CHEST

Quality Gaps and Comparative Effectiveness in Lung Cancer Staging and Diagnosis

David E. Ost, MD, MPH; Jiangong Niu, MS; Hui Zhao, PhD; Horiana B. Grosu, MD; and Sharon H. Giordano, MD

BACKGROUND: Guidelines recommend mediastinal sampling first for patients with mediastinal lymphadenopathy with suspected lung cancer. The objective of this study was to describe practice patterns and outcomes of diagnostic strategies in patients with lung cancer.

METHODS: This study included a retrospective cohort of 15,914 patients with lung cancer with T1-3N1-3M0 disease diagnosed from 2004 to 2013 in the National Cancer Institute's Surveillance, Epidemiology, and End Results or Texas Cancer Registry Medicare-linked databases. Patients who had mediastinal sampling as their first invasive test were classified as guideline consistent; all others were guideline inconsistent. Propensity matching was used to compare the number of tests performed, and multivariable logistic regression was used to compare the incidence of complications.

RESULTS: Guideline-consistent care increased from 23% to 34% of patients from 2004 to 2013 ($P < .001$). Use of endobronchial ultrasound-guided transbronchial needle aspiration increased from 0.1% to 25% of all patients ($P < .001$), and mediastinal sampling increased from 54% to 64% ($P < .0001$). Guideline-consistent care was associated with fewer thoracotomies (38% vs 71%; $P < .001$) and CT scan-guided biopsies (10% vs 75%; $P < .001$) than guideline-inconsistent care but more transbronchial needle aspirations (59% vs 12%; $P < .001$). Guideline-consistent care was associated with fewer pneumothoraxes (5.1% vs 22%; $P < .001$), chest tubes (0.9% vs 4.4%; $P < .001$), hemorrhages (3.5% vs 5.8%; $P < .001$), and respiratory failure events (2.7% vs 3.7%; $P = .047$) than guideline-inconsistent care. Bronchoscopic mediastinal sampling was associated with fewer complications than surgical mediastinal sampling.

CONCLUSIONS: Guideline-consistent care with mediastinal sampling first was associated with fewer tests and complications. Quality gaps decreased with the introduction of endobronchial ultrasound-guided transbronchial needle aspiration but persist. Gaps include failure to sample the mediastinum first, failure to sample the mediastinum at all, and overuse of thoracotomy.

CHEST 2020; 157(5):1322-1345

KEY WORDS: bronchoscopy; cancer; lung cancer; mediastinum; quality improvement

FOR EDITORIAL COMMENT, SEE PAGE 1062

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

AFFILIATIONS: From the Department of Pulmonary Medicine (Drs Ost and Grosu) and the Department of Health Services Research (Ms Niu and Drs Zhao and Giordano), The University of Texas MD Anderson Cancer Center, Houston, TX.

FUNDING/SUPPORT: This work was supported in part by the Center for Comparative Effectiveness Research on Cancer in Texas, a multi-university consortium funded by the Cancer Prevention and Research Institute of Texas [Grant RP160674] and by the Cancer Center Support Grant [National Cancer Institute, Grant P30 CA016672]. The collection of cancer incidence data 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

Current lung cancer guidelines from the American College of Chest Physicians and the National Cancer Care Network recommend mediastinal lymph node sampling as the first test in patients who have evidence of possible lymph node involvement by CT and PET imaging without distant metastases.[1-4] This guidance is because lymph node status in these patients will determine whether the disease is amenable to surgical resection. Although PET/CT imaging is useful for assessing the mediastinum, its limited accuracy makes lymph node sampling essential in these cases.[1-5]

Previous comparative effectiveness studies evaluating the impact of test sequencing have validated these evidence-based guidelines, showing that sampling of the mediastinum first (rather than following a peripheral biopsy) leads to fewer invasive tests and fewer complications.[6-8] These studies reported that guidelines were rarely followed, with only 21% of patients with lung cancer and regional lymph node involvement receiving guideline-consistent care. In the past, inadequate fellowship training in conventional transbronchial needle aspiration (TBNA) and the technical difficulty of the procedure may have contributed to this finding.[9-12] However, these large database studies only analyzed data up to 2007, and endobronchial ultrasound (EBUS)-guided TBNA was not widely available at that time. EBUS-TBNA has made lymph node sampling safer and more effective, such that the most recent lung cancer guidelines recommend use of EBUS-TBNA as the first invasive test for sampling of the mediastinum when indicated.[13]

The goal of the current study was to compare practice patterns and outcomes of diagnostic strategies in patients with lung cancer and mediastinal lymph node involvement without distant metastasis. A secondary objective was to compare EBUS-first strategies vs other practice patterns that have previously been considered guideline consistent (eg, mediastinoscopy and thoracotomy) in terms of their complication rates and resource utilization.

Patients and Methods

Data Source

A retrospective cohort analysis was performed by using the Texas Cancer Registry (TCR) and the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) database for the years 2004 to 2013. These registry data were linked to Medicare claims. The data also provided patients' socioeconomic information such as poverty and education at the census tract level based on the US 2000 and 2010 censuses. We compared the registries and analyzed practice patterns and outcomes. This study was approved by MD Anderson Cancer Center Institutional Review Board 4, and a waiver of informed consent was obtained.

Study Participants

The cohort consisted of 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 (e-Appendix 1). The patients included in the study were T1-3N1-3M0 based on the SEER database. SEER uses both clinical staging and cytology/pathology data to arrive at the final cancer stage. However, SEER does not provide specific data on PET or CT imaging results, and thus precise PET/CT staging is not available.

Diagnostic and Staging Strategy

The type and sequencing of invasive tests used for diagnosis and staging were determined by using CPT and *International Classification of Diseases, Ninth Revision,* codes (e-Tables 1 and 2) using a strategy similar to previously published reports[7,8] but updated to reflect changes in codes that occurred over the years. Invasive tests were defined as CT scan-guided needle biopsy, bronchoscopy with TBNA or EBUS-TBNA, endoscopy with ultrasound-guided needle aspiration, mediastinoscopy, or thoracotomy. Only tests performed within the 6 months prior to initiation of treatment were included. Patients with no evaluation recorded were excluded from the analysis.

Patients were categorized 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, non-small cell lung cancer (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. The definition of guideline-consistent care was based on the information available to physicians at the time of diagnosis as well as considerations of how fast and widely accepted these guidelines might be. The second and third editions of the lung cancer guidelines were published in 2007 and 2013. The main change that occurred in the third edition guidelines was that EBUS-TBNA became the recommended first test for mediastinal lymph node sampling, ahead of mediastinoscopy and other modalities.[3,13] Given the time needed for dissemination and implementation, we chose to consider any form of mediastinal sampling as guideline consistent in 2013, provided it was the first test, as guidelines prior to 2004 were all consistent on this point. Mediastinal sampling procedures were defined as bronchoscopy with TBNA or EBUS-TBNA, endoscopy with ultrasound-guided needle aspiration, mediastinoscopy, thoracoscopy, or thoracotomy with mediastinal lymph node sampling (the online supplement provides details on categories and criteria).

In a secondary analysis, we further subdivided the guideline-consistent group into bronchoscopic mediastinal sampling (EBUS-TBNA or

Prevention's National Program of Cancer Registries Cooperative Agreement [Grant 5U58/DP000824-05].

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

DOI: https://doi.org/10.1016/j.chest.2019.09.025

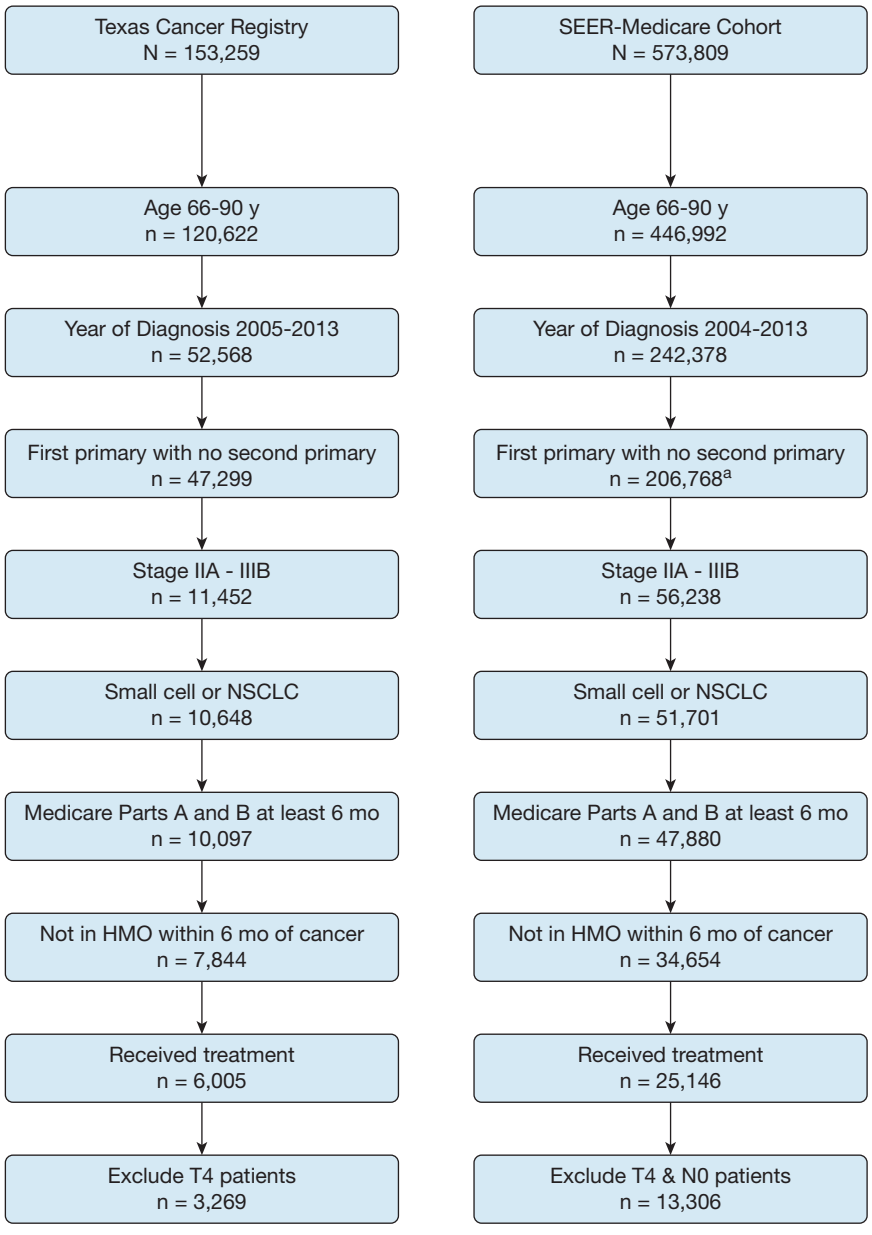

Figure 1 – *Study cohort selection results. SEER and Texas Cancer Registry 2004 to 2013. [a]Note that the SEER registry does not have complete staging information on all patients. There were 4,982 patients who met all the other inclusion criteria (NSCLC, year of diagnosis, first primary cancer with no second primary, Medicare Parts A and B at least 6 months, not in HMO within 6 months, and received treatment) who were not included because in the SEER database their stage was classified as stage unknown. These 4,982 patients were excluded. HMO = health maintenance organization; NSCLC = non-small cell lung cancer; SEER = Surveillance, Epidemiology, and End Results.*

TBNA) vs surgical mediastinal sampling (mediastinoscopy, thoracotomy, or both).

