ORIGINAL ARTICLE

Assessment of Updated Commission on Cancer Guidelines for Intraoperative Lymph Node Sampling in Early Stage NSCLC

Brendan T. Heiden, MD, MPHS,[a,*] Daniel B. Eaton Jr., MPH,[b] Su-Hsin Chang, PhD, SM,[b,c] Yan Yan, MD, PhD,[b,c] Martin W. Schoen, MD, MPH,[b,d] Mayank R. Patel, MD,[b] Daniel Kreisel, MD, PhD,[a,b] Ruben G. Nava, MD,[a,b] Bryan F. Meyers, MD, MPH,[a] Benjamin D. Kozower, MD, MPH,[a] Varun Puri, MD, MSCI[a,b]

[a]*Division of Cardiothoracic Surgery, Department of Surgery, Washington University School of Medicine, St. Louis, Missouri*
[b]*VA St. Louis Health Care System, St. Louis, Missouri*
[c]*Division of Public Health Sciences, Department of Surgery, Washington University School of Medicine, St. Louis, Missouri*
[d]*Division of Hematology and Medical Oncology, Department of Internal Medicine, Saint Louis University School of Medicine, St. Louis, Missouri*

Received 22 March 2022; revised 18 July 2022; accepted 2 August 2022
Available online - 30 August 2022

ABSTRACT

Introduction: The American College of Surgeons Commission on Cancer recently updated its sampling recommendations for early stage NSCLC from at least 10 lymph nodes to at least one N1 (hilar) and three N2 (mediastinal) lymph node stations. Nevertheless, intraoperative lymph node sampling minimums remain subject to debate. We sought to evaluate these guidelines in patients with early stage NSCLC.

Methods: We performed a cohort study using a uniquely compiled data set from the Veterans Health Administration. We manually abstracted data from operative notes and pathology reports of patients with clinical stage I NSCLC receiving surgery (2006–2016). Adequacy of lymph node sampling was defined using count-based (≥10 lymph nodes) and station-based (≥three N2 and one N1 nodal stations) minimums. Our primary outcome was recurrence-free survival. Secondary outcomes were overall survival and pathologic upstaging.

Results: The study included 9749 patients. Count-based and station-based sampling guidelines were achieved in 3302 (33.9%) and 2559 patients (26.3%), respectively, with adherence to either sampling guideline increasing over time from 35.6% (2006) to 49.1% (2016). Adherence to station-based sampling was associated with improved recurrence-free survival (multivariable-adjusted hazard ratio = 0.815, 95% confidence interval: 0.667–0.994, $p = 0.04$), whereas adherence to count-based sampling was not (adjusted hazard ratio = 0.904, 95% confidence interval: 0.757–1.078, $p = 0.26$). Adherence to either station-based or count-based guidelines was associated with improved overall survival and higher likelihood of pathologic upstaging.

Conclusions: Our study supports station-based sampling minimums (≥three N2 and one N1 nodal stations) for early stage NSCLC; however, the marginal benefit compared with count-based guidelines is minimal. Further efforts to promote widespread adherence to intraoperative lymph node sampling minimums are critical for improving patient outcomes after curative-intent lung cancer resection.

Keywords: Non–small cell lung cancer; Surgery; Lymph node sampling; Quality metrics

***Corresponding author.**

Disclosure: ***The authors declare no conflict of interest.***

Presented at Multidisciplinary Thoracic Cancers Symposium, December 2 to 4, 2021, Scottsdale, Arizona.

Address for correspondence: Brendan T. Heiden, MD, MPHS, Division of Cardiothoracic Surgery, Department of Surgery, Washington University School of Medicine, 660 South Euclid Avenue, Campus Box 8234, St Louis, MO 63110. E-mail: bheiden@wustl.edu

ISSN: 1556-0864

https://doi.org/10.1016/j.jtho.2022.08.009

Introduction

NSCLC is the leading cause of cancer-related mortality in the United States.[1] Surgical treatment remains the standard therapy for early stage disease, especially among functionally fit individuals.[2] A critical principle of lung cancer surgery is to systematically and routinely sample hilar and mediastinal lymph nodes or lymph node stations intraoperatively.[3] This is in contrast to mediastinal lymph node dissection where complete removal of all nodal tissues for a prespecified set of lymph node stations is performed. A pivotal randomized controlled trial (ACOSOG-Z0030) comparing nodal sampling with mediastinal lymph node dissection revealed that sampling is adequate, particularly for clinically node-negative disease.[4]

Nonetheless, what constitutes "adequate" intraoperative lymph node sampling remains unclear given multiple conflicting clinical guidelines. For example, the National Comprehensive Cancer Network (NCCN) recommends a station-based sampling strategy, assessing at least three N2 and one N1 nodal stations intraoperatively.[5] The American College of Surgeons (ACS) Commission on Cancer (CoC), conversely, long recommended a count-based sampling strategy, sampling at least 10 lymph nodes intraoperatively,[6] though these standards were recently changed to a station-based sampling strategy similar to the that in the NCCN (i.e., at least three N2 and one N1 nodal stations). These updated recommendations were based on the theoretical premise that lymph node location should supersede count, especially because several contemporary studies have revealed substantial variability between the number of lymph nodes sampled and long-term survival.[7,8] Despite this, comprehensive evaluations of these count- and station-based guidelines (i.e., $\geq$10 lymph nodes versus $\geq$three N2 and $\geq$one N1 stations) are lacking.

