                   | 209 (6.62)   | 353 (7.58)  | 475 (8.08)                                     | 171 (10.55)                    | 37 (6.02)                           | 1,245 (7.82)   |
| Louisiana                  | 98 (3.11)  | 215 (4.62)  | 454 (7.72)                                     | 106 (6.54)                     | 36 (5.85)                           | 909 (5.71)     |
| New Jersey                 | 395 (12.52)  | 462 (9.92)  | 576 (9.79)                                     | 144 (8.88)                     | 64 (10.41)                          | 1,641 (10.3)   |
| New Mexico                 | 49 (1.55)  | 82 (1.76)   | 80 (1.36)                                      | 28 (1.73)                      | < 11 (< 2) <sup>b</sup>             | < 250 (< 1.56) |
| Seattle                    | 146 (4.63)   | 198 (4.25)  | 217 (3.69)                                     | 45 (2.78)                      | 38 (6.18)                           | 644 (4.04)     |
| Texas                      | 720 (22.82)  | 946 (20.31)   | 1,393 (23.69)                                  | 374 (23.07)                    | 135 (21.95)                         | 3,568 (22.4)   |
| Utah                       | 39 (1.24)  | 44 (0.94)   | 45 (0.77)                                      | 18 (1.11)                      | < 11 (< 1) <sup>b</sup>             | < 157 (< 0.99) |
| Charlson comorbidity index |  |   |  |                                |                                     |                |
| 0                          | 1,620 (51.35)  | 2,563 (55.01)                                       | 2,740 (46.59)                                  | 784 (48.37)                    | 332 (53.98)                         | 8,039 (50.46)  |
| 1                          | 959 (30.4)   | 1,340 (28.76)                                       | 1,800 (30.61)                                  | 498 (30.72)                    | 148 (24.07)                         | 4,745 (29.79)  |
| 2+                         | 576 (18.26)  | 756 (16.23)   | 1,341 (22.8)                                   | 339 (20.91)                    | 135 (21.95)                         | 3,147 (19.75)  |

(Continued)

Table 1—Continued

| Variable   | Evaluation Consistent With Guidelines, Mediastinal Sampling Done First | Evaluation Inconsistent With Guidelines             |  |                                | No Evaluation Recorded <sup>a</sup> | Total          |
|--|--|---|--|--------------------------------|-------------------------------------|----------------|
|  |  | NSCLC Present, Mediastinal Sampling Second or Later | NSCLC Present, Mediastinal Sampling Never Done | Small Cell Lung Cancer Present |                                     |                |
| Poverty in patient's census tract <sup>c</sup>       |  |   |  |                                |                                     |                |
| ≤ 4.76%  | 886 (28.08)  | 1,183 (25.39)                                       | 1,316 (22.38)                                  | 369 (22.76)                    | 145 (23.58)                         | 3,899 (24.47)  |
| 4.77%-9.07%  | 742 (23.52)  | 1,194 (25.63)                                       | 1,401 (23.82)                                  | 390 (24.06)                    | 136 (22.11)                         | 3,863 (24.25)  |
| 9.08%-16.53%   | 779 (24.69)  | 1,071 (22.99)                                       | 1,461 (24.84)                                  | 417 (25.73)                    | 155 (25.2)                          | 3,883 (24.37)  |
| > 16.53%   | 619 (19.62)  | 997 (21.4)  | 1,544 (26.25)                                  | 391 (24.12)                    | 153 (24.88)                         | 3,704 (23.25)  |
| Unknown  | 129 (4.09)   | 214 (4.59)  | 159 (2.7)                                      | 54 (3.33)                      | 26 (4.23)                           | 582 (3.65)     |
| < 12 y education <sup>d</sup>                        |  |   |  |                                |                                     |                |
| ≤ 10.09%   | 864 (27.39)  | 1,239 (26.59)                                       | 1,330 (22.62)                                  | 336 (20.73)                    | 140 (22.76)                         | 3,909 (24.54)  |
| 10.1%-17.18%   | 767 (24.31)  | 1,157 (24.83)                                       | 1,392 (23.67)                                  | 413 (25.48)                    | 145 (23.58)                         | 3,874 (24.32)  |
| 17.19%-27.8%   | 765 (24.25)  | 1,040 (22.32)                                       | 1,471 (25.01)                                  | 417 (25.73)                    | 162 (26.34)                         | 3,855 (24.2)   |
| > 27.85%   | 630 (19.97)  | 1,009 (21.66)                                       | 1,529 (26)                                     | 401 (24.74)                    | 142 (23.09)                         | 3,711 (23.29)  |
| Unknown  | 129 (4.09)   | 214 (4.59)  | 159 (2.7)                                      | 54 (3.33)                      | 26 (4.23)                           | 582 (3.65)     |
| Specialty of physician doing first test <sup>e</sup> |  |   |  |                                |                                     |                |
| Internal medicine                                    | 368 (11.66)  | 1,040 (22.32)                                       | 1,381 (23.48)                                  | 375 (23.13)                    | 0 (0)                               | 3,164 (19.86)  |
| Pulmonary  | 945 (29.95)  | 2,300 (49.37)                                       | 2,998 (50.98)                                  | 878 (54.16)                    | 0 (0)                               | 7,121 (44.70)  |
| General surgery                                      | 298 (9.45)   | 196 (4.21)  | 244 (4.15)                                     | 74 (4.57)                      | 0 (0)                               | 812 (5.10)     |
| Thoracic surgery                                     | 1,431 (45.36)  | 304 (6.53)  | 292 (4.97)                                     | 80 (4.94)                      | 0 (0)                               | 2,107 (13.23)  |
| Other  | 19 (0.6)   | 141 (3.03)  | 236 (4.01)                                     | 46 (2.84)                      | 0 (0)                               | 442 (2.77)     |
| Unknown  | 94 (2.98)  | 678 (14.55)   | 730 (12.41)                                    | 168 (10.36)                    | 615 (100)                           | 2,285 (14.34)  |
| T stage  |  |   |  |                                |                                     |                |
| T1 <sup>b</sup>                                      | 908 (28.78)  | 1,105 (23.72)                                       | 869 (14.77)                                    | 322 (19.87)                    | 133 (21.63)                         | 3,337 (20.94)  |
| T2   | 1,342 (42.54)  | 2,633 (56.51)                                       | 2,993 (50.89)                                  | 728 (44.91)                    | 243 (39.51)                         | 7,939 (49.83)  |
| T3   | 176 (5.58)   | 279 (5.99)  | 762 (12.96)                                    | 117 (7.22)                     | 74 (12.03)                          | 1,408 (8.84)   |
| Unknown  | 729 (23.11)  | 642 (13.78)   | 1,257 (21.37)                                  | 454 (28.01)                    | 165 (26.83)                         | 3,247 (20.38)  |
| Cancer type  |  |   |  |                                |                                     |                |
| NSCLC  | 2,680 (84.95)  | 4,659 (100)   | 5,881 (100)                                    | 0 (0)                          | 500 (81.3)                          | 13,720 (86.12) |
| Small cell lung cancer                               | 475 (15.06)  | 0 (0)   | 0 (0)  | 1,621 (100)                    | 115 (18.7)                          | 2,211 (13.88)  |

Data are presented as No. (%). NSCLC = non-small cell lung cancer; SEER = Surveillance, Epidemiology, and End Results; TCR = Texas Cancer Registry.

<sup>a</sup>No evaluation recorded means that no Medicare payments were noted. For example, a patient who had their care delivered through the Veteran's Administration would not show up in the Medicare dataset.

<sup>b</sup>Strata with ≤ 10 patients were suppressed per National Cancer Institute policy and are reported as < 11 to ensure confidentiality. A total of 50 patients with T0 disease have been included in the T1 category to maintain confidentiality because there were too few patients with T0 disease to report them separately while maintaining confidentiality.

<sup>c</sup>Poverty is the percentage of the population in the patient's census tract living below the poverty level. Note that this does not mean that the patient's income is below the poverty line, just that the patient is living in a census tract with that level of poverty.

<sup>d</sup>Percentage of patients with < 12 y education is the percentage of the population in their census tract who did not graduate from high school.