Outcomes

The primary outcome was whether diagnostic test sequencing for each patient was consistent with guidelines. Secondary outcomes included whether mediastinal lymph node sampling was ever performed prior to treatment in patients with NSCLC, complications related to the diagnostic and staging evaluation, the number of invasive diagnostic tests performed, time to complete diagnosis and staging, and time to first treatment. We used a methodology similar to that previously published to identify complications, including pneumothorax, hemorrhage, and respiratory failure.[7,8,14] For thoracotomy and thoracoscopic surgery, 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 following the procedure.

When analyzing complications, we considered whether surgeries were for diagnosis, staging, or treatment. Because some surgeries are for diagnosis, staging, and treatment (eg, wedge biopsy with frozen section, mediastinal lymph node dissection, and then lobectomy), a system of attributing complications to either the diagnostic/staging evaluation or to treatment was developed. Complications arising from treatment were not counted. Because segmentectomy, lobectomy, pneumonectomy, tracheal resections, and chest wall resections for stage II disease would be part of treatment, if any of these major therapeutic resections were performed, the entire procedure was considered as part of treatment rather than diagnosis, provided the patient had stage II disease. The rationale behind this reasoning is that even if an alternative form of mediastinal staging had been chosen (eg, EBUS), eventually the patient would have had to have a therapeutic resection, and thus these procedures and their complications would have occurred and should properly be classified as complications of treatment rather than as diagnosis and staging. This approach tends to favor surgical staging in the analysis because,

by definition, patients with stage II disease who went straight to surgery would have no complications. However, if patients had stage III disease, and no previous mediastinal staging was conducted prior to the surgery, then even if a major therapeutic resection was performed, the procedures were counted and the complications were included. The rationale is that the standard guideline recommendation was chemotherapy and radiation for stage III disease, and in the absence of a previous staging procedure, the complications arising from the surgery would have been avoidable had previous staging been conducted.

This approach balances out the benefit provided by not counting major therapeutic resections and their complications in patients with stage II disease. For a population of patients without previous mediastinal sampling, an aggressive surgical approach would have fewer complications using this methodology if the surgeons accurately operated solely on patients with stage II disease; conversely, if surgeons operated on many patients with stage III disease, complications would then be considerably higher. It is worth noting that for patients with both stage II and III disease, surgeries that did not include a major resection (eg, mediastinal lymph node dissection and biopsies, pleural biopsies, wedge resections) were always counted, as were their complications, provided no major resection was performed on the same day. For the current discussion, this group of surgeries was labeled diagnostic/staging surgeries.

Time to complete diagnosis and staging was defined as date of the last invasive test minus the date of the first invasive test (eg, if only one test was required, the result would be zero). Time to treatment was defined as the date of the first treatment minus the date of the first invasive test. We also conducted an analysis to assess the relationship between diagnostic practice patterns, subsequent treatment modalities, and survival.

Statistical Analysis

This study used the χ^2 test for categorical variables, Student *t* tests for continuous normally distributed variables, and the Wilcoxon rank sum test for nonnormally distributed variables to compare characteristics of patients and outcomes. The Cochrane-Armitage trend test was used to evaluate utilization patterns over time, and the Kruskal-Wallis test was used to compare frequency of test utilization among different sampling groups. Mmultivariable logistic regression was used to identify factors associated with complications from diagnostic testing, and we compared this vs hierarchical models with patients nested within physicians; because there was no significant difference, we report ordinary multivariable logistic regression. We decided a priori that variables significantly associated with outcomes at the .2 level in univariate analysis would qualify as candidate variables for multivariable 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, and thus propensity scores were used 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 of having guideline-consistent care was estimated by using logistic regression analysis incorporating the following variables: age, sex, race, year of diagnosis, Charlson Comorbidity Index score, T stage, N stage, geographic region, and cancer type. Propensity matching was used for the cumulative incidence of complications. We also performed a stratified analysis of bronchoscopic vs surgical mediastinal sampling and mediastinal sampling performed as the first invasive test vs second or later. All statistical analyses were performed at a significance level of .05. No adjustment for multiple comparisons was made. All data were analyzed by using SAS version 9.4 (SAS Institute, Inc.).

Results

SEER-Medicare and TCR-Medicare Cohort

A total of 13,306 SEER-Medicare patients and 3,269 TCR-Medicare patients met the inclusion criteria (Fig 1). For subsequent analysis, we combined the two registries and controlled for geographic region. Patient characteristics for the combined cohort are shown in Table 1. Of the 16,575 patients eligible, 661 (4%) had no Medicare data indicating that any diagnostic testing was performed. The remaining 15,914 (96%) patients had Medicare data, and this group constituted the final study cohort.

Practice Patterns and Consistency With Guidelines

Mediastinal sampling was eventually performed in 8,298 of the 14,492 (57%; 95% CI, 56-58) patients with NSCLC. The method used to sample the mediastinum was thoracotomy without mediastinoscopy in 37%, mediastinoscopy alone in 24%, mediastinoscopy with thoracotomy in 7%, TBNA in 16%, and EBUS-TBNA in 16%.

From 2004 to 2013, a total of 4,044 of the 15,914 (25.4%; 95% CI, 24.7-26.1) patients received guideline-consistent care with mediastinal lymph node sampling performed first (Table 2). The percentage of patients receiving guideline-consistent care increased from 22.6% in 2004 to 33.6% in 2013 (difference, 10.9%; 95% CI, 7.7-14.2; $P < .001$, $P < .001$ test for trend) (Fig 2). The most common first invasive diagnostic test changed over time. Bronchoscopy without TBNA was the most frequent first test used from 2004 through 2006. From 2007 to 2014, CT scan-guided biopsy supplanted bronchoscopy without TBNA as the most common initial test.

Among patients with NSCLC who did not have mediastinal sampling performed during the first test, 5,776 went on to have a second invasive test. Mediastinal sampling occurred in 3,316 (57%) of second invasive tests (Table 3). When mediastinal sampling was performed, surgical methods were used in 88% of cases, whereas bronchoscopic approaches were used in 12%.

Among patients with NSCLC who did not have mediastinal sampling performed during the first or second test, 1,186 went on to have a third invasive test.

TABLE 1] Patient Characteristics

Variable	Evaluation Consistent With Guidelines, NSCLC, Mediastinal Sampling Done First	Evaluation Inconsistent With Guidelines, NSCLC Present, Mediastinal Sampling Performed on the Second or Later Biopsy	Evaluation Not Consistent With Guidelines, NSCLC Present, Mediastinal Sampling Never Done	Small Cell Present, Guideline-Inconsistent, No Mediastinal Sampling on First Test	No Evaluation Recorded[a]	Total	*P* Value, χ^2
All subjects	4,044 (100)	4,254 (100)	6,194 (100)	1,422 (100)	661 (100)	16,575 (100)	…
Age, y							< .0001
66-70	1,386 (34.3)	1,390 (32.7)	1,652 (26.7)	502 (35.3)	239 (36.2)	5,169 (31.2)	
71-75	1,225 (30.3)	1,312 (30.8)	1,672 (27)	425 (29.9)	191 (28.9)	4,825 (29.1)	
76-80	904 (22.4)	1,009 (23.7)	1526 (24.6)	302 (21.2)	138 (20.9)	3,879 (23.4)	
> 80	529 (13.1)	543 (12.8)	1,344 (21.7)	193 (13.6)	93 (14.1)	2,702 (16.3)	
Sex							< .0001
Female	1,999 (49.4)	2,034 (47.8)	2,861 (46.2)	750 (52.7)	259 (39.2)	7,903 (47.7)	
Male	2,045 (50.6)	2,220 (52.2)	3,333 (53.8)	672 (47.3)	402 (60.8)	8,672 (52.3)	
Race							< .0001
Non-Hispanic white	3,468 (85.8)	3,599 (84.6)	5,051 (81.6)	1,237 (87)	537 (81.2)	13,892 (83.8)	
Hispanic	167 (4.1)	208 (4.9)	301 (4.9)	62 (4.4)	37 (5.6)	775 (4.7)	
Non-Hispanic black	290 (7.2)	261 (6.1)	590 (9.5)	80 (5.6)	56 (8.5)	1,277 (7.7)	
Non-Hispanic other	119 (2.9)	186 (4.4)	252 (4.1)	43 (3)	31 (4.7)	631 (3.8)	
Urban/rural							.0009
Big metro	2,014 (49.8)	2,256 (53)	3,134 (50.6)	642 (45.2)	329 (49.8)	8,375 (50.5)	
Metro	1,257 (31.1)	1,172 (27.6)	1,860 (30)	477 (33.5)	183 (27.7)	4,949 (29.9)	
Urban	249 (6.2)	271 (6.4)	383 (6.2)	92 (6.5)	43 (6.5)	1,038 (6.3)	
Less urban	426 (10.5)	451 (10.6)	679 (11)	176 (12.4)	88 (13.3)	1,820 (11)	
Rural	98 (2.4)	104 (2.4)	138 (2.2)	35 (2.5)	18 (2.7)	393 (2.4)	
Year of diagnosis							< .0001
2004	343 (8.5)	385 (9.1)	622 (10)	166 (11.7)	66 (10)	1,582 (9.5)	
2005	375 (9.3)	540 (12.7)	777 (12.5)	212 (14.9)	68 (10.3)	1,972 (11.9)	
2006	372 (9.2)	475 (11.2)	729 (11.8)	164 (11.5)	58 (8.8)	1,798 (10.9)	
2007	396 (9.8)	452 (10.6)	715 (11.5)	153 (10.8)	71 (10.7)	1,787 (10.8)	
2008	413 (10.2)	443 (10.4)	651 (10.5)	149 (10.5)	78 (11.8)	1,734 (10.5)	
2009	418 (10.3)	449 (10.6)	625 (10.1)	126 (8.9)	69 (10.4)	1,687 (10.2)	
2010	374 (9.3)	396 (9.3)	534 (8.6)	136 (9.6)	54 (8.2)	1,494 (9)	

(Continued)