In this study, we assembled a novel data set from the Veterans Health Administration (VHA) of Veterans with clinical stage I NSCLC. Our team reviewed pathology reports and operative notes and abstracted data on operative lymph node sampling for nearly 10,000 Veterans. We evaluated the relationship between count-based (i.e., $\geq$10 lymph nodes) or station-based (i.e., $\geq$three N2 and $\geq$one N1 stations) sampling and several cancer-specific outcomes, including upstaging, recurrence-free survival (RFS), and overall survival (OS).

Materials and Methods

Study Population

We performed this study using a uniquely compiled cohort of Veterans with clinical stage I NSCLC (tumors $\leq$ 5 cm, N0) receiving surgical treatment (2006–2016) through the VHA. Cases of lung cancer were determined using the International Classification of Diseases for Oncology, Third Edition (ICD-O-3), codes. Surgical treatments were identified using ICD-9/ICD-10 procedure or Current Procedural Terminology codes. The VHA Informatics and Computing Infrastructure, a system that houses multiple administrative and clinical data sources within the Corporate Data Warehouse such as the Oncology Raw and the VA Surgical Quality Improvement Program data sets, was queried to generate the cohort of interest.[9,10] The study protocol was reviewed and approved by the St. Louis VHA Research and Development Committee and Institutional Review Board who granted a waiver of informed consent given the deidentified nature of the analysis. Results were reported according to the Strengthening the Reporting of Observational Studies in Epidemiology guidelines.[11]

Lymph Node Sampling

Lymph node sampling information was obtained from operative notes and pathology reports, which were accessed by means of the Compensation and Pension Record Interchange system. Data were abstracted regarding the total number of lymph nodes assessed and the location (station) from which these nodes were procured. Nodes and stations that were sampled during mediastinoscopy were included in the assessment. Two clinical research associates who received specialized training performed the data abstraction. To further ensure accuracy, the first 200 report abstractions were supervised by a board-certified thoracic surgeon. The subsequent 300 reports were independently abstracted by these two investigators, in which less than 3% discordance rate was achieved (which was the prespecified threshold for acceptable concordance, per the study protocol).

Patients were classified on the basis of whether they met count- or station-based sampling guidelines. The count-based guideline was defined as assessing at least 10 total hilar or mediastinal lymph nodes (i.e., old CoC guidelines).[6,12] The station-based guideline was defined as assessing at least one N1 (hilar) station and at least three N2 (mediastinal) stations (i.e., new CoC and current NCCN guidelines).[5,6] All lymph node stations were defined according to the International Association for the Study of Lung Cancer map.[13]

Covariates

Several patient-, treatment-, and tumor-related covariates were abstracted from the Corporate Data Warehouse system. Detailed information on these covariates has been described previously.[14–18] Mediastinal staging procedures were noted including whether patients underwent positron emission tomography or had invasive staging by means of mediastinoscopy or

endobronchial ultrasound (EBUS). All staging data are presented according to the American Joint Commission on Cancer, seventh edition.

Outcomes

Our primary outcome was RFS, defined as time from surgical treatment to disease recurrence. Recurrence was assessed using a combination of clinical documentation and billing codes suggestive of recurrence, as described previously by our group and others within the VHA.[17,19] Our secondary outcomes included OS, pathologic upstaging, and immediate recurrence (defined as recurrence within 6 mo of surgery). Immediate recurrence was included as an outcome because patients who experience such recurrences may have theoretically had nodal disease that was inadequately assessed at the time of surgery (i.e., incorrect or incomplete staging). Patients with pathologic stage IV disease were excluded from the OS and RFS analyses.

Statistical Analysis

Cohort demographics were reported as proportions (%) for categorical variables and means (SD) for continuous variables. Appropriate chi-square and *t* test statistics were reported when comparing groups. Factors associated with meeting either count-based or station-based lymph node sampling guidelines were assessed using a multivariable logistic regression model. In addition, multivariable logistic regression was used to assess the relationship between lymph node sampling adherence and pathologic upstaging and immediate recurrence. Multivariable Cox proportional hazards regression was used to assess the relationahip between lymph node sampling adequacy and OS. Multivariable cause-specific competing risk regression was used to evaluate the association between lymph node sampling adequacy and RFS, with recurrence as the event and death (before recurrence) as the competing event. OS was censored at the end of study follow-up (May 1, 2020) using the VHA Vital Status Files and displayed using Kaplan-Meier curves.[20] RFS was censored at date of last follow-up or death and displayed using cumulative incidence functions. In exploratory analyses, multivariable-restricted cubic spline models were used to assess the relationship between the number of sampled nodes and each outcome. All multivariable models were constructed using hierarchical techniques (clustering at hospital level). Missing values were imputed (n = 20) using the Proc MI MCMC function in SAS (SAS Institute, Cary, NC). All statistical tests were two tailed, and p values less than 0.05 were considered statistically significant. All analyses were performed in SAS version 9.3.