<sup>e</sup>Refers to the physician who ordered or performed the first invasive diagnostic test. For bronchoscopy and surgical procedures, this was the physician performing the procedure. For CT image-guided biopsy, this was the referring physician. Internal medicine includes family practice and all subspecialties of internal medicine other than oncology and pulmonary medicine. Surgery includes all other subspecialties of surgery other than thoracic and cardiothoracic. Thoracic surgery and cardiothoracic surgery are included under thoracic surgery.

mediastinal sampling prior to treatment. The most common first invasive diagnostic test was bronchoscopy without TBNA followed by CT image-guided biopsy (Table 2).

### Complications

On a per-procedure basis, the incidence of complications was different between groups (Table 3), with the guideline-consistent care group having a lower incidence of pneumothorax following CT image-guided

biopsy and mediastinoscopy, less hemorrhage following bronchoscopy, and less respiratory failure following thoracotomy. On a per-patient basis, guideline-consistent care with mediastinal sampling as the first test resulted in fewer pneumothoraxes, chest tubes, hemorrhages, and episodes of respiratory failure than guideline-inconsistent care with mediastinal sampling performed in the second or later test ( $P < .001$ ) (Table 4). In multivariate analysis (Table 5), for the outcome of any of these complications on a per-patient basis, guideline-consistent care was associated with a

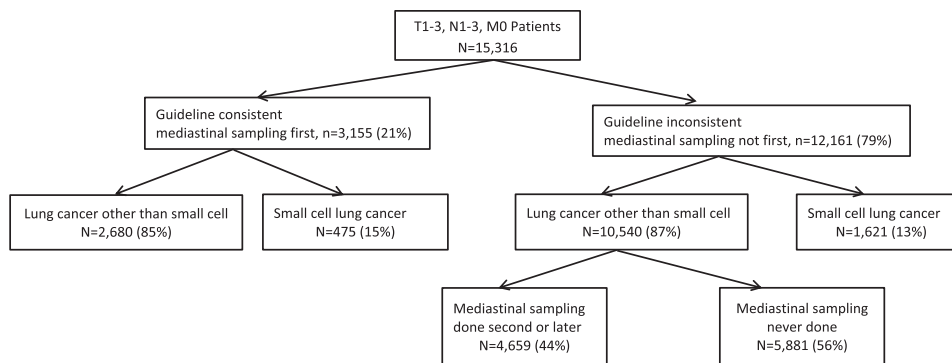


FIGURE 2. Practice patterns and diagnoses for the entire study cohort of SEER and Texas Cancer Registry patients from 1995 to 2007. See Figure 1 legend for expansion of abbreviation.

lower risk of complications (OR, 0.42; 95% CI, 0.37-0.48;  $P < .0001$ ).

Part of the difference between groups was due to the type of tests ordered. Patients receiving guideline-consistent care underwent fewer CT image-guided biopsies than those who had mediastinal sampling done second (9% vs 64%,  $P < .001$ ) (Table 3) and had fewer bronchoscopies without TBNA (5% vs 70%,  $P < .001$ ). As a result, far fewer pneumothoraxes occurred in the guideline-consistent group ( $P < .001$ ) (Table 4).

Patients receiving guideline-consistent care also had fewer episodes of hemorrhage and respiratory failure than those who had mediastinal sampling done in a second or later test ( $P < .001$ ) (Table 4) because patients with guideline-consistent care underwent fewer thoracotomies (49% vs 80%,  $P < .001$ ). Instead, patients in the guideline-consistent group had bronchoscopy with TBNA or mediastinoscopy. Patients who never had mediastinal lymph node sampling still had more pneumothoraxes than the guideline-consistent group, but they had fewer episodes of respiratory failure and hemorrhage because they never had surgery.

#### Number of Invasive Tests Performed

Another factor contributing to the incidence of complications on a per-patient basis was the number of invasive tests performed. Patients with guideline-

consistent care had fewer diagnostic tests ( $P < .001$ ) (Table 6) than those who had guideline-inconsistent care with mediastinal sampling done second. Propensity matching yielded 2,895 well-matched pairs of patients with lung cancer who did and did not receive guideline-consistent care, all of whom eventually had mediastinal sampling (e-Table 3). Compared with propensity-matched control subjects, patients with guideline-consistent care with mediastinal sampling first had fewer invasive diagnostic tests than similar patients with guideline-inconsistent care (median, 1 [interquartile range, 1-2] vs 2 [interquartile range, 2-3] respectively;  $P < .0001$ ). When we limited the analysis to just patients with NSCLC, the results were similar (Fig 3).

#### Subset Analysis Using SEER 2004 and Later Data With Precise T and N Staging

Practice patterns and consistency with guidelines for patients in SEER from 2004 to 2007 are shown in Figure 4. Only 19% of patients had guideline-consistent care with mediastinal sampling first. Only 52% of patients with NSCLC ever had mediastinal sampling prior to treatment. The frequency of guideline-consistent care varied according to stage, with patients with stage II disease having guideline-consistent care more frequently ( $P < .001$ ) (Table 7). Among patients with NSCLC who did not have mediastinal sampling

**Table 2—Initial Invasive Testing Procedures Used in Patients With Lung Cancer With Regional Spread**

| First Test  | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|-----------|---------|----------------------|--------------------|
| CT image-guided biopsy                            | 5,009     | 32.7    | 5,009                | 32.7               |
| Bronchoscopy without TBNA                         | 7,152     | 46.7    | 12,161               | 79.4               |
| Bronchoscopy with TBNA                            | 1,043     | 6.8     | 13,204               | 86.2               |
| Bronchoscopy with TBNA + EBUS or EUS <sup>a</sup> | 42        | 0.3     | 13,246               | 86.5               |
| Mediastinoscopy alone                             | 874       | 5.7     | 14,120               | 92.2               |
| Mediastinoscopy and thoracotomy                   | 181       | 1.2     | 14,301               | 93.4               |
| Thoracotomy alone                                 | 1,015     | 6.6     | 15,316               | 100                |

EBUS = endobronchial ultrasound; EUS = endoscopic ultrasound; TBNA = transbronchial needle aspiration.

<sup>a</sup>Strata with  $\leq 10$  patients were suppressed per National Cancer Institute policy and are reported as  $< 11$  to ensure confidentiality. EUS was done in  $< 11$  patients and, therefore, was included in the bronchoscopy with TBNA + EBUS category to protect patient confidentiality.