TABLE 1] *(Continued)*

Variable	Evaluation Consistent With Guidelines, NSCLC, Mediastinal Sampling Done First	Evaluation Inconsistent With Guidelines, NSCLC Present, Mediastinal Sampling Performed on the Second or Later Biopsy	Evaluation Not Consistent With Guidelines, NSCLC Present, Mediastinal Sampling Never Done	Small Cell Present, Guideline-Inconsistent, No Mediastinal Sampling on First Test	No Evaluation Recorded[a]	Total	*P* Value, χ^2
2011	416 (10.3)	363 (8.5)	563 (9.1)	98 (6.9)	62 (9.4)	1,502 (9.1)	
2012	472 (11.7)	395 (9.3)	512 (8.3)	120 (8.4)	69 (10.4)	1568 (9.5)	
2013	465 (11.5)	356 (8.4)	466 (7.5)	98 (6.9)	66 (10)	1,451 (8.8)	
SEER/TCR region							< .0001
California	617 (15.3)	926 (21.8)	1,243 (20.1)	258 (18.1)	130 (19.7)	3174 (19.2)	
Connecticut	244 (6)	226 (5.3)	268 (4.3)	58 (4.1)	43 (6.5)	839 (5.1)	
Detroit	265 (6.6)	227 (5.3)	324 (5.2)	70 (4.9)	33 (5)	919 (5.5)	
Georgia	602 (14.9)	448 (10.5)	663 (10.7)	162 (11.4)	85 (12.9)	1,960 (11.8)	
Hawaii	24 (0.6)	38 (0.9)	62 (1)	11 (0.8)	< 11 (1.7)	< 146 (< 0.9)	
Iowa	183 (4.5)	188 (4.4)	269 (4.3)	69 (4.9)	28 (4.2)	737 (4.5)	
Kentucky	297 (7.3)	373 (8.8)	525 (8.5)	169 (11.9)	51 (7.7)	1,415 (8.5)	
Louisiana	138 (3.4)	215 (5.1)	514 (8.3)	113 (8)	36 (5.5)	1,016 (6.1)	
New Jersey	587 (14.5)	502 (11.8)	756 (12.2)	159 (11.2)	79 (12)	2,083 (12.6)	
New Mexico	40 (1)	45 (1.1)	101 (1.6)	19 (1.3)	< 11 (< 1.7)	< 216 (< 1.3)	
Seattle	198 (4.9)	189 (4.4)	238 (3.8)	45 (3.2)	27 (4.1)	697 (4.2)	
Texas	816 (20.2)	844 (19.8)	1,201 (19.4)	278 (19.6)	130 (19.7)	3,269 (19.7)	
Utah	33 (0.8)	33 (0.8)	30 (0.5)	11 (0.8)	< 11 (< 1.7)	< 118 (0.7)	
Marital status							< .0001
Married	2,233 (55.2)	2,323 (54.6)	3,117 (50.3)	721 (50.7)	343 (51.9)	8,737 (52.7)	
Not married	1,525 (37.7)	1,627 (38.3)	2,619 (42.3)	600 (42.2)	269 (40.7)	6,640 (40.1)	
Unknown	286 (7.1)	304 (7.2)	458 (7.4)	101 (7.1)	49 (7.4)	1,198 (7.2)	
Comorbidities							< .0001
0	1,716 (42.4)	1,981 (46.6)	2,480 (40)	550 (38.7)	344 (52)	7,071 (42.7)	
1	1,324 (32.7)	1,302 (30.6)	1,913 (30.9)	467 (32.8)	156 (23.6)	5,162 (31.1)	
≥ 2	1,004 (24.8)	971 (22.8)	1,801 (29.1)	405 (28.5)	161 (24.4)	4,342 (26.2)	
% Poverty in that patient's census tract[b]							< .0001
≤ 4.73%	1,017 (25.2)	1,084 (25.5)	1,370 (22.1)	312 (21.9)	150 (22.7)	3,933 (23.7)	

(Continued)

TABLE 1] *(Continued)*

Variable	Evaluation Consistent With Guidelines, NSCLC, Mediastinal Sampling Done First	Evaluation Inconsistent With Guidelines, NSCLC Present, Mediastinal Sampling Performed on the Second or Later Biopsy	Evaluation Not Consistent With Guidelines, NSCLC Present, Mediastinal Sampling Never Done	Small Cell Present, Guideline-Inconsistent, No Mediastinal Sampling on First Test	No Evaluation Recorded[a]	Total	*P* Value, χ^2
4.74%-9.28%	932 (23.1)	1,046 (24.6)	1,432 (23.1)	310 (21.8)	136 (20.6)	3,856 (23.3)	
9.29%-17.48%	976 (24.1)	1,030 (24.2)	1,507 (24.3)	401 (28.2)	166 (25.1)	4,080 (24.6)	
> 17.48%	1,119 (27.7)	1,094 (25.7)	1,885 (30.4)	399 (28.1)	209 (31.6)	4,706 (28.4)	
% With < 12 y education[c]							< .0001
≤ 9.38%	841 (20.8)	937 (22)	1,135 (18.3)	250 (17.6)	108 (16.3)	3,271 (19.7)	
9.39%-17.23%	979 (24.2)	1,019 (24)	1,441 (23.3)	343 (24.1)	145 (21.9)	3,927 (23.7)	
17.24%-28.46%	978 (24.2)	1,007 (23.7)	1,540 (24.9)	360 (25.3)	163 (24.7)	4,048 (24.4)	
> 28.46%	1,246 (30.8)	1,291 (30.4)	2,078 (33.6)	469 (33)	245 (37.1)	5,329 (32.2)	
Specialty of physician performing first test[d]							<.0001
Internal medicine	320 (7.9)	1,013 (23.8)	1,453 (23.5)	292 (20.5)	0	3,078 (18.6)	
Pulmonary	1,576 (39)	2,105 (49.5)	3,154 (50.9)	811 (57)	0	7,646 (46.1)	
General surgery	299 (7.4)	137 (3.2)	222 (3.6)	53 (3.7)	0	711 (4.3)	
Thoracic surgery	1,560 (38.6)	254 (6)	294 (4.8)	52 (3.7)	0	2,160 (13)	
Other	19 (0.5)	469 (11)	708 (11.4)	135 (9.5)	0	1,331 (8)	
Unknown	270 (6.7)	276 (6.5)	363 (5.9)	79 (5.6)	661 (100)	1,649 (10)	
Tumor grade							< .0001
Well differentiated	92 (2.3)	165 (3.9)	186 (3)	< 11 (< 0.8)	13 (2)	< 467 (< 3)	
Moderately differentiated	722 (17.9)	1,370 (32.2)	1,163 (18.8)	< 11 (< 0.8)	70 (10.6)	<3336 (< 21)	
Poorly differentiated	1,418 (35.1)	1,869 (43.9)	2,139 (34.5)	545 (38.3)	184 (27.8)	6,155 (37.1)	
Unknown	1,812 (44.8)	850 (20)	2,706 (43.7)	868 (61)	394 (59.6)	6,630 (40)	
AJCC stages[e]							< .0001
IIA	392 (9.7)	496 (11.7)	251 (4.1)	78 (5.5)	24 (3.6)	1,241 (7.5)	
IIB	560 (13.9)	1,260 (29.6)	660 (10.7)	141 (9.9)	50 (7.6)	2,671 (16.1)	
IIIA	2,611 (64.6)	2,270 (53.4)	4,300 (69.4)	1,014 (71.3)	329 (49.8)	10,524 (63.5)	
IIIB	481 (11.9)	228 (5.4)	983 (15.9)	189 (13.3)	258 (39)	2,139 (12.9)	

(Continued)

TABLE 1] *(Continued)*

Variable	Evaluation Consistent With Guidelines, NSCLC, Mediastinal Sampling Done First	Evaluation Inconsistent With Guidelines, NSCLC Present, Mediastinal Sampling Performed on the Second or Later Biopsy	Evaluation Not Consistent With Guidelines, NSCLC Present, Mediastinal Sampling Never Done	Small Cell Present, Guideline-Inconsistent, No Mediastinal Sampling on First Test	No Evaluation Recorded[a]	Total	*P* Value, χ^2
T stages							< .0001
T1[f]	1,312 (32.4)	1,120 (26.3)	1,147 (18.5)	340 (23.9)	145 (22.0)	4,064 (24.5)	
T2	1,962 (48.5)	2,614 (61.5)	3,561 (57.5)	769 (54.1)	295 (44.6)	9,201 (55.5)	
T3	337 (8.3)	377 (8.9)	1,042 (16.8)	124 (8.7)	92 (13.9)	1,972 (11.9)	
Unknown	433 (10.7)	143 (3.4)	444 (7.2)	189 (13.3)	129 (19.5)	1,338 (8.1)	
N stages							< .0001
N1	1,023 (25.3)	1,926 (45.3)	1,132 (18.3)	234 (16.5)	93 (14.1)	4,408 (26.6)	
N2	2,540 (62.8)	2,100 (49.4)	4,079 (65.9)	999 (70.3)	310 (46.9)	10,028 (60.5)	
N3	481 (11.9)	228 (5.4)	983 (15.9)	189 (13.3)	258 (39)	2,139 (12.9)	
Cancer type							
NSCLC	3,311 (81.9)	4,254 (100)	6,194 (100)	0	527 (79.7)	14,286 (86.2)	< .0001
Small cell	733 (18.1)	0	0	1,422 (100)	134 (20.3)	2,289 (13.8)	

Data are presented as No. (%). AJCC = American Joint Committee on Cancer; NSCLC = non-small cell lung cancer; SEER = Surveillance, Epidemiology, and End Results; TCR = Texas Cancer Registry.
[a]No evaluation recorded means that there are no Medicare payments noted. For example, a patient who had his or her care delivered through the Veteran's Administration would not show up in the Medicare dataset.
[b]Percentage of the population in the patient's census tract living below the poverty level. Note that this does not mean the patient's income is below the poverty line, just that he or she is living in a census tract with that level of poverty.
[c]Percentage of the population in the patient's census tract that did not graduate high school.
[d]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 scan-guided biopsy, this was the referring physician. Internal medicine includes family practice and all subspecialties of internal medicine other than oncology and pulmonary medicine physicians. Surgery includes all other subspecialties of surgery other than thoracic or cardiothoracic surgery. Thoracic surgery and cardiothoracic surgery are included under thoracic surgery.
[e]Strata with cell counts of ≤ 10 patients and its row margins were suppressed as per National Cancer Institute policy and are reported as "< 11" to ensure confidentiality. A total of 115 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.
[f]Stage from SEER data is based on all TNM data; thus, for patients who had mediastinal sampling, this is based on tissue. For patients who never had sampling, it is clinical-radiographic N stage.

TABLE 2] Initial Invasive Testing Procedures Used in Patients With Lung Cancer; Regional Spread From 2004 to 2013

First Test	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Guideline-inconsistent				
CT scan-guided biopsy	6,161	38.7	6,161	38.7
Bronchoscopy without TBNA	5,703	35.8	11,864	74.6
Guideline-consistent				
Bronchoscopy with TBNA	1,192	7.5	13,056	82.1
Bronchoscopy with TBNA + EBUS or EUS[a]	954	6.0	14,010	88.0
Mediastinoscopy alone	824	5.2	14,834	93.2
Mediastinoscopy + one of the following thoracotomy-related surgeries				
Any major resection[b]	127	0.8	14,961	94.0
Wedge resection or wedge resection plus other biopsies of pleura or mediastinum	15	0.1	14,976	94.1
Only biopsies of pleura or mediastinum	20	0.1	14,996	94.2
Thoracotomy, one of the following types:				
Any major resection[b]	745	4.7	15,741	98.9
Wedge resection only	68	0.4	15,809	99.3
Wedge resection plus other biopsies of pleura or mediastinum	43	0.3	15,852	99.6
Only biopsies of pleura or mediastinum	62	0.4	15,914	100

EBUS = endobronchial ultrasound; EUS = endoscopy with ultrasound-guided needle aspiration; TBNA = transbronchial needle aspiration.
[a]Strata with ≤ 10 patients were suppressed as per National Cancer Institute policy and are reported as "< 11" to ensure confidentiality. EUS was performed in < 11 patients and was therefore included in the bronchoscopy with TBNA + EBUS category to protect patient confidentiality.
[b]Major resection is anything more than a wedge resection. This is defined as a segmentectomy, lobectomy, pneumonectomy, tracheal resection, or chest wall resection.