Results

Study Cohort

The study cohort included 9749 Veterans. The mean (SD) age was 67.6 (7.9) years, 9391 (96.3%) were of male sex, and 8067 (82.8%) identified as white. Smoking at the time of surgical treatment was noted in 5701 patients (58.5%). The median (interquartile range) time between cancer diagnosis and surgery was 63 (41–96) days, with 3046 (31.2%) experiencing delayed care (>12 wk). The most common type of resection was lobectomy (n = 6913, 70.9%), and the most common approach was thoracotomy (n = 5690, 58.5%). Most tumors displayed adenocarcinoma histology (n = 5195, 53.3%) with higher-grade features (II–IV, n = 7927, 86.8%). Additional demographic and treatment-related factors are presented in Table 1.

Adherence to Sampling Guidelines

Count-based sampling minimum (i.e., ≥10 lymph nodes) was achieved in 1728 patients (17.7%); station-based sampling minimum (i.e., ≥three N2 and ≥one N1 stations) was achieved in 985 (10.1%); and both sampling minimums (i.e., ≥10 lymph nodes and ≥three N2 and ≥one N1 stations) were achieved in 1574 (16.2%). Among patients who did not meet either guideline, 1033 (10.6%) had no nodal sampling whereas 4429 (45.4%) had some degree of nodal sampling that did not meet guideline criteria (i.e., <10 lymph nodes and <three N2 or <one N1 stations). Adherence to either lymph node sampling guideline increased in the study period from 35.6% in 2006 to 49.1% in 2016 (Fig. 1).

Factors associated with higher likelihood of meeting either guideline in multivariable analysis included more recent surgical year (multivariable-adjusted OR [aOR] = 1.071, 95% confidence interval [CI]: 1.054–1.090, $p < 0.001$), receiving a pneumonectomy (aOR = 1.932, 95% CI: 1.352–2.760, $p < 0.001$), squamous cell carcinoma histology (aOR = 1.165, 95% CI: 1.050–1.292, $p = 0.004$), and larger tumor size (e.g., 31–40 mm versus ≤10 mm, aOR = 1.328, 95% CI: 1.090–1.618, $p = 0.005$). Factors associated with lower likelihood of meeting either guideline included receiving a segmentectomy (aOR = 0.529, 95% CI: 0.431–0.649, $p < 0.001$) and wedge resection (aOR = 0.529, 95% CI: 0.431–0.649, $p < 0.001$).

Primary Outcome

With a median (interquartile range) follow-up of 6.2 (2.5–11.4) years, 2283 patients (23.4%) had recurrent disease. Adherence to station-based sampling was associated with improved RFS (multivariable-adjusted hazard ratio [aHR] = 0.815, 95% CI: 0.667–0.994, $p = 0.04$;

Table 1. Cohort Characteristics

Demographics	Full Population N = 9749	Neither Guideline Met N = 5462	Either Guideline Met N = 4287	*p* Value
Age (SD)	67.61 (7.89)	67.84 (8.14)	67.31 (7.57)	<0.001
Sex (%)				0.95
Male	9391 (96.33)	5262 (96.34)	4129 (96.31)	
Female	358 (3.67)	200 (3.66)	158 (3.69)	
Race (%)				0.04
White	8,067 (82.75)	4,491 (82.22)	3,576 (83.41)	
Black	1,457 (14.95)	858 (15.71)	599 (13.97)	
Other	131 (1.34)	66 (1.21)	65 (1.52)	
Unknown	94 (0.96)	47 (0.86)	47 (1.10)	
Smoking status (at time of surgical treatment, %)				0.40
Current	5701 (58.48)	3223 (59.01)	2478 (57.80)	
Former	3915 (40.16)	2169 (39.71)	1746 (40.73)	
Never	133 (1.36)	70 (1.28)	63 (1.47)	
BMI (SD)	27.16 (5.39)	27.15 (5.51)	27.18 (5.25)	0.74
Area deprivation index (IQR)	61.21 (42.78-75.65)	61.57 (42.15-75.89)	60.64 (43.31-75.40)	0.58
Distance from hospital (IQR)	33.18 (11.79-72.72)	30.50 (10.85-67.23)	36.86 (13.27-81.42)	<0.001
Charlson/Deyo score (IQR)	7 (5-8)	7 (5-8)	7 (5-8)	<0.001
FEV1% (IQR)	78 (66-95)	76 (64-93)	81 (69-97)	<0.001
Frailty score (IQR)	41 (39-43)	41 (40-43)	40 (39-43)	<0.001
Number of prescriptions (IQR)	12 (8-18)	13 (8-19)	12 (8-17)	<0.001
Treatment characteristics				
Wait time to surgery, d				
Median (IQR)	63 (41-96)	63 (41-97)	62 (40-94)	0.02
>12 wk (%)	3046 (31.24)	1730 (31.67)	1316 (30.70)	0.30
Resection (%)				<0.001
Lobectomy	6913 (70.91)	3316 (60.71)	3597 (83.90)	
Wedge	2139 (21.94)	1749 (32.02)	390 (9.10)	
Segment	541 (5.55)	345 (6.32)	196 (4.57)	
Pneumonectomy	152 (1.60)	52 (0.95)	104 (2.43)	
Surgical approach (%)				0.62
Vats	4032 (41.47)	2245 (41.25)	1787 (41.75)	
Thoracotomy	5690 (58.53)	3197 (58.75)	2493 (58.25)	
Tumor size (%)				<0.001
≤10 mm	861 (8.83)	553 (10.12)	308 (7.18)	
11-20 mm	3831 (39.30)	2291 (41.94)	1540 (35.92)	
21-30 mm	2653 (27.21)	1456 (26.66)	1197 (27.92)	
31-40 mm	1476 (15.14)	736 (13.47)	740 (17.26)	
≥40 mm	719 (7.38)	313 (5.73)	406 (9.47)	
Unknown	209 (2.14)	113 (2.07)	96 (2.24)	
Histology (%)				0.002
Adenocarcinoma	5195 (53.29)	2943 (53.88)	2252 (52.53)	
Squamous cell	3294 (33.79)	1774 (32.48)	1520 (35.46)	
Other	1260 (12.92)	745 (13.64)	515 (12.01)	
Grade (%)				0.16
I	1201 (13.16)	688 (13.52)	513 (12.70)	
II	4818 (52.78)	2709 (53.22)	2109 (52.23)	
III	2980 (32.65)	1616 (31.75)	1364 (33.78)	
IV	129 (1.41)	77 (1.51)	52 (1.29)	
Pathologic stage (%)				<0.001
I	8327 (87.69)	4769 (90.19)	3558 (84.55)	
II	754 (7.94)	330 (6.24)	424 (10.08)	
III	372 (3.92)	158 (2.99)	214 (5.09)	
IV	43 (0.45)	31 (0.59)	12 (0.29)	
Adjuvant treatment	1194 (12.25)	625 (11.44)	569 (13.27)	0.006