**Table 3—Procedure Use and Incidence of Complications Per Procedure**

| Invasive Diagnostic Test and Associated Complications | Evaluation Inconsistent With Guidelines                                |   |  |                                | P Value                            |   |
|---|--|---|--|--------------------------------|------------------------------------|---|
|   | Evaluation Consistent With Guidelines, Mediastinal Sampling Done First | NSCLC Present, Mediastinal Sampling Performed Second or Later | NSCLC Present, Mediastinal Sampling Never Done | Small Cell Lung Cancer Present | Frequency of Test Use <sup>a</sup> | Incidence of Complications <sup>b</sup> |
| Total patients <sup>c</sup>                           | 3,155  | 4,659   | 5,881  | 1,621                          | ...                                | ...                                     |
| Procedure   |  |   |  |                                |                                    |   |
| CT image-guided biopsy <sup>c</sup>                   | 283 (9)  | 2,961 (64)  | 3,635 (62)                                     | 743 (46)                       | < .001                             |   |
| Complications within 1 d                              |  |   |  |                                |                                    |   |
| Pneumothorax <sup>b</sup>                             | 68 (24)  | 857 (29)  | 843 (23)                                       | 179 (24)                       |                                    | < .001                                  |
| Pneumothorax requiring chest tube <sup>b</sup>        | 11 (3.9)   | 182 (6.1)   | 245 (6.7)                                      | 52 (7)                         |                                    | .22                                     |
| Hemorrhage <sup>b</sup>                               | < 11 (< 3.9) <sup>d</sup>  | 30 (1)  | 27 (0.7)                                       | < 11 (< 1.5) <sup>d</sup>      |                                    | .4                                      |
| Respiratory failure <sup>b</sup>                      | < 11 (< 3.9) <sup>d</sup>  |   |  | 0 (0)                          |                                    | .14                                     |
| Bronchoscopy without TBNA <sup>c</sup>                | 155 (5)  | 3,268 (70)  | 4,189 (71)                                     | 1,304 (80)                     | < .001                             |   |
| Complications within 1 d                              |  |   |  |                                |                                    |   |
| Pneumothorax  | < 11 (< 7.1) <sup>d</sup>  | 79 (2.4)  | 87 (2.1)                                       | 17 (1.3)                       |                                    | .09                                     |
| Pneumothorax requiring chest tube                     | 0 (0)  | 14 (0.4)  | 15 (0.4)                                       | < 11 (< 0.8) <sup>d</sup>      |                                    | .28                                     |
| Hemorrhage  | 0 (0)  | 39 (1.2)  | 39 (0.9)                                       | < 11 (< 0.8) <sup>d</sup>      |                                    | .02                                     |
| Respiratory failure                                   | 0 (0)  | 24 (0.7)  | 32 (0.8)                                       | < 11 (0.8) <sup>d</sup>        |                                    | .78                                     |
| Bronchoscopy with TBNA <sup>c</sup>                   | 1,150 (36)   | 183 (4)   | ...  | 31 (2)                         | < .001                             |   |
| Complications within 1 d                              |  |   |  |                                |                                    |   |
| Pneumothorax  | 26 (2.3)   | 4 (2.2)   | ...  | 0 (0)                          |                                    | 1                                       |
| Pneumothorax requiring chest tube                     | < 11 (< 0.9) <sup>d</sup>  | 0 (0)   | ...  | 0 (0)                          |                                    | 1                                       |
| Hemorrhage  | 13 (1.1)   | < 11 (< 6.0) <sup>d</sup>                                     | ...  | < 11 (< 35) <sup>d</sup>       |                                    | .32                                     |
| Respiratory failure                                   | < 11 (< 0.9) <sup>d</sup>  | 0 (0)   | ...  | 0 (0)                          |                                    | 1                                       |
| Mediastinoscopy alone (no thoracotomy) <sup>c</sup>   | 1,218 (39)   | 1,390 (30)  | ...  | 132 (8)                        | < .001                             |   |
| Complications within 14 d                             |  |   |  |                                |                                    |   |
| Pneumothorax  | 21 (1.7)   | 43 (3.1)  | ...  | 5 (3.8)                        |                                    | .04                                     |
| Pneumothorax requiring chest tube                     | < 11 (< 0.9)   | 6 (0.4)   | ...  | 0 (0)                          |                                    | .69                                     |
| Hemorrhage  | < 11 (< 0.9)   | 12 (0.9)  | ...  | < 11 (< 8.3)                   |                                    | .07                                     |
| Respiratory failure                                   | < 11 (< 0.9)   | 3 (0.2)   | ...  | < 11 (< 8.3)                   |                                    | .33                                     |
| Thoracotomy <sup>c</sup>                              | 1,630 (52)   | 3,903 (84)  | ...  | 93 (6)                         | < .001                             |   |
| Complications within 14 d                             |  |   |  |                                |                                    |   |
| Hemorrhage  | 146 (9)  | 408 (10.5)  | ...  | < 11 (< 12)                    |                                    | .14                                     |
| Respiratory failure                                   | 157 (9.6)  | 454 (11.6)  | ...  | 15 (16)                        |                                    | .03                                     |

Data are presented as No. (%) unless otherwise indicated. See Table 1 and 2 legends for expansion of abbreviations.

<sup>a</sup>P value compares the frequency of testing use between groups on a per-patient basis. For bronchoscopy with TBNA, mediastinoscopy alone, and thoracotomy, the comparison is only between guideline-consistent care and guideline-inconsistent care with sampling done second or later.

<sup>b</sup>P value compares incidence of complications on a per-procedure basis among guideline-consistent care, guideline-inconsistent care with sampling done second or later, and guideline-inconsistent care with sampling never done.

<sup>c</sup>Reflects percentage of patients in that group (guideline-consistent care, guideline-inconsistent care with NSCLC sampling done second or later, guideline-inconsistent care with NSCLC sampling never done, and guideline-inconsistent care with small cell lung cancer) who had the test done.

<sup>d</sup>Strata with ≤ 10 patients were suppressed per National Cancer Institute policy and are reported as < 11 to ensure patient confidentiality.

**Table 4—Cumulative Incidence of Complications During the Entire Diagnostic Evaluation Per Patient**

| Complication                      | Evaluation Consistent With Guidelines, Mediastinal Sampling Done First | Evaluation Inconsistent With Guidelines                       |  |                                | P Value <sup>a</sup> |
|-----------------------------------|--|---|--|--------------------------------|----------------------|
|                                   |  | NSCLC Present, Mediastinal Sampling Performed Second or Later | NSCLC Present, Mediastinal Sampling Never Done | Small Cell Lung Cancer Present |                      |
| No. patients                      | 3,155  | 4,659   | 5,881  | 1,621                          |                      |
| Pneumothorax                      | 3.8  | 21.1  | 15.8   | 12.4                           | < .001               |
| Pneumothorax requiring chest tube | 0.6  | 4.3   | 4.4  | 3.3                            | < .001               |
| Hemorrhage                        | 5.4  | 10.5  | 1.1  | 1.2                            | < .001 <sup>b</sup>  |
| Respiratory failure               | 5.2  | 10.5  | 0.6  | 1.4                            | < .001 <sup>b</sup>  |

Data are presented as percentages unless otherwise indicated. A single patient could have more than one complication. Percentages reflect the number of complications per 100 patients evaluated. See Table 1 legend for expansion of abbreviation.

<sup>a</sup>Guideline-consistent group vs other groups.

<sup>b</sup>Significant difference among the guideline-consistent group, guideline-inconsistent group with NSCLC and mediastinal sampling performed on the second or later invasive test, and guideline-inconsistent group without sampling (all  $P < .001$ ).

as the first test, those with stage II disease were more likely to have mediastinal sampling than those with stage IIIA or IIIB disease (67% vs 34% vs 16%, respectively;  $P < .001$ ).

#### Diagnostic Strategy, Subsequent Treatment, and Survival

Among patients with stage II NSCLC, those who had mediastinal sampling were more likely to have surgery as part of their treatment (90% vs 96% vs 8% for guideline consistent, guideline inconsistent with sampling second, and mediastinal sampling never done, respectively;  $P < .001$ ) (Fig 4). Among patients with stage IIIA NSCLC, surgery was performed more commonly in patients who had mediastinal lymph node sampling than in those who never had sampling (46% vs 68% vs 5%, respectively;  $P < .001$ ) (Fig 4).

Patients with NSCLC who had mediastinal sampling survived longer than those who never had mediastinal sampling ( $P < .0001$ ) (Fig 5). In patients with small cell lung cancer, those who had guideline-consistent care with mediastinal sampling as the first diagnostic test survived longer than those who had guideline-inconsistent care ( $P = .03$ ) (Fig 6), but the magnitude of the effect was much smaller.

## DISCUSSION

The findings indicate that a significant quality gap exists in the diagnostic evaluation of patients with lung cancer. Quality gaps are defined as the difference between the outcomes and processes found in practice and those obtainable using the best current knowledge.<sup>19,20</sup> We found that in patients with lung cancer with regional spread to the mediastinal lymph nodes without distant metastases, sampling of the mediastinum was performed first as per guidelines in only 21% of patients. Among patients with NSCLC, 44%

never had mediastinal sampling prior to treatment. The consequence of not sampling the mediastinum first was unnecessary testing, more thoracotomies, and more complications. We identified three main practice errors: improper sequencing of invasive tests, failure to sample the mediastinum, and overuse of thoracotomy. We cross-validated the findings by using two large, independently collected datasets and found similar results across multiple regions of the country. Both data sources suggest that a guideline-consistent strategy with mediastinal lymph node sampling done first, preferably with EBUS-TBNA, results in fewer tests and complications than alternative approaches, such as CT image-guided biopsy or thoracotomy.