Mediastinal sampling occurred in 811 (68%) of the third invasive tests (Table 3). When mediastinal sampling was performed, surgical methods were used in 88% of cases, and bronchoscopic approaches were used in 12%.

From 2004 to 2013, utilization of bronchoscopic methods of mediastinal sampling increased (Fig 2), predominantly because of increased utilization of EBUS-TBNA. EBUS-TBNA utilization as the first invasive diagnostic test increased from 2% in 2007 to 18% in 2013 (Cochran-Armitage trend test, $P < .001$). EBUS-TBNA utilization at any point in the diagnostic evaluation (first test or later) increased from 4% in 2007 to 25% in 2013 ($P < .001$).

However, of the 14,492 patients eventually found to have NSCLC, 6,194 (43%) never had mediastinal lymph node sampling of any type. The clinical stage of the patient was associated with the probability of having mediastinal lymph node staging performed at any time. Patients with stage IIIA and IIIB disease were less likely to have mediastinal sampling than patients with stage II disease (50% vs 38% vs 74%; $P < .001$) (Table 4).

Complications

On a per-procedure basis, the incidence of complications was generally similar between guideline-consistent and guideline-inconsistent groups for CT scan-guided biopsy, bronchoscopy with TBNA, mediastinoscopy, and thoracotomy (Table 5). The incidence of pneumothorax and respiratory failure following bronchoscopy without TBNA was higher in the guideline-consistent group compared with the guideline-inconsistent groups ($P = .009$ and $P = .029$, respectively); the incidence of pneumothorax following CT scan-guided biopsy was higher in the guideline-inconsistent group with sampling performed second or later ($P < .001$).

On a per-patient basis, the cumulative incidence of complications during the diagnostic evaluation was lower in patients receiving guideline-consistent care with mediastinal sampling first compared with patients who had mediastinal sampling performed second or later (Table 6). Mediastinal sampling first was associated with fewer pneumothoraxes, chest tubes, hemorrhages, and

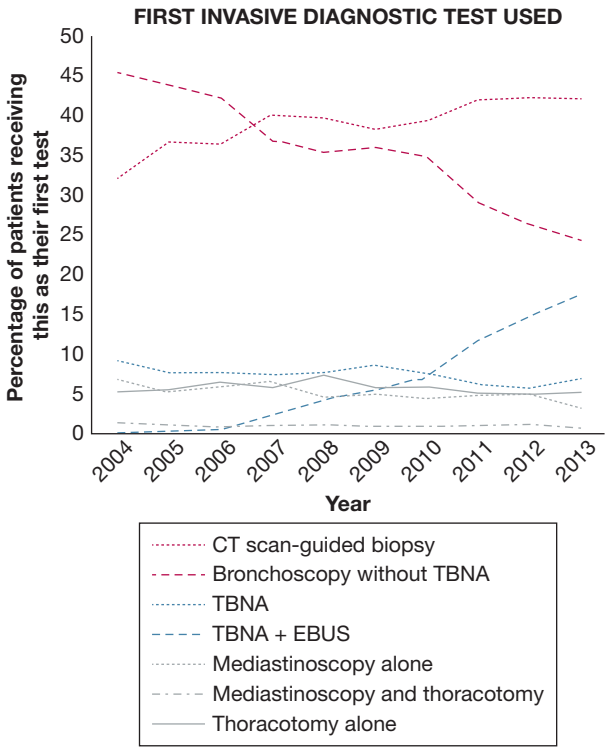

Figure 2 – *First invasive test used. These data are for the entire study cohort, which includes SEER and Texas Cancer Registry patients from 2004 to 2013. EBUS = endobronchial ultrasound; TBNA = transbronchial needle aspiration. See Figure 1 legend for expansion of other abbreviation.*

episodes of respiratory failure than guideline-inconsistent care with mediastinal sampling performed on the second or later biopsy ($P < .001$).

Propensity matching yielded 3,048 well-matched pairs of patients with lung cancer who did and did not receive guideline-consistent care, all of whom eventually underwent mediastinal sampling (e-Table 3). Compared with propensity-matched control subjects, patients who had guideline-consistent care with mediastinal sampling first had a lower cumulative incidence of complications during the diagnostic evaluation than similar patients who had guideline-inconsistent care with sampling performed second or later (Table 7). In multivariable analysis (Table 8) for the outcome of any of these complications on a per-patient basis, guideline-consistent care was associated with a lower risk of complications (OR, 0.37; 95% CI, 0.33-0.42; $P < .001$).

There were differences between groups in terms of invasive test selection and complications from particular types of tests. Patients receiving guideline-consistent care underwent fewer CT scan-guided biopsies than those who had mediastinal sampling performed second or later (10% vs 75%; $P < .001$) (Table 5) and fewer bronchoscopies without TBNA (2% vs 52%; $P < .001$). Among patients with stage III disease, those who received guideline-consistent care had fewer major therapeutic thoracotomy resections than guideline-inconsistent care patients (518 of 3,092 [17%] vs 1,154 of 2,498 [46%], respectively; $P < .001$). There were far fewer CT scan-guided, biopsy-related pneumothoraxes in the guideline-consistent group than in the group with mediastinal sampling performed second or later (78 of 4,044 vs 788 of 4,254; $P < .001$). Similarly, there were fewer pneumothoraxes from bronchoscopy without TBNA in the guideline-consistent group vs the mediastinal sampling second group (< 11 of 4,044 vs 60 of 4,254; $P < .001$). The guideline-consistent group had fewer thoracotomy-related hemorrhages (88 of 4,044 vs 144 of 4,254; $P < .001$) and fewer thoracotomy-related respiratory failure events (49 of 4,044 vs 111 of 4,254; $P < .001$) than patients who had mediastinal sampling performed second or later. It is worth noting that when considering thoracotomy-related complications, complications from major resections were only counted if patients had stage III disease (Fig 3). Complications from diagnostic/staging thoracotomies were always counted, but thoracotomy with major therapeutic resections in stage II disease did not contribute complications.

Patients who never had mediastinal lymph node sampling performed had more CT scan-guided biopsies than patients receiving guideline-consistent care (72% vs 10%; $P < .001$) (Table 5). Pneumothoraxes from CT scan-guided biopsy procedures were more frequent as well (845 of 6,194 vs 78 of 4,044; $P < .001$). Patients who never had mediastinal lymph node sampling never had surgery and thus never experienced any surgery-related complications.

Bronchoscopic Mediastinal Sampling vs Surgical Mediastinal Sampling

Patients who had mediastinal sampling performed were stratified based on whether the first method of mediastinal sampling was bronchoscopic or surgical, and whether mediastinal sampling was done as the first invasive test vs the second or later test (e-Table 4, Table 9).

Among patients who had mediastinal sampling performed first (Table 9), bronchoscopic mediastinal sampling techniques were associated with a higher

TABLE 3] Invasive Test Performed Among Patients With NSCLC Who Did Not Have Mediastinal Sampling Performed First

Variable	Stage II NSCLC				Stage III NSCLC			
	Frequency		Cumulative Frequency		Frequency		Cumulative Frequency	
	No.	%	No.	%	No.	%	No.	%
Second test performed when first test did not sample the mediastinum								
CT scan-guided biopsy	441	22.1	441	22.1	1,275	33.7	1,275	33.7
Bronchoscopy without TBNA	180	9.0	621	31.2	564	14.9	1,839	48.6
TBNA	18	0.9	639	32.1	91	2.4	1,930	51.0
TBNA + EBUS	68	3.3	707	35.4	222	5.8	2,152	56.8
Mediastinoscopy alone	180	9.0	887	44.5	710	18.8	2,862	75.7
Mediastinoscopy and one of the following thoracotomy surgeries								
Any major resection[a]	167	8.4	1,054	52.9	132	3.5	2,994	79.1
Wedge resection or wedge resection plus other biopsies of pleura or mediastinum[b]	< 11	< 0.6	…	…	< 37	< 1.0	…	…
Only biopsies of pleura or mediastinum	< 11	< 0.6	1,066	53.5	< 37	< 1.0	3,031	80.1
Thoracotomy, one of the following types								
Any major resection[a]	879	44.1	1,945	97.6	644	17.0	3,675	97.1
Wedge resection or wedge resection plus other biopsies of pleura or mediastinum	32	1.6	1,977	99.2	57	1.5	3,732	98.7
Only biopsies of pleura or mediastinum	16	0.8	1,993	100	51	1.3	3,783	100
Third test done when first and second test did not sample the mediastinum								
CT scan-guided biopsy	66	15.3	66	15.3	197	26.1	197	26.1
Bronchoscopy without TBNA	26	6.0	92	21.3	86	11.4	283	37.5
TBNA or TBNA + EBUS[b]	23	5.3	115	26.6	72	9.5	355	47.1
Mediastinoscopy alone	29	6.7	144	33.3	171	22.7	526	69.8
Mediastinoscopy and thoracotomy[b]	43	10.0	187	43.3	45	6.0	571	75.7
Thoracotomy, one of the following types:								
Any major resection[a]	233	53.9	420	97.2	150	19.9	721	95.6
Wedge resection or wedge resection plus other biopsies of pleura or mediastinum	< 11	< 2.5	425	98.4	21	2.8	742	98.4
Only biopsies of pleura or mediastinum	< 11	< 2.5	432	100	12	1.6	754	100

(Continued)

TABLE 3 *(Continued)*

Variable	Stage II NSCLC				Stage III NSCLC			
	Frequency		Cumulative Frequency		Frequency		Cumulative Frequency	
	No.	%	No.	%	No.	%	No.	%
Fourth to sixth tests done when first, second, and third tests did not sample the mediastinum								
CT scan-guided biopsy or bronchoscopy without TBNA[b]	15	25.9	15	25.9	57	42.2	57	42.2
TBNA or TBNA + EBUS	0	0	0	0	12	8.9	69	51.1
Mediastinoscopy, thoracotomy, or both[b]	43	74.1	58	100	66	48.9	135	100

See Table 1 and 2 legends for expansion of abbreviations.
[a]Major resection is anything more than a wedge resection. This is defined as a segmentectomy, lobectomy, pneumonectomy, tracheal resection, or chest wall resection.
[b]Some strata have been combined to keep cell counts ≥ 11 as per National Cancer Institute policy to ensure confidentiality.

cumulative incidence of pneumothoraxes (5.7% vs 3.6%; $P = .001$) but a lower incidence of pulmonary hemorrhages (2.3% vs 4.2%; $P < .001$) than surgical mediastinal sampling. Among patients who had mediastinal sampling performed second or later, bronchoscopic mediastinal sampling was associated with a lower incidence of pneumothoraxes (18.5% vs 21.9%; $P = .045$), pulmonary hemorrhages (< 2.2% vs 6.4%; $P < .001$), and respiratory failure (< 2.2% vs 3.9%; $P = .007$) than surgical mediastinal sampling.

Surgical mediastinal sampling as the first test was guideline consistent during this time period, but mediastinal sampling second or later with bronchoscopic techniques was still better in some situations (Table 9). Specifically, bronchoscopic sampling performed second or later was associated with a lower incidence of pulmonary hemorrhage (< 2.2% vs 4.2%; $P = .01$) but a higher incidence of pneumothorax (18.5% vs 3.6%; $P < .001$) than surgical sampling performed first.