BMI, body mass index; FEV1, forced expiratory volume in 1 second; IQR, interquartile range.

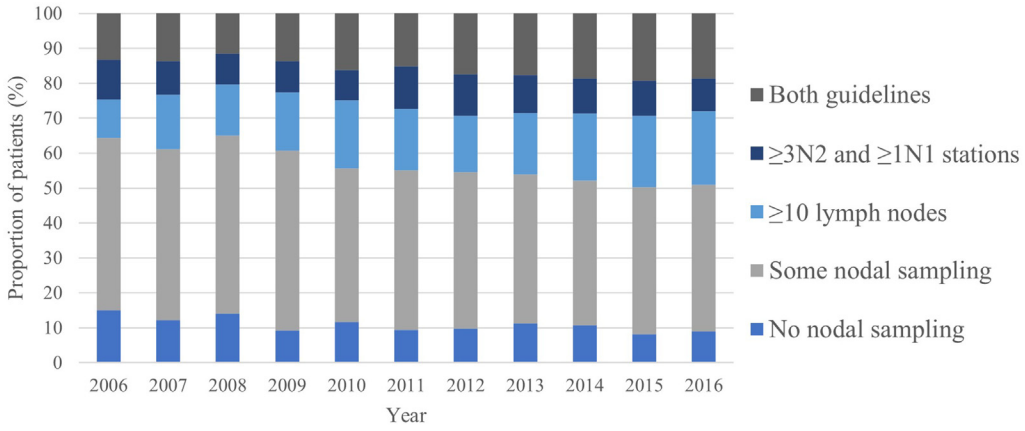

Figure 1. Trend of intraoperative lymph node sampling adherence, 2006 to 2016.

Table 2) whereas adherence to count-based sampling was not associated with improved RFS (aHR = 0.904, 95% CI: 0.757–1.078, $p = 0.26$). Adherence to both count- and station-based sampling did not seem to have an additive benefit in terms of improving RFS (aHR = 0.831, 95% CI: 0.690–1.001, $p = 0.051$; Fig. 2). In stratified analyses by tumor size, station-based sampling remained associated with improved RFS in tumors less than 3 cm (aHR = 0.757, 95% CI: 0.602–0.952, $p = 0.02$). Neither sampling strategy was associated with improved RFS in 3- to 5-cm tumors.

Secondary Outcomes

Adherence to count-based (aHR = 0.822, 95% CI: 0.727–0.930, $p = 0.002$), station-based (aHR = 0.797, 95% CI: 0.712–0.893, $p < 0.001$), or both sampling guidelines (aHR = 0.816, 95% CI: 0.726–0.917, $p < 0.001$) was each associated with improved risk-adjusted OS (Fig. 3). Survival benefit seemed to be most prominent in patients receiving sublobar resections (Supplementary Table 1). Similarly, adherence to count-based (aOR = 2.017, 95% CI: 1.409–2.886, $p < 0.001$), station-based (aOR = 2.596, 95% CI: 1.866–3.611, $p < 0.001$), or both sampling guidelines (aOR = 3.026, 95% CI: 2.169–4.221, $p < 0.001$) was each associated with significantly higher likelihood of pathologic upstaging. Even patients with some nodal sampling (i.e., <10 lymph nodes and <three N2 or <one N1 stations) had improved OS (aHR = 0.848, 95% CI: 0.772–0.932, $p < 0.001$) and higher likelihood of pathologic upstaging (aOR = 1.774, 95% CI: 1.305–2.412, $p < 0.001$) compared with patients with no nodal sampling. Non-adherence to these sampling strategies was unassociated with immediate recurrence on multivariable analysis. Of note, station-based sampling strategies were associated with similar outcome when compared with count-based sampling strategies.