This suggestion may seem counterintuitive because the sensitivity of CT image-guided biopsy is 90%, whereas the sensitivity of bronchoscopy for peripheral lesions is 34% if the lesion is  $< 2$  cm and 63% if  $> 2$  cm.<sup>3</sup> However, treatment decisions require both staging information and a tissue diagnosis. CT image-guided biopsy can only lead to a tissue diagnosis, whereas bronchoscopy with EBUS-TBNA provides both staging and diagnostic information. Even if the CT image-guided biopsy finding is positive, additional mediastinal sampling is still required if adenopathy is present. However, physicians often approach this problem by thinking in series, asking what the diagnosis is and only later asking what the stage is. This linear thinking has logical appeal and has been ingrained in training programs. The real goal is to answer both staging and diagnosis questions at the same time, essentially working on multiple questions in parallel rather than in series. Thus, multiple guidelines recommend biopsy of the mediastinal lymph nodes first in patients with evidence of nodal disease rather than biopsy of the peripheral mass.<sup>2-4,11-16</sup>

A previous single-center retrospective study suggested that a strategy of sampling the mediastinum first resulted in fewer tests and complications than a



**Table 5—Analysis of Factors Associated With Complications During the Diagnostic Evaluation**

| Variable                                       | Univariate Analysis |         | Multivariate Analysis |         |
|--|---------------------|---------|-----------------------|---------|
|  | OR (95% CI)         | P Value | OR (95% CI)           | P Value |
| Age, y   |                     |         |                       |         |
| 71-75 vs 66-70                                 | 1.14 (1.02-1.26)    | .02     | ...                   | ...     |
| 76-80 vs 66-70                                 | 1.11 (0.99-1.25)    | .06     | ...                   | ...     |
| >80 vs 66-70                                   | 1.16 (1.02-1.32)    | .03     | ...                   | ...     |
| Female vs male                                 | 1.04 (0.96-1.13)    | .35     | ...                   | ...     |
| Year of diagnosis vs 1995-1996                 |                     |         |                       |         |
| 1997   | 2.31 (1.69-3.16)    | < .0001 | 2.38 (1.73-3.28)      | < .0001 |
| 1998   | 1.66 (1.19-2.32)    | .0032   | 1.73 (1.23-2.42)      | .002    |
| 1999   | 2.28 (1.63-3.17)    | < .0001 | 2.43 (1.73-3.41)      | < .0001 |
| 2000   | 2.06 (1.58-2.70)    | < .0001 | 2.29 (1.73-3.01)      | < .0001 |
| 2001   | 1.76 (1.37-2.26)    | < .0001 | 2.17 (1.67-2.82)      | < .0001 |
| 2002   | 2.01 (1.57-2.58)    | < .0001 | 2.45 (1.89-3.18)      | < .0001 |
| 2003   | 1.96 (1.53-2.51)    | < .0001 | 2.41 (1.86-3.13)      | < .0001 |
| 2004   | 2.05 (1.61-2.62)    | < .0001 | 2.24 (1.73-2.89)      | < .0001 |
| 2005   | 2.30 (1.81-2.94)    | < .0001 | 2.41 (1.87-3.11)      | < .0001 |
| 2006   | 2.01 (1.57-2.57)    | < .0001 | 2.13 (1.64-2.76)      | < .0001 |
| 2007   | 2.19 (1.72-2.81)    | < .0001 | 2.31 (1.79-2.99)      | < .0001 |
| Race vs non-Hispanic white                     |                     |         |                       |         |
| Hispanic                                       | 1.03 (0.85-1.25)    | .76     | 0.97 (0.79-1.19)      | .77     |
| Non-Hispanic black                             | 0.72 (0.60-0.86)    | .0002   | 0.68 (0.56-0.81)      | < .0001 |
| Non-Hispanic other                             | 1.09 (0.88-1.36)    | .42     | 0.96 (0.75-1.22)      | .71     |
| Geographic region vs California                |                     |         |                       |         |
| Atlanta and rural Georgia                      | 0.71 (0.55-0.92)    | .01     | 0.73 (0.56-0.96)      | .02     |
| Connecticut                                    | 0.97 (0.82-1.16)    | .73     | 1.06 (0.88-1.27)      | .51     |
| Detroit  | 1.08 (0.92-1.27)    | .34     | 1.27 (1.08-1.51)      | .005    |
| Hawaii   | 0.933 (0.64-1.35)   | .71     | 0.93 (0.62-1.38)      | .71     |
| Iowa   | 0.60 (0.49-0.73)    | < .0001 | 0.60 (0.49-0.74)      | < .0001 |
| Kentucky                                       | 0.77 (0.64-0.92)    | .004    | 0.74 (0.61-0.89)      | .001    |
| Louisiana                                      | 0.87 (0.72-1.06)    | .16     | 0.82 (0.67-1.002)     | .052    |
| New Jersey                                     | 0.81 (0.69-0.95)    | .009    | 0.82 (0.70-0.97)      | .02     |
| New Mexico                                     | 0.92 (0.66-1.30)    | .65     | 1.05 (0.74-1.48)      | .80     |
| Seattle  | 1.02 (0.82-1.27)    | .85     | 1.02 (0.82-1.28)      | .86     |
| Texas  | 0.86 (0.76-0.98)    | .02     | 1.02 (0.88-1.18)      | .78     |
| Utah   | 0.61 (0.38-1.0)     | .05     | 0.70 (0.43-1.16)      | .17     |
| Charlson comorbidity index vs 0                |                     |         |                       |         |
| 1  | 1.10 (1.002-1.21)   | .046    | 1.08 (0.98-1.19)      | .11     |
| 2+   | 1.18 (1.05-1.31)    | .004    | 1.15 (1.03-1.28)      | .02     |
| T stage vs 1                                   |                     |         |                       |         |
| 2  | 0.67 (0.60-0.74)    | < .0001 | 0.62 (0.56-0.68)      | < .0001 |
| 3  | 0.48 (0.40-0.57)    | < .0001 | 0.42 (0.35-0.50)      | < .0001 |
| Unknown  | 0.48 (0.42-0.54)    | < .0001 | 0.43 (0.37-0.50)      | < .0001 |
| Small cell lung cancer vs NSCLC                | 0.52 (0.45-0.60)    | < .0001 | 0.53 (0.46-0.61)      | < .0001 |
| Guideline consistent vs guideline inconsistent | 0.47 (41-0.53)      | < .0001 | 0.42 (0.37-0.48)      | < .0001 |

See Table 1 legend for expansion of abbreviation.

strategy of sampling peripheral lung masses first.<sup>17</sup> The present study adds to the existing body of knowledge in this area. It confirms that sampling the mediastinum first is more effective, resulting in fewer tests and complications. To our knowledge, this multicenter study is the first to compare alternative diagnostic and staging strategies and to quantify the differences in outcome. It suggests that a significant quality gap exists in terms of test sequencing in many areas of the country because the mediastinum was sampled first in only 21% of patients.

The second quality gap we found was failure to sample the mediastinum at all. The result of relying

solely on imaging to stage the mediastinum is false upstaging, leading to missed opportunities for surgery and possibly cure. Conversely, false understaging would lead to unnecessary thoracotomies and complications.<sup>1,3</sup> This problem of underuse of mediastinal lymph node sampling has been described previously.<sup>5,21,22</sup> Studies of patients undergoing surgery for lung cancer have shown a low rate of mediastinal lymph node biopsy.<sup>5,22</sup> Although there are fewer data on TBNA and EBUS underuse, surveys of pulmonologists and pulmonary fellows have shown that bronchoscopy training varies significantly, and presumably, practice patterns vary accordingly.<sup>9,10,23-25</sup> This is in the context of randomized