The differences in complications on a per-patient basis were due to differences in the frequency of tests used and the types of tests performed (e-Table 4) rather than to major differences in the incidence of complications on a per- procedure basis. The incidence of complications on a per-procedure basis was similar between TBNA/EBUS and surgical groups for all procedures except mediastinoscopy, which was associated with a higher incidence of respiratory failure and hemorrhage per procedure in the EBUS/TBNA sampling first group, and pneumothorax requiring chest tube, which was associated with a higher incidence in the surgical group.

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 for each patient. Patients who had guideline-consistent care had fewer diagnostic tests ($P < .001$) (Table 10) than patients who had guideline-inconsistent care with mediastinal sampling performed second. Propensity matching did not significantly change these results.

Time to Complete Diagnosis and Staging and Time to Treatment

Guideline-consistent strategies with mediastinal sampling performed first were associated with more rapid completion of the diagnostic and staging evaluation than guideline-inconsistent strategies with

TABLE 4] Practice Patterns and Guideline Consistency Stratified According to Stage in Patients With NSCLC for SEER, 2004 to 2013

Variable	Evaluation Consistent With Guidelines, Sampling Done First	Evaluation Inconsistent With Guidelines, NSCLC Present, Mediastinal Sampling Done on the Second or Later Biopsy	Evaluation Not Consistent With Guideline, NSCLC Present, Mediastinal Sampling Never Done	Total
Stage IIA or IIB	869 (24.6)	1,755[a] (49.7)	911 (25.8)	3,535
Stage IIIA	1,783 (22.6)	2,153[a] (27.3)	3,961 (50.2)	7,897
Stage IIIB	336 (24.0)	197[a] (14.1)	864 (61.9)	1,397
Total	2,988 (23.3)	4,105 (32.0)	5,736 (44.7)	12,829

Data are presented as No. (%). See Table 1 legend for expansion of abbreviations.
[a]$P < .001$ for comparison of stage vs practice pattern.

mediastinal sampling performed second or later (median time from first to last invasive test, 0 vs 34 days; $P < .001$) (Table 11). Guideline-consistent care also was associated with more rapid treatment than guideline-inconsistent strategies (median time from first test to treatment, 18 days vs 41 days; $P < .001$). Propensity matching did not significantly change this outcome.

Diagnostic Strategy, Subsequent Treatment, and Survival

Among patients with NSCLC, stage II patients who had mediastinal sampling were more likely to have surgery as part of their treatment (85% vs 93% vs 6% for guideline-consistent, guideline-inconsistent sampling second, and mediastinal sampling never performed, 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 patients who never had sampling (37% vs 63% vs 4%; $P < .001$). Patients with NSCLC who underwent mediastinal sampling survived longer than patients who never had mediastinal sampling ($P < .0001$) (Fig 5). Survival was similar for patients with stage II NSCLC, but for stage III NSCLC, survival was longer in patients who had mediastinal sampling performed second or later ($P < .001$). The results were similar even after propensity matching (details are given in e-Fig 1).

Discussion

A quality gap can be defined as the difference between the outcomes and processes found in practice and those obtainable using the best current knowledge.[15,16] The current study found that although previously identified quality gaps in the diagnostic staging of lung cancer[7,8] have improved with the introduction of EBUS, significant gaps remain. Mediastinal sampling among patients with NSCLC with lymph node involvement without distant metastases (T1-3N1-3M0), whether done first or subsequently, improved from 2004 to 2013, increasing from 54% to 64% of patients ($P < .001$). Delivery of guideline-consistent care with mediastinal lymph node sampling performed first also improved, increasing from 23% to 34%. Guideline-inconsistent strategies that sampled the mediastinum second or later were associated with unnecessary testing, more major resections in patients who had stage III disease, more complications, and longer time to treatment.

We also found that bronchoscopic methods of mediastinal lymph node sampling were associated with significantly fewer complications than surgical methods. We cross-validated our findings by using two large independently collected datasets. Results were similar between datasets and across multiple regions of the country. The data suggest that a guideline-consistent strategy with mediastinal lymph node sampling performed first by using EBUS-TBNA will result in fewer tests and complications than strategies that perform biopsies on peripheral lesions first, such as CT scan-guided biopsy, and fewer complications than strategies that use surgical mediastinal staging as the first test. The study identified three common practice errors: improper sequencing of invasive tests, failing to sample the mediastinum, and overuse of thoracotomy.

Multiple guidelines address the issue of invasive test sequencing, and all recommend biopsy of the mediastinal lymph nodes first in patients with evidence of nodal disease rather than biopsy of the peripheral mass.[2-4,13,17-21] This recommendation is despite the fact that CT scan-guided biopsy of peripheral lesions has a sensitivity of 90% compared with a sensitivity of 34% to 63% for conventional bronchoscopy.[4,21] The reasoning behind the

TABLE 5] Procedure Utilization and Incidence of Complications According to Procedure

InvasiveDiagnostic Test and Associated Complications	Evaluation Consistent With Guidelines, Mediastinal Sampling Done First	Evaluation Inconsistent With Guidelines, NSCLC, Mediastinal Sampling on the Second or Later Biopsy	Evaluation Inconsistent With Guidelines, NSCLC Present, Mediastinal Sampling Never Done	Small Cell Present, Inconsistent With Guidelines, No Mediastinal Sampling on First Test	Frequency of Test Utilization *P* Value[a]	Incidence of Complications *P* Value[b]
Total No. of patients	4,044	4,254	6,194	1,422		
Procedures						
CT scan-guided biopsy[c]	392 (10)	3,171 (75)	4,469 (72)	805 (57)	< .001	
Complications within 1 d						
Pneumothorax[b]	78 (20)	788 (25)	845 (19)	157 (20)		< .001
Pneumothorax requiring chest tube[b]	14 (3.6)	151 (4.8)	209 (4.7)	34 (4.2)		.612
Hemorrhage[b]	< 11 (< 2.8)[d]	34 (1.1)	44 (1)	< 11 (< 7)[d]		.892
Respiratory failure[b]	< 11 (< 2.8)[d]	< 11 (<0.3)[d]	< 11 (< 0.2)[d]	0		.109
Bronchoscopy without TBNA[c]	101 (2)	2,224 (52)	3,433 (55)	974 (68)	< .001	
Complications within 1 d						
Pneumothorax	< 11 (< 11)[d]	60 (2.7)	72 (2.1)	11 (1.1)		.009
Pneumothorax requiring chest tube	< 11 (< 11)[d]	12 (0.5)	11 (0.3)	< 11 (< 1.1)[d]		.184
Hemorrhage	< 11 (< 11)[d]	24 (1)	36 (1)	< 11 (< 1.1)[d]		.461
Respiratory failure	< 11 (< 11)[d]	13 (0.6)	33 (1)	< 11 (< 1.1)[d]		.029
Bronchoscopy with TBNA[c]	2,393 (59)	530 (12)	…	46 (3.2)	< .001	
Complications within 1 d						
Pneumothorax	40 (1.7)	< 11 (< 2)[d]	…	0		.444
Pneumothorax requiring chest tube	< 11 (0.5)	< 11 (< 2)[d]	…	0		.701
Hemorrhage	17 (0.7)	< 11 (< 2)[d]	…	0		.226
Respiratory failure	25 (1)	< 11 (< 2)[d]	…	0		.808
Mediastinoscopy alone (no thoracotomy)[c]	1,228 (30)	1,431 (34)	…	96 (6.8)	< .001	
Complications within 14 d			…			
Pneumothorax	65 (5.3)	82 (5.7)	…	< 11 (< 11)[d]		.671
Pneumothorax requiring chest tube	< 11 (0.9)	< 11 (< 0.8)[d]	…	< 11 (< 11)[d]		.990
Hemorrhage	19 (1.5)	45 (3.1)	…	< 11 (< 11)[d]		.008
Respiratory failure	19 (1.5)	22 (1.5)	…	< 11 (< 11)[d]		.990
Thoracotomy, one of the following types of surgery:[c]	1,527 (38)	3,029 (71)	…	61 (4.3)	< .001	
Complications within 14 d						

(Continued)

TABLE 5] *(Continued)*

InvasiveDiagnostic Test and Associated Complications	Evaluation Consistent With Guidelines, Mediastinal Sampling Done First	Evaluation Inconsistent With Guidelines, NSCLC, Mediastinal Sampling on the Second or Later Biopsy	Evaluation Inconsistent With Guidelines, NSCLC Present, Mediastinal Sampling Never Done	Small Cell Present, Inconsistent With Guidelines, No Mediastinal Sampling on First Test	Frequency of Test Utilization *P* Value[a]	Incidence of Complications *P* Value[b]
Any major resection[e]	1,253	2,714	…	35	< .001	
Hemorrhage	63 (5)	120 (4.4)	…	0		.416
Respiratory failure	34 (2.7)	91 (3.4)	…	0		.328
Wedge resection or wedge resection plus other biopsies of pleura or mediastinum	156	149		13	< .001	
Hemorrhage	< 25 (< 16)	< 24 (< 16)		<11		.849
Respiratory failure	< 11	< 20 (< 13)		<11		.661
Only biopsies of pleura or mediastinum	118	166	…	13	< .001	
Hemorrhage	< 25 (< 21)	< 24 (< 14)	…	0		.635
Respiratory failure	< 11 (< 9)	< 20 (< 12)	…	0		.990

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

[a]*P* value using the χ^2 test compares the frequency of testing utilization between groups on a per-patient basis. For bronchoscopy with TBNA, mediastinoscopy alone, and thoracotomy, the comparison is only between guideline-consistent and guideline-inconsistent, with sampling performed second or later.

[b]*P* value using the Fisher exact test is for comparing incidence of complications on a per-procedure basis between guideline-consistent, guideline-inconsistent with sampling performed second or later, and guideline-inconsistent sampling never done.

[c]Percentage reflects percentage of patients in that group (guideline-consistent, guideline-inconsistent with NSCLC sampling performed second or later, guideline-inconsistent with NSCLC sampling never done, and guideline-inconsistent with small cell) that had the test done.

[d]Strata with ≤ 10 patients were suppressed as per National Cancer Institute policy and are reported as "< 11" to ensure confidentiality. In other cases, aggregate numbers are used. For example, there were 25 hemorrhages associated with thoracotomy involving (wedge resection or wedge plus other biopsies) or (other biopsies of pleura) in the guideline-consistent group. The total hemorrhage complications for thoracotomy (all types) in the guideline-consistent group is 88 as noted in the text.

[e]Major resection is anything more than a wedge resection. This is defined as a segmentectomy, lobectomy, pneumonectomy, tracheal resection, or chest wall resection.