We further evaluated the relationship between count-based sampling minimums and each outcome using multivariable restricted cubic spline functions. Patients seemed to derive maximum RFS and OS benefit after sampling nine and 16 nodes, respectively (Supplementary Fig. 1). No discrete cutoffs were observed for either upstaging or immediate recurrence.

Discussion

In this study, we procured a novel data set of nearly 10,000 Veterans with early stage NSCLC receiving surgical treatment. High-quality operative lymph node sampling data were obtained through intensive review of pathology and operative reports. We found that few patients met either count-based (≥10 lymph nodes) or station-based (≥three N2 and ≥one N1 station) criteria. Adherence to station-based sampling was associated with improved RFS, while count-based sampling was unassociated with RFS. Adherence to either station-based or count-based guidelines was associated with improved OS and higher likelihood of pathologic upstaging. Finally, these findings were most pronounced in tumors less than 3 cm, highlighting the importance of guideline-concordant sampling even in small tumors. Overall, these findings support the updated recommendations from the CoC of station-based sampling minimums for early stage NSCLC.

Intraoperative nodal assessment is a critical component of the surgical treatment of NSCLC. Indeed, nodal assessment is often cited as one of the leading benefits of surgery versus stereotactic body radiotherapy.[21] Despite this, what constitutes adequate intraoperative nodal assessment is subject to much debate. For example, longstanding ACS CoC guidelines advocated for at least 10 lymph nodes to be sampled intraoperatively.[22] On the basis of limited evidence,[7] these guidelines were updated in 2021 to instead recommend sampling three N2

Table 2. Hierarchical Multivariable Models for RFS, OS, Pathologic Upstaging, and Immediate Recurrence

	Entire Cohort		Tumor < 3 cm		Tumor 3-5 cm	
Outcome	aOR or aHR (95% CI)	*p* value	aOR or aHR (95% CI)	*p* value	aOR or aHR (95% CI)	*p* value
RFS						
No nodal sampling	[1 reference]		[1 reference]		[1 reference]	
Some nodal sampling	0.887 (0.764-1.030)	0.11	0.869 (0.740-1.020)	0.09	1.090 (0.700-1.698)	0.70
≥10 LN	0.904 (0.757-1.078)	0.26	0.889 (0.731-1.082)	0.24	1.096 (0.681-1.764)	0.71
3 N2 + 1 N1	0.815 (0.667-0.994)	0.04	0.757 (0.602-0.952)	0.02	1.032 (0.629-1.693)	0.90
Both guidelines	0.831 (0.690-1.001)	0.05	0.821 (0.665-1.013)	0.07	0.971 (0.602-1.565)	0.90
OS						
No nodal sampling	[1 reference]		[1 reference]		[1 reference]	
Some nodal sampling	0.848 (0.772-0.932)	<0.001	0.829 (0.749-0.917)	<0.001	0.973 (0.742-1.276)	0.84
≥10 LN	0.797 (0.712-0.893)	<0.001	0.767 (0.677-0.870)	<0.001	0.952 (0.709-1.278)	0.74
3 N2 + 1 N1	0.822 (0.727-0.930)	0.002	0.789 (0.686-0.908)	<0.001	0.990 (0.731-1.340)	0.95
Both guidelines	0.816 (0.726-0.917)	<0.001	0.807 (0.709-0.919)	0.001	0.934 (0.696-1.254)	0.65
Pathologic upstaging						
No nodal sampling	[1 reference]		[1 reference]		[1 reference]	
Some nodal sampling	1.774 (1.305-2.412)	<0.001	1.825 (1.301-2.516)	<0.001	1.546 (0.662-3.611)	0.31
≥10 LN	2.596 (1.866-3.611)	<0.001	2.983 (2.068-4.302)	<0.001	1.809 (0.788-4.150)	0.16
3 N2 + 1 N1	2.017 (1.409-2.886)	<0.001	2.498 (1.671-3.732)	<0.001	1.161 (0.469-2.871)	0.75
Both guidelines	3.026 (2.169-4.221)	<0.001	3.183 (2.187-4.632)	<0.001	2.558 (1.064-6.148)	0.04
Immediate recurrence						
No nodal sampling	[1 reference]		[1 reference]		[1 reference]	
Some nodal sampling	0.836 (0.565-1.236)	0.37	0.934 (0.584-1.495)	0.78	0.691 (0.267-1.790)	0.45
≥10 LN	0.891 (0.553-1.436)	0.64	0.869 (0.478-1.580)	0.65	0.749 (0.252-2.23)	0.60
3 N2 + 1 N1	0.599 (0.332-1.080)	0.09	0.613 (0.478-1.580)	0.20	0.408 (0.120-1.390)	0.15
Both guidelines	0.930 (0.571-1.517)	0.77	0.919 (0.488-1.728)	0.79	0.794 (0.275-2.300)	0.67