**Table 6—Number of Invasive Tests Per Patient by Practice Pattern in Different Registries**

| Evaluation  | SEER (n = 11,883) |                      | TCR (n = 3,422) |                      | Total (N = 15,316) |                      |
|---|-------------------|----------------------|-----------------|----------------------|--------------------|----------------------|
|   | Mean ± SD         | Median (IQR)         | Mean ± SD       | Median (IQR)         | Mean ± SD          | Median (IQR)         |
| Guideline consistent, mediastinal sampling done first                                 | 1.4 ± 0.7         | 1 <sup>a</sup> (1-2) | 1.3 ± 0.7       | 1 <sup>a</sup> (1-2) | 1.4 ± 0.7          | 1 <sup>a</sup> (1-2) |
| Guideline inconsistent, NSCLC present, mediastinal sampling performed second or later | 2.5 ± 0.8         | 2 <sup>a</sup> (2-3) | 2.5 ± 0.8       | 2 <sup>a</sup> (2-3) | 2.5 ± 0.8          | 2 <sup>a</sup> (2-3) |
| Guideline inconsistent, NSCLC present, mediastinal sampling not done                  | 1.3 ± 0.6         | 1 (1-2)              | 1.4 ± 0.7       | 1 (1-2)              | 1.3 ± 0.6          | 1 (1-2)              |
| Guideline inconsistent, small cell lung cancer present                                | 1.4 ± 0.7         | 1 (1-2)              | 1.4 ± 0.7       | 1 (1-2)              | 1.4 ± 0.7          | 1 (1-2)              |

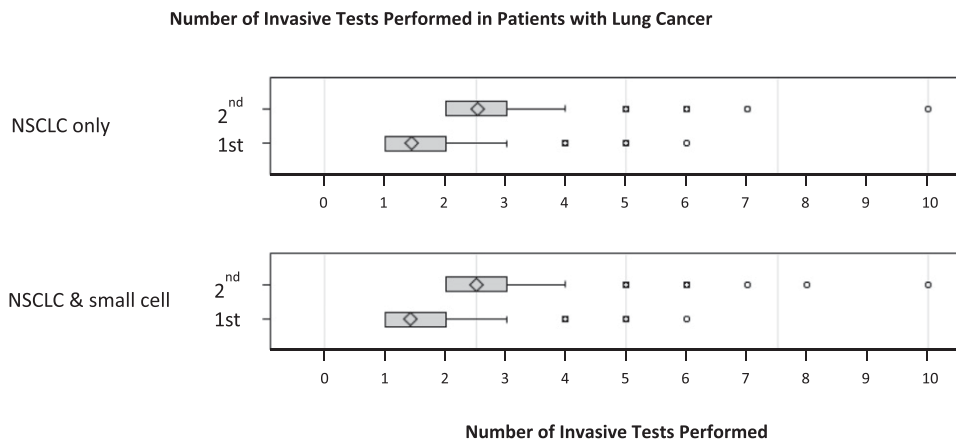
IQR = interquartile range. See Table 1 legend for expansion of other abbreviations.

<sup>a</sup>*P* < .0001 by nonparametric comparison for guideline-consistent care vs guideline-inconsistent care with sampling performed second or later.

controlled studies demonstrating that EBUS-TBNA is at least as good as mediastinoscopy and, in some cases, is more effective and less costly than mediastinoscopy.<sup>26-28</sup> More recent studies suggested that EBUS-TBNA is gaining traction,<sup>25</sup> but how far it has penetrated into the community is not clear. The present study builds on the existing literature in this area by more directly measuring the quality gap. It demonstrates that underuse of mediastinal staging techniques persists despite recent advances, with only 56% of patients with NSCLC ever undergoing mediastinal sampling.

Although problems with improper sequencing of testing and underuse of mediastinal lymph node sampling techniques clearly exist, the present study also highlights a third potential pitfall: being too aggressive, with resulting overuse of thoracotomy. EBUS-TBNA has fewer complications than mediastinoscopy, and mediastinoscopy has fewer complications than

thoracotomy.<sup>3,26,27,29</sup> Although thoracotomy allows sampling of the mediastinal lymph nodes, and many guidelines are not explicit about what comprises an optimal strategy for mediastinal lymph node sampling (thoracotomy could technically qualify as guideline consistent), thoracotomy is not an optimal first choice for evaluation of patients with a high probability of N2 or N3 disease. However, the present data suggest that thoracotomy without prior mediastinal lymph node sampling is not uncommon. Thoracotomy was the first invasive test performed in 6% of patients (Table 2). In addition, in patients with NSCLC who did not have mediastinal sampling done first, thoracotomy without prior sampling of the mediastinum was frequently performed (Table 3). The consequence of thoracotomy overuse was a significantly higher incidence of respiratory failure and hemorrhage (Table 4). EBUS-TBNA or mediastinoscopy first with thoracotomy to follow if the nodes are normal are both



**FIGURE 3.** Number of invasive diagnostic tests performed. Box plots represent median and interquartile range (25th-75th percentile) for the number of invasive tests performed. Patients who had mediastinal sampling as their first test are labeled as first. These patients received guideline-consistent care. Propensity-matched control patients who had mediastinal sampling as a second or later test are labeled as second. Patients who had mediastinal sampling first underwent fewer total tests (*P* < .0001). The first plot shows the comparison limited to patients with NSCLC given as the final diagnosis. The second plot shows the comparison limited to patients with NSCLC with patients with small cell lung cancer. See Figure 1 legend for expansion of abbreviation.

**Table 7—Practice Patterns and Guideline Consistency Stratified by Stage for SEER 2004 to 2007**

| Stage      | Evaluation Consistent With Guidelines, Sampling Done First | Evaluation Inconsistent With Guidelines                       |  |                                | Total       |
|------------|--|---|--|--------------------------------|-------------|
|            |  | NSCLC Present, Mediastinal Sampling Performed Second or Later | NSCLC Present, Mediastinal Sampling Never Done | Small Cell Lung Cancer Present |             |
| IIA or IIB | 282 (22) <sup>a</sup>                                      | 613 (48)  | 302 (24)                                       | 84 (7)                         | 1,281 (100) |
| IIIA       | 565 (19)   | 725 (24)  | 1,391 (47)                                     | 310 (10)                       | 2,991 (100) |
| IIIB       | 86 (16)  | 63 (12)   | 325 (62)                                       | 52 (10)                        | 526 (100)   |
| Total      | 933 (19)   | 1,401 (29)  | 2,018 (42)                                     | 446 (9)                        | 4,798 (100) |

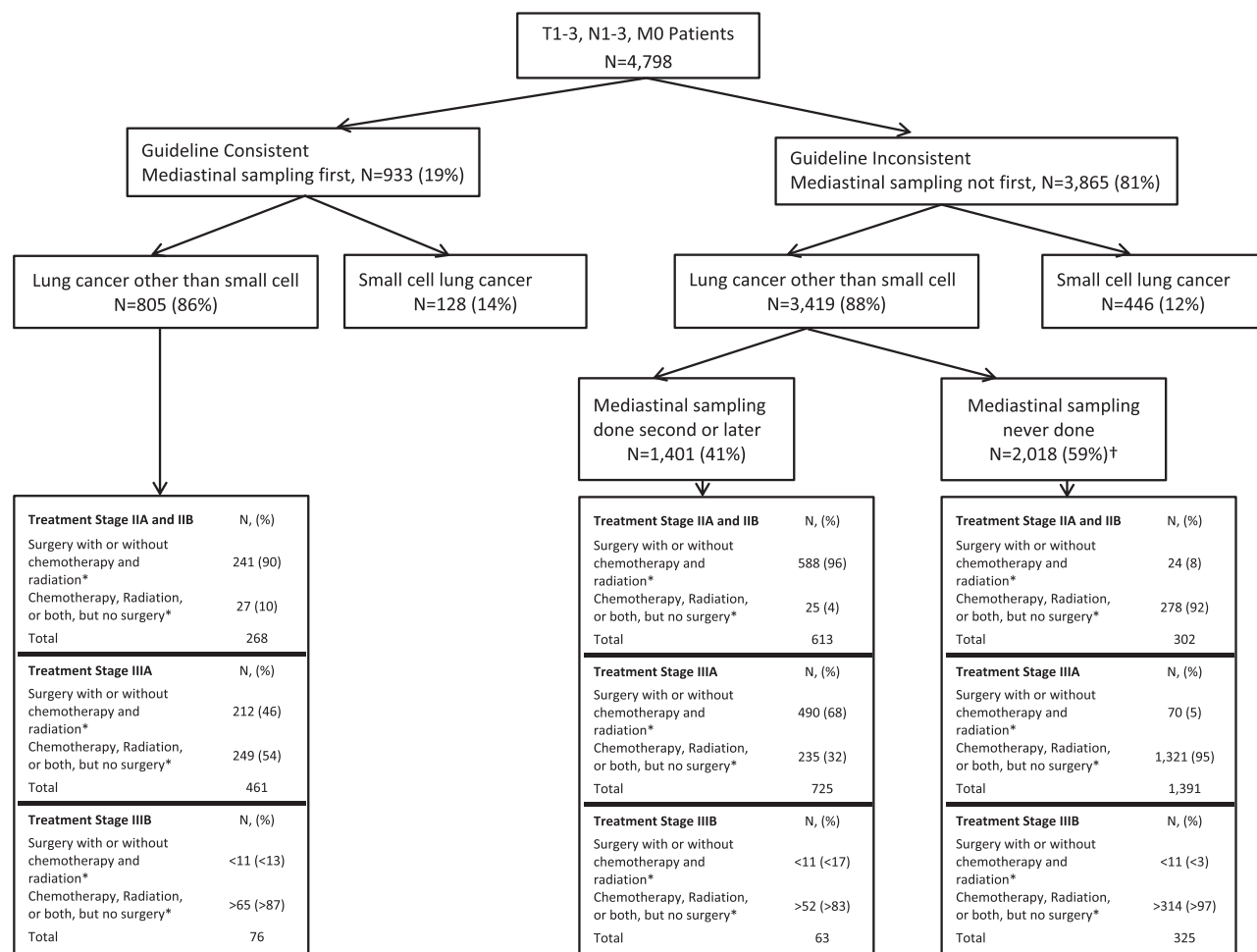
Data are presented as No. (%). See Table 1 legend for expansion of abbreviation.