TABLE 6] Cumulative Incidence of Complications During the Entire Diagnostic Evaluation per Patient

Variable	Evaluation Consistent With Guidelines, NSCLC Present, Mediastinal Sampling Done First (n = 4,044)	Evaluation Inconsistent With Guidelines, NSCLC Present, Mediastinal Sampling Done on the Second or Later Biopsy (n = 4,254)	Evaluation Inconsistent With Guideline, NSCLC Present, Mediastinal Sampling Never Done (n = 6,194)	Small Cell Present, Inconsistent With Guidelines, No Mediastinal Sampling on First Test (n = 1,422)	*P* Value Guideline-Consistent vs Other Groups
Pneumothorax	4.7	22	14.8	12.2	< .001[a]
Pneumothorax requiring chest tube	0.8	4.1	3.6	2.5	< .001[a]
Hemorrhage	3.2	5.8	1.3	1.6	< .001[a]
Respiratory failure	2.4	3.6	0.6	0.1	< .001[a]

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.
[a]Significant difference between guideline-consistent group with NSCLC and guideline-inconsistent groups with NSCLC with mediastinal sampling performed on the second or later invasive test, and guideline-inconsistent group without sampling (all, $P < .001$ using the Kruskal-Wallis test).

guidelines is that treatment decisions depend on complete staging information. In patients with clinical-radiographic T1-3N1-3M0 disease, mediastinal lymph node sampling is required for staging. In such circumstances, if the first test is a CT scan-guided biopsy of the peripheral lesion, any finding of lung cancer will require a second procedure to stage the mediastinum. Thus, a strategy of bronchoscopic mediastinal lymph node sampling with peripheral biopsy as needed to both diagnose and stage the patient at one time is likely to be superior when cancer is present.[6,8] This study confirms the findings from previous studies that sampling the mediastinum first is indeed more effective, resulting in fewer tests and complications.[6,8,22]

The second quality gap identified in this study was failing to sample the mediastinum at all. This problem was more common in patients with stage IIIA and IIIB cancer, suggesting that physicians are frequently relying on imaging as the sole method to stage the mediastinum. The consequence of overreliance on imaging is that some patients will be falsely upstaged, leading to missed opportunities for surgery and possibly cure. Conversely, patients may be falsely understaged, leading to unnecessary thoracotomies and complications.[1,3,7,8] This problem of underutilization of both bronchoscopic and surgical mediastinal lymph node sampling has been previously described.[7,8,23-25]

EBUS-TBNA has revolutionized mediastinal lymph node staging and is now common as part of pulmonary fellowship training, but how far it has penetrated into community practice has not been well studied.[26] Similarly, the American Board of Thoracic Surgery has designated a minimum number of EBUS procedures for residents in the thoracic track to

TABLE 7] Propensity-Matched Cohorts: Cumulative Incidence of Complications During the Entire Diagnostic Evaluation per Patient

Propensity-Matched Cohorts	Evaluation Consistent With Guidelines, NSCLC Present, Mediastinal Sampling Done First (n = 3,048)	Evaluation Inconsistent With Guidelines, NSCLC Present, Mediastinal Sampling Done on the Second or Later Biopsy (n = 3,048)	*P* Value Guideline-Consistent vs Other Groups
Pneumothorax	5.1	22	< .001[a]
Pneumothorax requiring chest tube	0.9	4.4	< .001[a]
Hemorrhage	3.5	5.8	< .001[a]
Respiratory failure	2.7	3.7	.047[a]

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.
[a]Significant difference between guideline-consistent group with NSCLC and guideline-inconsistent groups with NSCLC with mediastinal sampling performed on the second or later invasive test (all, $P < .001$ using Kruskal-Wallis test).

TABLE 8] Factors Associated With Complications During the Diagnostic Evaluation

Variable	Univariable		Multivariable[a]	
	OR (95% CI)	*P* Value	OR (95% CI)	*P* Value
Age group				
66-70 y	Ref		…	…
71-75 y	1.14 (1.02-1.27)	.019	…	…
76-80 y	1.06 (0.95-1.19)	.296	…	…
> 80	1.21 (1.06-1.37)	.003	…	…
Sex				
Male	Ref		…	…
Female	1.05 (0.97-1.14)	.223	…	…
Race				
Non-Hispanic white	Ref		Ref	
Hispanic	1.08 (0.90-1.31)	.405	1.05 (0.86-1.28)	.659
Non-Hispanic black	0.81 (0.69-0.96)	.014	0.81 (0.68-0.96)	.016
Non-Hispanic other	1.21 (0.98-1.48)	.070	1.07 (0.85-1.34)	.578
Year of diagnosis				
2004	Ref		…	…
2005	1.14 (0.95-1.37)	.155	…	…
2006	0.98 (0.81-1.18)	.794	…	…
2007	1.11 (0.92-1.33)	.289	…	…
2008	1.13 (0.94-1.36)	.191	…	…
2009	1.18 (0.98-1.42)	.089	…	…
2010	1.24 (1.03-1.51)	.024	…	…
2011	1.09 (0.90-1.32)	.384	…	…
2012	0.97 (0.79-1.18)	.730	…	…
2013	1.01 (0.83-1.23)	.921	…	…
Region				
California	Ref		Ref	
Connecticut	0.93 (0.76-1.14)	.479	1.04 (0.85-1.28)	.701
Detroit	1.14 (0.95-1.37)	.151	1.28 (1.06-1.55)	.011
Georgia	0.67 (0.57-0.79)	< .001	0.77 (0.66-0.91)	.002
Hawaii	1.04 (0.68-1.59)	.868	1.00 (0.63-1.58)	.990
Iowa	0.80 (0.64-1.00)	.046	0.82 (0.65-1.03)	.082
Kentucky	0.67 (0.56-0.80)	< .001	0.70 (0.58-0.84)	< .001
Louisiana	0.69 (0.56-0.84)	< .001	0.69 (0.56-0.85)	< .001
New Jersey	0.77 (0.66-0.89)	.001	0.86 (0.74-1.01)	.064
New Mexico	1.08 (0.76-1.53)	.671	1.07 (0.75-1.53)	.721
Seattle	1.04 (0.84-1.28)	.703	1.14 (0.92-1.41)	.230
Texas	0.87 (0.77-0.99)	.037	0.95 (0.83-1.09)	.470
Utah	1.37 (0.87-2.14)	.170	1.54 (0.97-2.44)	.067
T stage				
T1	Ref		Ref	
T2	0.71 (0.64-0.78)	< .001	0.63 (0.57-0.69)	< .001
T3	0.54 (0.47-0.63)	< .001	0.47 (0.40-0.55)	< .001
Unknown	0.36 (0.29-0.44)	< .001	0.40 (0.32-0.49)	< .001

(Continued)

TABLE 8] *(Continued)*

Variable	Univariable		Multivariable[a]	
	OR (95% CI)	*P* Value	OR (95% CI)	*P* Value
Guideline consistent				
No	Ref		Ref	
Yes	0.39 (0.35-0.44)	< .001	0.37 (0.33-0.42)	< .001
Cancer type				
NSCLC	Ref		Ref	
Small cell	0.58 (0.50-0.67)	< .001	0.65 (0.56-0.75)	< .001
Charlson Comorbidity Index score				
0	Ref		Ref	
1	1.11 (1.00-1.22)	.043	…	…
≥ 2	1.07 (0.97-1.19)	.193	…	…

See Table 1 legend for expansion of abbreviation.
[a]Backward selection with $P < .05$ to remain in the model.

perform to be eligible for the boards. This study is among the first to quantify changes in EBUS-TBNA utilization patterns at the national level. It shows that while EBUS use has increased and more patients are receiving guideline-consistent care, incorrect test sequencing and underutilization of mediastinal sampling persist.

The third quality gap identified was overutilization of thoracotomy. During the time span of this study,

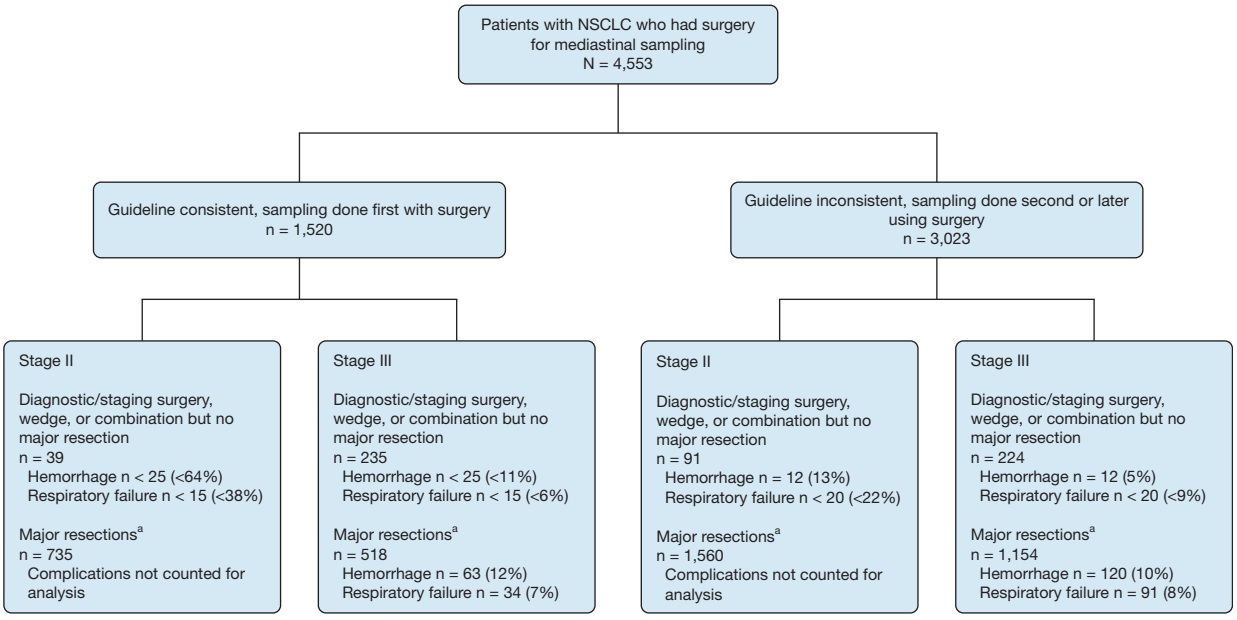

Figure 3 – *Types of thoracotomy surgeries and their complications recorded in guideline-consistent vs guideline-inconsistent care stratified according to disease stage.* [a]*Major resection is defined as any thoracotomy with segmentectomy, lobectomy, pneumonectomy, tracheal resection, or chest wall resection. Diagnostic/staging surgery is defined as thoracotomy not involving any major resection. If a major therapeutic resection was performed and the patient had stage II disease, the complications were not counted. If a major therapeutic resection was performed and the patient had stage III disease, complications were counted. N is the number of patients for each group, and n is the number with that complication. Strata with ≤ 10 patients were suppressed as per National Cancer Institute policy and are reported as "< a number" to ensure confidentiality. For the guideline-consistent group, there were a total of 88 hemorrhage complications (63 arising from major resections in patients with stage III disease, and 25 arising from patients with stage II and III disease undergoing diagnostic/staging surgeries). There were a total of 49 respiratory failure complications (34 arising from major resections in patients with stage III disease, and 15 arising from patients with stage II and III disease undergoing diagnostic/staging surgeries). In the guideline-inconsistent group, there were a total of 111 respiratory failure complications, 91 from major resections, and 20 from patients with stage II and III disease undergoing diagnostic/staging surgeries. See Figure 1 legend for expansion of abbreviation.*

TABLE 9] Bronchoscopic vs Mediastinal Sampling Strategies: Cumulative Incidence of Complications per Patient During the Entire Diagnostic Evaluation