Note: All models controlling for age, sex, race, BMI, smoking status, Charlson comorbidity score, frailty score, number of unique prescriptions, ADI score, distance from treating facility, FEV1, surgical year, delayed treatment, surgical approach, extent of resection, histology, grade, and positive margin. RFS, OS, and immediate recurrence models also control for postoperative readmission, postoperative major complications, prolonged stay (>14 d), pathologic stage, and adjuvant therapy.

ADI, area deprivation index; aHR, adjusted hazard ratio; aOR, adjusted OR; BMI, body mass index; CI, confidence interval; FEV1, forced expiratory volume in 1 second; LN, lymph node; OS, overall survival; RFS, recurrence-free survival.

and one N1 nodal stations, similar to recommendations from the NCCN.[5] Nevertheless, a number of other organizations have also proposed guidelines. For example, the International Association for the Study of Lung Cancer advocates for side-specific sampling (stations 2R, 4R, 7, 10R, and 11R for right-sided tumors; stations 5, 6, 7, 10L, and 11L for left-sided tumors; ±station 9 for lower lobe tumors). The Union for International Cancer Control recommends sampling three N1 and three N2 "lymph nodes/stations."[23] Other groups have advocated

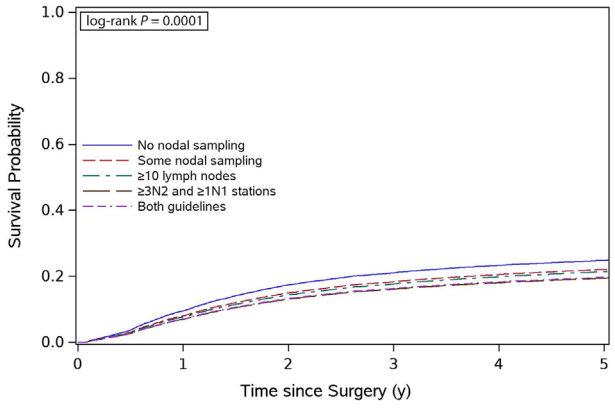

Figure 2. Cumulative incidence function for recurrence stratified by adherence to intraoperative lymph node sampling.

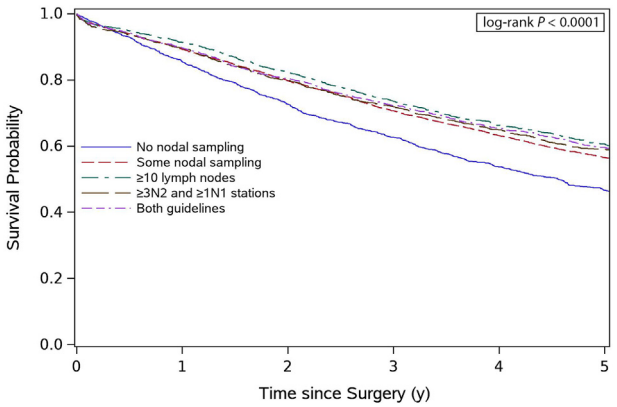

Figure 3. Overall survival curve stratified by adherence to intraoperative lymph node sampling.

for even more complex sampling strategies, including lobe-specific sampling and differential sampling on the basis of tumor histology.[3] Although these standards are each distinct, an emerging theme is the centrality of station-based sampling as opposed to count-based sampling. Indeed, studies question the interpretability of total lymph node counts, especially because fragmented nodes are often erroneously counted as unique lymph nodes. In addition, recent work has suggested that nodal count is a proxy measure of a patient's immune reaction in the tumor microenvironment; higher lymph node yield may reflect a beneficial antitumor immunosurveillance as opposed to a modifiable surgeon-level skill.[24] Therefore, the improved prognosis observed with highly immunogenic tumors may explain the survival benefit associated with higher lymph node yields.

Furthermore, our data reveal that the added benefit of achieving a station-based sampling minimum over a count-based strategy (such as $\geq$10 lymph nodes) is likely small. Indeed, when the two guidelines were compared, outcomes were similar. Nonetheless, it is important to be parsimonious when considering surgical principles. Indeed, the risk of additional mediastinal dissection, although small, can be significant (vascular injury, etc.).[25] Moreover, station-based guidelines, with a greater degree of objectivity (compared with counts, where fragments of a nodes can each be counted as individual nodes), promote more consistent practices and allow comparisons across institutions.