<sup>a</sup>P value < .001 for guideline-consistent care and stage.

superior to proceeding directly to thoracotomy for this patient population.<sup>26,30</sup>

We also found that sampling of the mediastinal lymph nodes was associated with better survival, which is consistent with previous studies of multimodality mediastinal staging.<sup>31</sup> This is clinically plausible, but

caution is warranted when interpreting the results. Alternative explanations would include stage migration, provider effects, patient effects, or lead time bias. Stage migration is particularly relevant in the present study. It occurs when more accurate staging results in higher stage-based survival.<sup>32</sup> For example, a patient



**FIGURE 4.** Practice patterns, diagnoses, stages of disease, and treatment patterns in the SEER database from 2004 to 2007 for which there was detailed T and N stage information. \*If surgery was performed without mediastinal lymph node sampling, this was considered as not consistent with guidelines. Similarly, if surgery with lymph node sampling was performed but was not the first test and there was no prior sampling done, then this was classified as not consistent with guidelines. †If a patient received any type of treatment, such as chemotherapy or radiation, without prior lymph node sampling and went on to surgery with lymph node sampling at that time, then this was considered as no lymph node sampling prior to the first treatment. See Figure 1 legend for expansion of abbreviation.

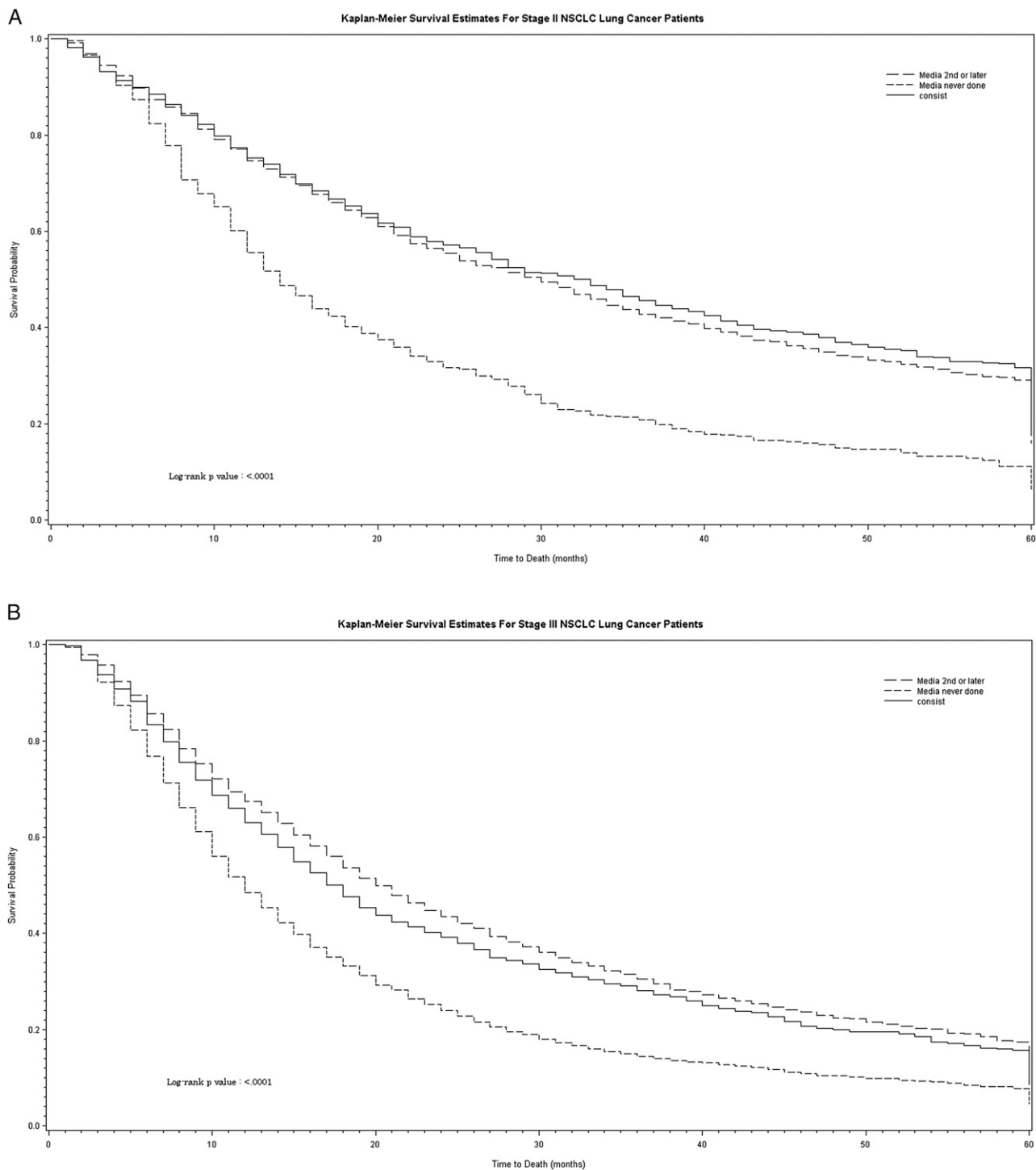


FIGURE 5. Survival in patients with NSCLC according to stage and diagnostic strategy. A, Patients with stage II NSCLC. B, Patients with stage III NSCLC. Diagnostic strategy is shown for each stage as follows: guideline-consistent care with mediastinal sampling performed first (solid line), guideline-inconsistent care with mediastinal sampling performed second or later (long-dashed line), and guideline-inconsistent care with no mediastinal sampling (short-dashed line). Staging was done using the American Joint Commission on Cancer third edition staging guidelines because cases occurring prior to 2004 did not have sufficient detail to use the sixth edition. consist = guideline-consistent care; media = mediastinal sampling. See Figure 1 legend for expansion of other abbreviation.

with clinical T2N1M0 disease on CT-PET imaging and positive EBUS-TBNA findings for N2 disease would be included in the stage III guideline-consistent group. If the same patient did not have mediastinal

sampling, he or she would be included in the stage II guideline-inconsistent group. This would result in stage migration, meaning that essentially, the guideline-consistent stage II group will have improved survival