Method of Mediastinal Sampling	Mediastinal Sampling First	Mediastinal Sampling Second or Later
Bronchoscopic sampling	N = 2,140 PTx, 122 (5.7%)[a]	N = 502 PTx, 93 (18.5%)
	PTx requiring chest tube, 21 (1%)	PTx requiring chest tube, 14 (2.8%)
	Hemorrhage, 50 (2.3%)[b]	Hemorrhage, < 11 (< 2.2%)
	Respiratory failure, 51 (2.4%)[c]	Respiratory failure, < 11 (2.2%)
Surgical sampling	N = 1,904 PTx, 68 (3.6%)[a]	N = 3,752 PTx, 843 (21.9%)
	PTx requiring chest tube, 12 (0.6%)	PTx requiring chest tube, 159 (4.1%)
	Hemorrhage, 80 (4.2%)[b]	Hemorrhage, 239 (6.4%)
	Respiratory failure, 47 (2.5%)[c]	Respiratory failure, 145 (3.9%)

N is the total number of patients who had the corresponding method of mediastinal sampling performed first vs second or later. Within each category, the number of patients with that complication is provided, along with the percentage of patients in that cell that had this complication (n/N). PTx = pneumothorax.
[a]$P = .001$.
[b]$P < .001$.
[c]$P = .86$ comparing bronchoscopic vs surgical sampling when mediastinal sampling was done as the first test. Among patients who had mediastinal sampling second or later, bronchoscopic mediastinal sampling was associated with a lower incidence of Ptxs ($P = .045$), pulmonary hemorrhages ($P < .001$), and respiratory failure ($P < .012$) than surgical mediastinal sampling.

mediastinoscopy, bronchoscopic TBNA, and even video-assisted thoracotomy could all have been classified as guideline-consistent practices. However, for patients with a high probability of N2 or N3 disease, thoracotomy is not an optimal method of mediastinal sampling. Thoracotomy without any previous mediastinoscopy or TBNA was the method of mediastinal lymph node sampling in 1,640 of 2,627 (62%) patients with stage II NSCLC and 1,390 of 4,938 (28%) patients with stage III NSCLC who eventually had staging performed (Table 3). This finding indicates that thoracotomy without previous staging is occurring in a significant percentage of patients with stage III disease. Many of these futile thoracotomies in stage III disease were probably avoidable.[27] The definition of futile thoracotomy used here is based on previous research[28-30] that designates thoracotomy in the presence of N2 or N3 disease as futile.

It is worthwhile to distinguish these futile thoracotomies from multimodality approaches to stage IIIA disease that include surgery, because the two are not the same.[31,32] If the mediastinum is properly staged and N2 disease is detected, surgery may then have a role to play following neoadjuvant chemotherapy according to the most recent guidelines.[32] However, that scenario is not what we

TABLE 10] Number of Invasive Tests per Patient According to Practice Pattern

	Entire Cohort		Following Propensity Matching	
Practice Pattern	Mean ± SD	Median (IQR)	Mean ± SD	Median (IQR)
Evaluation consistent with guidelines, NSCLC, mediastinal sampling done first	1.4 ± 0.7	1 (1-2)[a]	1.5 ± 0.8	1 (1-2)[a]
Evaluation inconsistent with guidelines, NSCLC present, mediastinal sampling done on the second or later biopsy	2.5 ± 0.8	2 (2-3)[a]	2.5 ± 0.8	2 (2-3)[a]
Evaluation not consistent with guideline, NSCLC present, mediastinal sampling never done	1.3 ± 0.6	1 (1-1)	…	…
Small cell present	1.4 ± 0.7	1 (1-2)	…	…

IQR = interquartile range, 25% percentile to 75% percentile. See Table 1 legend for expansion of abbreviation.
[a]$P < .001$ using nonparametric comparison for guideline-consistent care vs guideline-inconsistent care with sampling performed second. Comparisons using the entire cohort are shown in the middle section. Following propensity matching, we compared guideline-consistent NSCLC with guideline-inconsistent NSCLC with sampling performed second or later (right-hand section).

TABLE 11] Time to Finish Diagnosis and Staging and Time to Treatment Stratified According to Mediastinal Sampling Strategy

	Mediastinal Sampling First: Guideline Consistent	Mediastinal Sampling Second or Later: Guideline Inconsistent	Mediastinal Sampling Never Done
Entire cohort			
Median No. of days from first to last invasive test (median, IQR)	0 (0, 5)[a,b]	34 (21, 52)[a,b]	0 (0, 0)[a,b]
Median No. of days from first invasive test to first treatment (median, IQR)	18 (0, 35)[a,b]	41 (26, 63)[a,b]	27 (15, 44)[a,b]
Propensity-matched cohorts			
Median No. of days from first to last invasive test (median, IQR)	0 (0, 7)[a]	34 (21, 52)[a]	…
Median No. of days from first invasive test to first treatment (median, IQR)	18 (0, 35)[a]	42 (27, 64)[a]	…

See Table 10 legend for expansion of abbreviation.
[a]$P < .001$, nonparametric Wilcoxon rank sum test comparing mediastinal sampling first vs second or later.
[b]$P < .001$, nonparametric Kruskal-Wallis test for all three groups.

observed, in which patients receive no mediastinal sampling prior to thoracotomy and turn out to have pN2 disease. In addition, none of these patients had chemotherapy prior to their surgery, and therefore it was not surgery following neoadjuvant chemotherapy for stage IIIA disease. Finally, during the time period of the study, guidelines did not support the routine use of neoadjuvant chemotherapy.[31]

Our findings also build on the existing body of evidence regarding the use of EBUS vs mediastinoscopy for lung cancer staging. Previous clinical research studies found that in select centers of excellence, EBUS and mediastinoscopy have comparable sensitivity and specificity.[33-35] A large meta-analysis also found indirect evidence that sensitivity of EBUS-TBNA and mediastinoscopy were similar (0.84 and 0.86, respectively), with mediastinoscopy having a higher complication rate but fewer false-negative results.[36] This analysis is the first large comparative effectiveness study to show that EBUS-TBNA as used in everyday practice is indeed associated with lower complication rates than surgical mediastinal lymph node sampling. This finding reaffirms the more recent lung cancer guidelines, published in 2013, recommending that EBUS-TBNA be the first method used to sample the mediastinum.[13] The data in this study only extend up to 2013 because, subsequent to that year, the definition of guideline-consistent care has changed, and it is possible that adherence to guidelines may be better.

We found that mediastinal sampling first was associated with lower survival in patients with stage III NSCLC than mediastinal sampling performed second. Most likely, this finding is the result of selection bias. Figure 4 shows that 32% of patients with stage III NSCLC with mediastinal sampling performed first eventually had surgical treatment, whereas 59% of patients with stage III NSCLC with mediastinal sampling performed second or later had surgical treatment. Therefore, treatments for the two groups in Figure 4 were different. It is probable that patients with good performance status were more likely to have surgery, whereas those with poor performance status and more comorbidities would undergo bronchoscopic sampling and would never undergo surgery, which would explain the observed difference in survival.

Limitations of this study include restricting the data to Medicare patients only. Thus, the results may not be generalizable to younger patients. In a similar manner, the results may not be generalizable to non-SEER registry sites. These were administrative data, and we therefore do not know whether lymph nodes were enlarged on CT scanning or positive on PET imaging. If the lymph nodes were negative on both CT and PET imaging, then mediastinal sampling in smaller peripheral nodules would not be required. Such patients would probably have a CT scan-guided biopsy and subsequently undergo surgery, and they would be misclassified as guideline inconsistent if they were found to harbor microscopic nodal disease. Although such misclassification is no doubt present, previous studies have shown PET/CT-scanned N0 patients have only a 5% to 7% incidence of N2 disease;

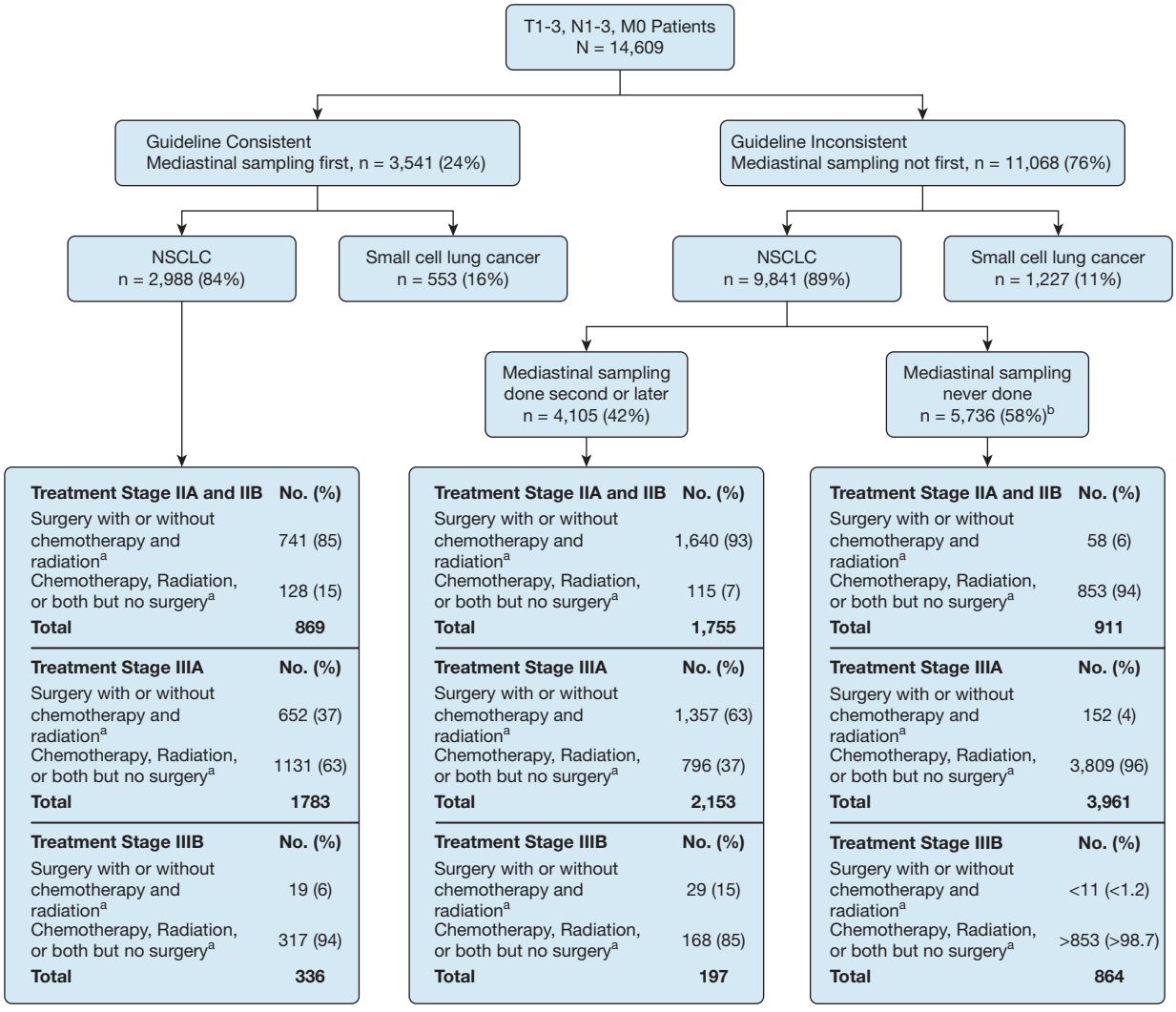

Figure 4 – *Practice patterns, diagnoses, stage of disease, and treatment patterns in the SEER database from 2004 to 2013 for which there was detailed T and N stage information.* [a]*If surgery was performed without mediastinal lymph node sampling, this was considered not consistent with guidelines. Similarly, if surgery with lymph node sampling was performed, but it was not the first test and there was no previous sampling performed, this was then classified as not consistent with guidelines.* [b]*If a patient received any type of treatment, such as chemotherapy or radiation, without previous lymph node sampling and went on to receive surgery and had lymph node sampling at that time, but they never had lymph node sampling prior to the chemotherapy or radiation, this was then considered as no lymph node sampling prior to first treatment. See Figure 1 legend for expansion of abbreviations.*

therefore, this should be a fairly infrequent occurrence.[37-39] Similarly, patients with PET/CT-scanned N1 or N2 disease eventually proven to have pathologic stage I disease do not appear in the data. On balance, although misclassification probably is present, the magnitude of the true quality gap is likely to be large. In addition, while nodal sampling should occur during resection, it does not always. We classified patients having surgery as having mediastinal sampling "first," based on the assumption that systematic lymph node sampling was always performed. However, if true systematic nodal sampling was not performed during surgery, the estimates of guideline-consistent care would then be even lower than reported.