Given the higher likelihood of nodal involvement in larger tumors, there is a natural tendency to question whether extensive mediastinal nodal sampling (i.e., three N2 stations) is beneficial for small, clinically node-negative tumors. This logic could potentially be extended to other "low-risk" categories as well (i.e., well-differentiated tumors and peripheral tumors).[26] Nevertheless, previous studies have revealed that even NSCLC with low-risk features benefit from systematic nodal sampling. For example, a study by Dai et al.[27] revealed a strong association between inadequate nodal assessment and diminished survival, even in T1a ($<$1 cm) tumors. Although other fields have benefited from selective nodal assessments (such as sentinel lymph node biopsy), selective sampling strategies have not been adopted for lung cancer.[28] One reason for this is that preoperative assessments remain imperfect[29,30] and, therefore, occult mediastinal nodal involvement is not infrequently encountered at the time of surgical resection.[31] Indeed, Bille et al.[32] reported in a cohort of 1667 patients with T1 to T2 ($\leq$5 cm), N0 NSCLC that the rate of occult pathologic N2 disease was 9% (6.7% in tumors $\leq$ 2 cm). Furthermore, 34% of these patients with pathologic N2 disease had no associated pN1 disease—so-called skip lesions. The considerable incidence of skip lesions abrogates the entire notion of sequential nodal sampling (i.e., sampling N2 nodes only after N1 disease is identified) in early-stage NSCLC. In conjunction with these previous data, our study further highlights the critical importance of guideline-concordant, systematic, station-based, intraoperative nodal sampling strategies to adequately stage and treat patients with NSCLC.

A critical finding of our study is the benefit of systematic nodal sampling even in small tumors amenable to sublobar resection. A recent, highly anticipated, pivotal trial (JCOG0802/WJOG4607L) found that segmentectomy was associated with improved survival compared with lobectomy for patients with early stage ($\leq$2 cm, peripheral) NSCLC.[33] Importantly, nodal sampling was robust in this study because "systematic or selective lymph node dissection was mandatory" in both groups. Indeed, more than 95% of patients underwent mediastinal nodal assessment in each arm which may partially explain the excellent 5-year relapse-free survival in this study (87.9% [95% CI: 84.8–90.3] for lobectomy versus 88.0% [95% CI: 85.0–90.4] for segmentectomy). We found that guideline-concordant nodal sampling in small tumors amenable to sublobar resection was associated with improved outcomes. Nevertheless, we also found that sublobar resections were associated with higher rates of inadequate nodal sampling. Together, these findings highlight the importance of adequate nodal sampling which will continue to be critical as segmentectomies may become "the standard surgical procedure" for select patients with early-stage disease.[33]

An advantage of the VHA data is the granularity of information available from the electronic medical records, such as comprehensive lymph node data. Conversely, there is a paucity of comprehensive, station-based lymph node assessment and involvement data in other multi-institutional, nationally inclusive data sets. Indeed, the National Cancer Database; Surveillance, Epidemiology, and End Results; and Society of Thoracic Surgery platforms only collect data on the total number of sampled lymph nodes.[34,35] With the dissemination of the newer station-based guidelines, it will be important to collect these more complex data elements in various tumor registries. Future research will need to determine whether different station-based sampling strategies (i.e., lobe-specific sampling) are superior to the current (three N2 and one N1) guidelines.

An area that merits further investigation is the complex interplay between preoperative staging (i.e., high-resolution computed tomography and positron emission tomography imaging), invasive mediastinal assessment (i.e., EBUS or mediastinoscopy, which can occur either preoperatively or at the time of surgery), and intraoperative assessment. In this study, we included nodes

that were sampled by means of mediastinoscopy in our final assessment (consistent with current CoC standards). One could envision, however, that with higher resolution or functionally based imaging modalities in conjunction with EBUS, the rate of occult mediastinal disease will continue to diminish.[36,37] Therefore, lymph node sampling strategies will need to evolve to best reflect the diagnostic accuracy of modern-day imaging and staging modalities.

This study has several strengths. We queried operative and pathology reports from nearly 10,000 patients to assemble a large, homogeneous, multi-institutional, national cohort of patients with clinically node-negative NSCLC undergoing definitive surgical treatment. The study also has some limitations. First, the study group consists of Veterans who are predominantly of male sex with heavy comorbidity burden and smoking histories. Despite this, lung cancer treatment patterns and outcomes are similar between the Veterans and the general U.S. population, suggesting our study is highly relevant to the typical patient with early-stage lung cancer.[16] Nevertheless, further validation of station-based sampling guidelines in non-Veteran populations is encouraged. Second, we were unable to determine whether lymph nodes were assessed in multiple fragments or as whole nodes. This is a common challenge of count-based sampling minimums, but it should not bias the station-based sampling minimums. Third, we assessed recurrence using a combination of clinical and administrative data sources.[38] Although robust, it is possible that untreated episodes of recurrence were undercaptured. Finally, we did not evaluate other sampling strategies in this study (i.e., lobe-specific sampling). Future research will need to focus on these other station-based approaches.

In conclusion, in a cohort study of nearly 10,000 Veterans with clinically node-negative NSCLC, we sought to evaluate count-based (≥10 lymph nodes) and station-based (≥three N2 and one N1 nodal stations) nodal sampling minimums on the basis of updated guidelines from the ACS CoC. Our findings support station-based sampling minimums (≥three N2 and one N1 nodal stations) for early-stage NSCLC although the relative benefit of meeting one guideline over the other is modest. Efforts to enhance guideline-concordant intraoperative lymph node sampling could have a disproportionate impact on patient outcomes after curative-intent lung cancer resection.