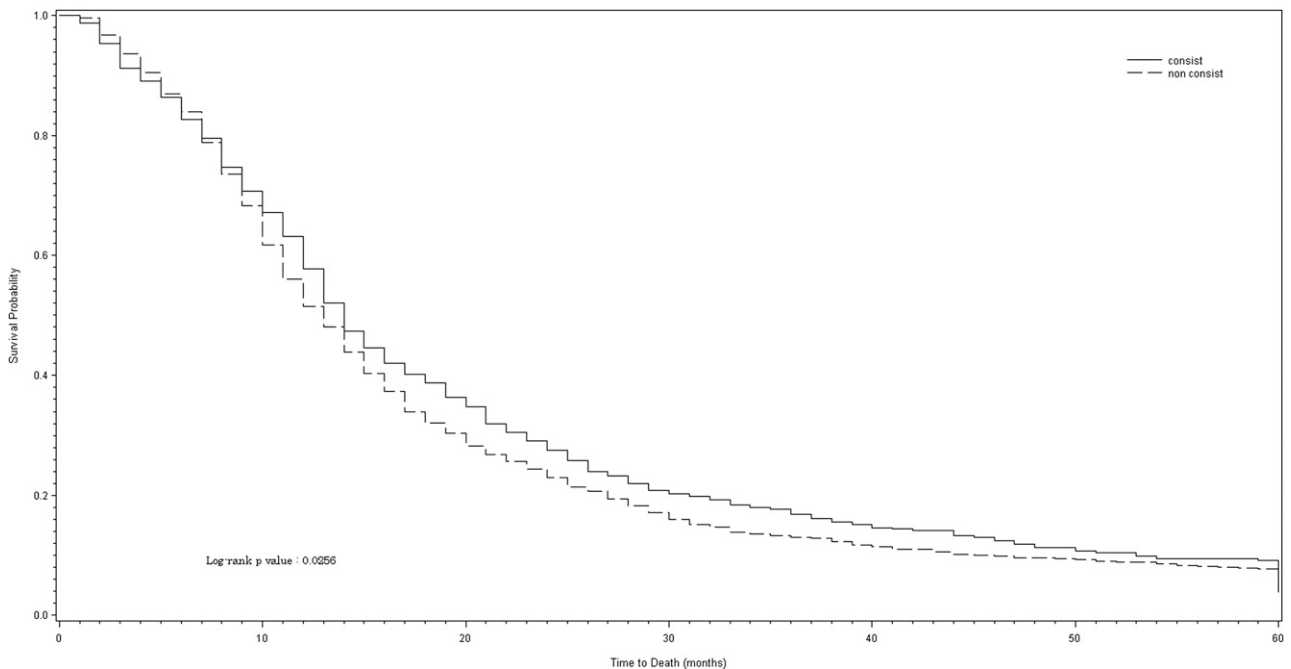


FIGURE 6. Survival in patients with small cell lung cancer according to diagnostic strategy. Diagnostic strategy is shown as guideline-consistent care with mediastinal sampling first (solid line) vs guideline-inconsistent care (dashed line). non consist = guideline-inconsistent care. See Figure 5 legend for expansion of other abbreviation.

compared with the guideline-inconsistent stage II group because the guideline-inconsistent group actually contained some patients with stage III disease. Residual confounding resulting from unmeasured patient or provider characteristics also may have affected the association between mediastinal sampling and survival if patients who had sampling were healthier and received care from high-quality providers and centers of excellence.

Although the present study adds to the existing body of evidence, several limitations are important to consider. The dataset was from administrative databases and included Medicare patients only. As such, the findings may not be generalizable to younger patients. In addition, we were limited to patients enrolled in SEER and TCR, so the findings may not be generalizable to sites not providing data to these databases.

Importantly, because these were administrative data, we had no way to verify that the lymph nodes were enlarged on either CT or PET scan. If the lymph nodes were PET scan negative and not enlarged on CT scan (ie, false negative), then mediastinal sampling would not have been warranted by the guidelines. Such patients would go to surgery, and only later would nodal disease be staged. However, previous studies have shown that patients with negative CT and PET scans have a very low incidence of occult N2 disease (around 5%-7%).<sup>33-35</sup> In addition, the SEER data from 2004 to 2007 indicate that many patients never had mediastinal sampling and yet received IIIA or IIIB staging

and no surgery. Thus, although a small fraction of the patients included in the present study probably did have normal lymph nodes on CT and PET scans and might have been misclassified as receiving guideline-inconsistent care, it is very likely that the magnitude of the quality gap is still significant, even after accounting for this limitation.

Another limitation of the dataset arising from the absence of CT and PET imaging data is that we cannot report on patients who had mediastinal adenopathy by CT or PET scan but who truly had N0 disease (ie, false positive). Some of the patients who never had mediastinal sampling probably fell within this category. However, patients who had mediastinal sampling and were subsequently identified as having N0 disease would not show up in the present cohort. Guideline-consistent care would still dictate mediastinal sampling first, preferably by EBUS-TBNA plus transbronchial biopsy and brush and lavage for diagnosis of the mass. The expected diagnostic yield of bronchoscopy for the mass in such cases would vary depending on the location and size of the mass. For central endobronchial lesions, the diagnostic yield would be about 88%; for peripheral lesions > 2 cm, about 63%; and for peripheral lesions < 2 cm, about 34%.<sup>14</sup> Peripheral radial EBUS in those cases would probably increase the yield further to around 70%.<sup>36</sup> Because radial EBUS probes use the same platform as the convex probes used for lymph node sampling, a practical approach is to first do EBUS-TBNA for lymph

node sampling and then to follow this with peripheral biopsy using radial EBUS when needed. This EBUS-TBNA/radial EBUS transbronchial biopsy strategy is superior to CT image-guided biopsy of the peripheral lesion, even in this subset of patients, because the CT image-guided biopsy strategy would still require subsequent mediastinal lymph node sampling in all cases (ie, two tests) because the mediastinal stage would be unknown. In contrast, up to 70% of cases using the EBUS-TBNA/radial EBUS transbronchial biopsy strategy would have sufficient tissue after one test.

Finally, because these are administrative data, we cannot determine which patients were not surgical candidates because of concurrent severe COPD, other comorbidities, or patient preferences. We used the Charlson comorbidity index to adjust for comorbidities, but there probably was residual confounding. No doubt, many of the patients with stage II disease who did not have surgery would fall into this category. However, we restricted the cohort to patients who received treatment. If a patient can have treatment, albeit not surgical, then sampling of the mediastinum is still necessary because radiation and chemotherapy protocols would change significantly depending on the results. In addition, because we cannot be certain about the severity of the comorbidities present, it may be that the excess complications observed on a per-procedure basis were the result of differences in the clinical characteristics of the patients who we could not measure, explaining the observed higher incidence of complications per procedure in the guideline-inconsistent groups. However, if the guideline-inconsistent group with mediastinal sampling second or later had more severe unmeasured comorbidities, it is unlikely that such differences would lead physicians to order more invasive diagnostic tests, which is what occurred. What is more likely is that physicians in the guideline-inconsistent group ordered invasive tests that provided useful information regarding diagnosis but were insufficient to fully stage the disease. As such, these physicians had to order another test to develop a treatment plan.

In summary, we found large quality gaps in the diagnosis and staging of patients with lung cancer. The three main quality gaps were (1) failure to sample the mediastinum first in 79% of patients, (2) failure to sample the mediastinum at all in 44% of patients with NSCLC, and (3) overuse of thoracotomy without prior mediastinal lymph node sampling in 6% of patients. Educational initiatives to address this problem should focus on a clear message: In patients with suspected lung cancer with hilar or mediastinal adenopathy without evidence of distant disease, sample the mediastinum first, preferably with EBUS-TBNA. Following this simple dictum will result in lower costs and

morbidity because it will eliminate many unnecessary tests and their associated complications.

#### ACKNOWLEDGMENTS

**Author contributions:** Dr Ost had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

*Dr Ost:* contributed as principal investigator and to the study oversight; data analysis; and writing, editing, and review of the manuscript.

*Dr Niu:* contributed to the data analysis and writing, editing, and review of the manuscript.

*Dr Elting:* contributed to the data analysis and writing, editing, and review of the manuscript.

*Dr Buchholz:* contributed to the writing, editing, and review of the manuscript.

*Dr Giordano:* contributed to the data analysis and writing, editing, and review of the manuscript.