Finally, because these are administrative data, we could not determine which patients did not receive procedures because severe comorbidities precluded them. Differences in comorbidities might also contribute to differences in outcomes. We did adjust for comorbidities by using the Charlson Comorbidity Index, but the possibility of residual confounding still exists. However, the analysis was limited to patients who eventually received treatment. If treatment can be tolerated, albeit perhaps nonsurgical, then minimally invasive staging

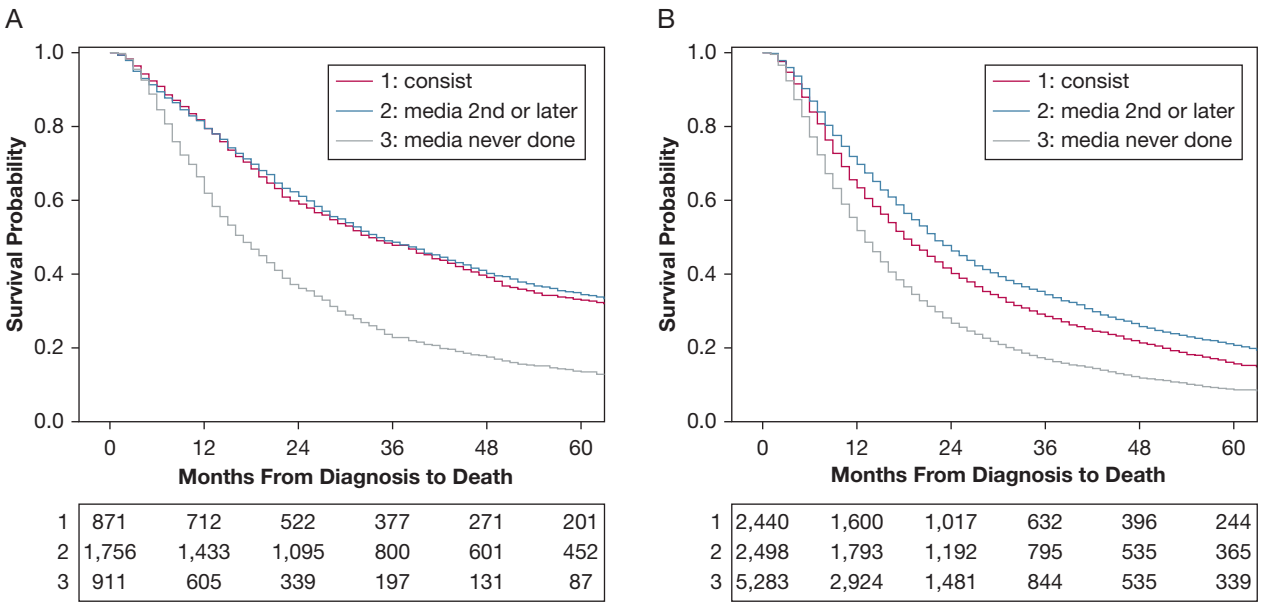

Figure 5 – *Survival in patients with NSCLC according to stage and diagnostic strategy. A, Patients with stage II NSCLC. B, Patients with stage III NSCLC. Diagnostic strategy is shown for each stage as follows: guideline-consistent care with mediastinal sampling performed first (blue), guideline-inconsistent care with mediastinal sampling performed second or later (red), and guideline-inconsistent care that never had mediastinal sampling (gray). See Figure 1 legend for expansion of abbreviation.*

procedures such as EBUS are usually warranted when there is nodal disease. The reason for the observed differences in cumulative complication rates was due to patients in the guideline-inconsistent arm undergoing more tests, in particular more CT scan-guided biopsies and more thoracotomies. However, the differences in survival between groups could be due to residual confounding, and thus the results should be interpreted with caution.

Conclusions

We found that lung cancer diagnosis and staging quality gaps have decreased but are still large. Three quality gaps were identified: (1) failure to sample the mediastinum first in 75% of patients; (2) failure to sample the mediastinum at all in 43% of patients; and (3) overreliance on thoracotomy without previous staging. These quality gaps were associated with increased complications, including pneumothorax, respiratory failure, and hemorrhage. Bronchoscopic mediastinal staging was associated with fewer complications than surgical mediastinal staging. Future quality initiatives should include a simple message: in patients with suspected lung cancer with hilar or mediastinal lymphadenopathy without evidence of distant disease, the mediastinum should be sampled first with EBUS-TBNA.

Acknowledgments

Author contributions: D. E. O. was the principal investigator and was responsible for study oversight. D. E. O. and J. N. were responsible for data analysis, with S. H. B. G., H. Z., and H. B. G. offering input. D. E. O., J. N., and H. Z. contributed to writing, editing, and review.

Financial/nonfinancial disclosures: None declared.

Other contributions: The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the Department of State Health Services (Texas) Cancer Prevention and Research Institute of Texas, or the Centers for Disease Control and Prevention. This study used the linked SEER-Medicare database. The interpretation and reporting of these data are the sole responsibility of the authors. 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.

Role of sponsors: The sponsor had no role in the design of the study, the collection and analysis of the data, or the preparation of the manuscript.

Additional information: The e-Appendix, e-Figure, and e-Tables can be found in the Supplemental Materials section of the online article.

References

1. Silvestri GA, Gould MK, Margolis ML, et al. Noninvasive staging of non-small cell lung cancer: ACCP evidenced-based clinical practice guidelines (2nd edition). *Chest*. 2007;132(suppl 3):178S-201S.
2. Almeida FA, Uzbeck M, Ost D. Initial evaluation of the nonsmall 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(suppl 3): 202S-220S.
4. Rivera MP, Mehta AC. Initial diagnosis of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest*. 2007;132(suppl 3):131S-148S.
5. Schmidt-Hansen M, Baldwin DR, Hasler E, Zamora J, Abraira V, Roque IFM. PET-CT for assessing mediastinal lymph node involvement in patients with suspected resectable non-small cell lung cancer. *Cochrane Database Syst Rev*. 2014;11:CD009519.
6. 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.
7. Ost DE, Niu J, Elting LS, Buchholz TA, Giordano SH. Determinants of practice patterns and quality gaps in lung cancer staging and diagnosis. *Chest*. 2014;145(5): 1097-1113.
8. Ost DE, Niu J, Elting LS, Buchholz TA, Giordano SH. Quality gaps and comparative effectiveness in lung cancer staging and diagnosis. *Chest*. 2014;145(2): 331-345.
9. Haponik EF, Shure D. Underutilization of transbronchial needle aspiration: experiences of current pulmonary fellows. *Chest*. 1997;112(1):251-253.
10. Smyth CM, Stead RJ. Survey of flexible fibreoptic bronchoscopy in the United Kingdom. *Eur Respir J*. 2002;19(3):458-463.
11. Prakash UB, Offord KP, Stubbs SE. Bronchoscopy in North America: the ACCP survey. *Chest*. 1991;100(6):1668-1675.
12. Colt HG, Prakash UBS, Offord KP. Bronchoscopy in North America: survey by the American Association for Bronchology, 1999. *J Bronchol Intervention Pulmonol*. 2000;7(1):8-25.
13. 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(suppl 5):e211S-e250S.
14. 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.
15. Shojania K, McDonald K, Wachter R, Owens D. Series overview and methodology. In: AHRQ, ed. *Closing the Quality Gap: A critical analysis of quality improvement strategies*. Rockville, MD: Agency for Healthcare Research and Quality; 2004:1-37.
16. 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.
17. Baldwin DR, White B, Schmidt-Hansen M, Champion AR, Melder AM. Diagnosis and treatment of lung cancer: summary of updated NICE guidance. *BMJ*. 2011;342:d2110.
18. Crino L, Weder W, van Meerbeeck J, Felip E. 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.
19. 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.
20. Ost DE, Jim Yeung SC, 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(suppl 5):e121S-e141S.
21. 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.
22. Ost DE, Niu J, Zhao H, Grosu HB, Giordano SH. Quality gaps and comparative effectiveness of management strategies for recurrent malignant pleural effusions. *Chest*. 2018;153(2):438-452.
23. 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.
24. Little AG, Rusch VW, Bonner JA, et al. Patterns of surgical care of lung cancer patients. *Ann Thorac Surg*. 2005;80(6): 2051-2056.
25. 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.
26. Tanner NT, Pastis NJ, Silvestri GA. Training for linear endobronchial ultrasound among US pulmonary/critical care fellowships. a survey of fellowship directors. 2013;143(2):423-428.
27. Annema JT, van Meerbeeck JP, Rintoul RC, et al. Mediastinoscopy vs endosonography for mediastinal nodal staging of lung cancer: a randomized trial. *JAMA*. 2010;304(20):2245-2252.
28. Smoragiewicz M, Laskin J, Wilson D, et al. Using PET-CT to reduce futile thoracotomy rates in non-small-cell lung cancer: a population-based review. *Curr Oncol*. 2014;21(6):e768-e774.
29. Fischer B, Lassen U, Mortensen J, et al. Preoperative staging of lung cancer with combined PET-CT. *N Engl J Med*. 2009;361(1):32-39.
30. van Tinteren H, Hoekstra OS, Smit EF, et al. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial. *Lancet*. 2002;359(9315):1388-1393.
31. Robinson LA, Ruckdeschel JC, Wagner H Jr, Stevens CW. Treatment of non-small cell lung cancer-stage IIIA: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest*. 2007;132(suppl 3):243S-265S.
32. Ramnath N, Dilling TJ, Harris LJ, et al. Treatment of stage III 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(suppl 5): e314S-e340S.
33. 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.

34. 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.e1.

35. Um SW, Kim HK, Jung SH, et al. Endobronchial ultrasound versus mediastinoscopy for mediastinal nodal staging of non-small-cell lung cancer. *J Thorac Oncol*. 2015;10(2):331-337.

36. Ge X, Guan W, Han F, Guo X, Jin Z. Comparison of endobronchial ultrasound-guided fine needle aspiration and video-assisted mediastinoscopy for mediastinal staging of lung cancer. *Lung*. 2015;193(5):757-766.

37. Cerfolio RJ, Bryant AS. Survival of patients with unsuspected N2 (stage IIIA) nonsmall-cell lung cancer. *Ann Thorac Surg*. 2008;86(2):362-366.

38. 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.

39. 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.