**Financial/nonfinancial disclosures:** The authors have reported to *CHEST* that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

**Role of sponsors:** The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the Texas Department of State Health Services, Cancer Prevention Research Institute of Texas, or the Centers for Disease Control and Prevention. The interpretation and reporting of Surveillance, Epidemiology, and End Results (SEER)-Medicare data are the sole responsibility of the authors.

**Other contributions:** The authors acknowledge the efforts of the Applied Research Program, National Cancer Institute; the Office of Research, Development and Information, Centers for Medicare & Medicaid Services; Information Management Services, Inc; and the SEER program tumor registries in the creation of the SEER-Medicare database.

**Additional information:** The e-Appendix and e-Tables can be found in the "Supplemental Materials" area of the online article.

#### REFERENCES

1. Silvestri GA, Gould MK, Margolis ML, et al. Noninvasive staging of non-small cell lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest*. 2007;132(3\_suppl):178S-201S.
2. Almeida FA, Uzbek M, Ost D. Initial evaluation of the non-small cell lung cancer patient: diagnosis and staging. *Curr Opin Pulm Med*. 2010;16(4):307-314.
3. Detterbeck FC, Jantz MA, Wallace M, Vansteenkiste J, Silvestri GA. Invasive mediastinal staging of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest*. 2007;132(3\_suppl):202S-220S.
4. Rivera MP, Mehta AC. Initial diagnosis of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest*. 2007;132(3\_suppl):131S-148S.
5. Little AG, Rusch VW, Bonner JA, et al. Patterns of surgical care of lung cancer patients. *Ann Thorac Surg*. 2005;80(6):2051-2056.
6. Schipper P, Schofield M. Minimally invasive staging of N2 disease: endobronchial ultrasound/transesophageal endoscopic ultrasound, mediastinoscopy, and thoracoscopy. *Thorac Surg Clin*. 2008;18(4):363-379.
7. Haponik EF, Shure D. Underutilization of transbronchial needle aspiration: experiences of current pulmonary fellows. *Chest*. 1997;112(1):251-253.
8. Smyth CM, Stead RJ. Survey of flexible fiberoptic bronchoscopy in the United Kingdom. *Eur Respir J*. 2002;19(3):458-463.

9. Prakash UB, Offord KP, Stubbs SE. Bronchoscopy in North America: the ACCP survey. *Chest*. 1991;100(6):1668-1675.
10. Colt HG, Prakash UBS, Offord KP. Bronchoscopy in North America: survey by the American Association for Bronchology, 1999. *J Bronchol*. 2000;7(1):8-25.
11. Baldwin DR, White B, Schmidt-Hansen M, Champion AR, Melder AM; Guideline Development Group. Diagnosis and treatment of lung cancer: summary of updated NICE guidance. *BMJ*. 2011;342:d2110.
12. Crinò L, Weder W, van Meerbeeck J, Felip E; ESMO Guidelines Working Group. Early stage and locally advanced (non-metastatic) non-small-cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2010;21(suppl 5):v103-v115.
13. De Leyn P, Lardinois D, Van Schil PE, et al. ESTS guidelines for preoperative lymph node staging for non-small cell lung cancer. *Eur J Cardiothorac Surg*. 2007;32(1):1-8.
14. Rivera MP, Mehta AC, Wahidi MM. Establishing the diagnosis of lung cancer: diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2013;143(5\_suppl):e142S-e165S.
15. Silvestri GA, Gonzalez AV, Jantz MA, 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*. 2013;143(5\_suppl):e211S-e250S.
16. Ost DE, Yeung SCJ, Tanoue LT, Gould MK. Clinical and organizational factors in the initial evaluation of patients with lung cancer: diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. 2013;143(5\_suppl):e121S-e141S.
17. Almeida FA, Casal RF, Jimenez CA, et al. Quality gaps and comparative effectiveness in lung cancer staging: the impact of test sequencing on outcomes. *Chest*. 2013;144(6):1776-1782.
18. Wiener RS, Schwartz LM, Woloshin S, Welch HG. Population-based risk for complications after transthoracic needle lung biopsy of a pulmonary nodule: an analysis of discharge records. *Ann Intern Med*. 2011;155(3):137-144.
19. Shojania K, McDonald K, Wachter R, Owens DK. *Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies, Volume 1—Series Overview and Methodology*. Rockville, MD: Agency for Healthcare Research and Quality; 2004:1-37. AHRQ publication 04-0051-1.
20. Institute of Medicine (Committee on Quality of Health Care in America). *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academies Press; 2001.
21. Allen JW, Farooq A, O'Brien TF, Osarogiagbon RU. Quality of surgical resection for nonsmall cell lung cancer in a US metropolitan area. *Cancer*. 2011;117(1):134-142.
22. Osarogiagbon RU, Allen JW, Farooq A, Berry A, O'Brien T. Pathologic lymph node staging practice and stage-predicted survival after resection of lung cancer. *Ann Thorac Surg*. 2011;91(5):1486-1492.
23. Haponik EF, Russell GB, Beamis JF Jr, et al. Bronchoscopy training: current fellows' experiences and some concerns for the future. *Chest*. 2000;118(3):625-630.
24. Pastis NJ, Nietert PJ, Silvestri GA; American College of Chest Physicians Interventional Chest/Diagnostic Procedures Network Steering Committee. Variation in training for interventional pulmonary procedures among US pulmonary/critical care fellowships: a survey of fellowship directors. *Chest*. 2005;127(5):1614-1621.
25. Tanner NT, Pastis NJ, Silvestri GA. Training for linear endobronchial ultrasound among US pulmonary/critical care fellowships: a survey of fellowship directors. *Chest*. 2012;143(2):423-428.
26. Sharples LD, Jackson C, Wheaton E, et al. Clinical effectiveness and cost-effectiveness of endobronchial and endoscopic ultrasound relative to surgical staging in potentially resectable lung cancer: results from the ASTER randomised controlled trial. *Health Technol Assess*. 2012;16(18):1-75.
27. Yasufuku K, Pierre A, Darling G, et al. A prospective controlled trial of endobronchial ultrasound-guided transbronchial needle aspiration compared with mediastinoscopy for mediastinal lymph node staging of lung cancer. *J Thorac Cardiovasc Surg*. 2011;142(6):1393-1400.
28. Ernst A, Anantham D, Eberhardt R, Krasnik M, Herth FJ. Diagnosis of mediastinal adenopathy-real-time endobronchial ultrasound guided needle aspiration versus mediastinoscopy. *J Thorac Oncol*. 2008;3(6):577-582.
29. Eapen GA, Shah AM, Lei X, et al; American College of Chest Physicians Quality Improvement Registry, Education, and Evaluation (AQuIRE) Participants. Complications, consequences, and practice patterns of endobronchial ultrasound-guided transbronchial needle aspiration: results of the AQuIRE Registry. *Chest*. 2013;143(4):1044-1053.
30. Annema JT, Versteegh MI, Veselić M, et al. Endoscopic ultrasound added to mediastinoscopy for preoperative staging of patients with lung cancer. *JAMA*. 2005;294(8):931-936.
31. Farjah F, Flum DR, Ramsey SD, Heagerty PJ, Symons RG, Wood DE. Multi-modality mediastinal staging for lung cancer among Medicare beneficiaries. *J Thorac Oncol*. 2009;4(3):355-363.
32. 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.
33. Cerfolio RJ, Bryant AS. Survival of patients with unsuspected N2 (stage IIIA) nonsmall-cell lung cancer. *Ann Thorac Surg*. 2008;86(2):362-366.
34. Cerfolio RJ, Bryant AS, Eloubeidi MA. Routine mediastinoscopy and esophageal ultrasound fine-needle aspiration in patients with non-small cell lung cancer who are clinically N2 negative: a prospective study. *Chest*. 2006;130(6):1791-1795.
35. Herth FJ, Eberhardt R, Krasnik M, Ernst A. Endobronchial ultrasound-guided transbronchial needle aspiration of lymph nodes in the radiologically and positron emission tomography-normal mediastinum in patients with lung cancer. *Chest*. 2008;133(4):887-891.
36. Wang Memoli JS, Nietert PJ, Silvestri GA. Meta-analysis of guided bronchoscopy for the evaluation of the pulmonary nodule. *Chest*. 2012;142(2):385-393.