CRediT Authorship Contribution Statement

Brendan T. Heiden: Conceptualization, Methodology, Formal analysis, Investigation, Writing—original draft.

Daniel B. Eaton: Data Curation, Formal analysis, Visualization.

Su-Hsin Chang: Conceptualization, Writing—review and editing.

Yan Yan: Conceptualization, Writing—review and editing.

Martin W. Schoen: Conceptualization, Writing—review and editing.

Mayank R. Patel: Conceptualization, Writing—review and editing.

Daniel Kreisel: Conceptualization, Writing—review and editing.

Ruben G. Nava: Conceptualization, Writing—review and editing.

Bryan F. Meyers: Conceptualization, Writing—review and editing.

Benjamin D. Kozower: Conceptualization, Writing—review and editing.

Varun Puri: Conceptualization, Methodology, Formal analysis, Investigation, Writing—original draft, Funding acquisition.

Acknowledgments

This work was supported by Merit Award # 1I01HX002475-01A2 from the U.S. Department of Veterans Affairs (VP, SHC, YY, DBE) and 5T32HL007776-25 from the National Institutes of Health (BTH). Funded in part by NIH 5T32HL007776-25 (BTH), 1 I01 HX002475-01A2 (VP, S-HC, YY). The contents do not represent the views of the U.S. Department of Veterans Affairs or the U.S. Government. This cohort study required extensive data acquisition, biostatistical, and technical review. Our team consists of five attending surgeons from Washington University School of Medicine, three attending surgeons from the St. Louis VA Hospital (including two with dual appointments at Washington University), a medical oncologist at the St. Louis VA Hospital (dual appointment at St. Louis University School of Medicine), three biostatisticians (at Washington University and St. Louis VA Hospital), and a general surgery resident.

Supplementary Data

Note: To access the supplementary material accompanying this article, visit the online version of the *Journal of Thoracic Oncology* at www.jto.org and at https://doi.org/10.1016/j.jtho.2022.08.009.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70:7-30.
2. Puri V, Crabtree TD, Bell JM, et al. Treatment outcomes in stage I lung cancer: a comparison of surgery and

stereotactic body radiation therapy. *J Thorac Oncol*. 2015;10:1776–1784.
3. Lardinois D, De Leyn P, Van Schil P, et al. ESTS guidelines for intraoperative lymph node staging in non-small cell lung cancer. *Eur J Cardio Thorac Surg*. 2006;30:787–792.
4. Darling GE, Allen MS, Decker PA, et al. Randomized trial of mediastinal lymph node sampling versus complete lymphadenectomy during pulmonary resection in the patient with N0 or N1 (less than hilar) non-small cell carcinoma: results of the American College of Surgery Oncology Group Z0030 trial. *J Thorac Cardiovasc Surg*. 2011;141:662–670.
5. National Comprehensive Cancer Network. NCCN Guidelines in Oncology: non-small cell lung cancer, Version 6. 2020. https://www.nccn.org/professionals/physician_gls/pdf/nscl_blocks.pdf. Accessed January 9, 2020.
6. Nissen AP, Vreeland TJ, Teshome M, et al. American College of Surgeons Commission on Cancer standard for curative-intent pulmonary resection. *Ann Thorac Surg*. 2022;113:5–8.
7. Smeltzer MP, Faris NR, Ray MA, Osarogiagbon RU. Association of pathologic nodal staging quality with survival among patients with non-small cell lung cancer after resection with curative intent. *JAMA Oncol*. 2018;4:80–87.
8. Dezube AR, Mazzola E, Bravo-Iñiguez CE, et al. Analysis of lymph node sampling minimums in early stage non-small-cell lung cancer. *Semin Thorac Cardiovasc Surg*. 2021;33:834–845.
9. Zullig LL, Smith VA, Jackson GL, et al. Colorectal cancer statistics from the Veterans Affairs central cancer registry. *Clin Colorectal Cancer*. 2016;15:e199–e204.
10. Massarweh NN, Kaji AH, Itani KMF. Practical guide to surgical data sets: Veterans Affairs Surgical Quality Improvement Program (VASQIP). *JAMA Surg*. 2018;153:768–769.
11. Ghaferi AA, Schwartz TA, Pawlik TM. STROBE reporting guidelines for observational studies. *JAMA Surg*. 2021;156:577–578.
12. Samson P, Crabtree T, Broderick S, et al. Quality measures in clinical stage I non-small cell lung cancer: improved performance is associated with improved survival. In: *Ann Thorac Surg*. 2017;103:303–311.
13. Goldstraw P, Chansky K, Crowley J, et al. The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM Classification for lung cancer. *J Thorac Oncol*. 2016;11:39–51.
14. Heiden BT, Eaton DB Jr, Engelhardt KE, et al. Analysis of delayed surgical treatment and oncologic outcomes in clinical stage I non-small cell lung cancer. *JAMA Netw Open*. 2021;4:e2111613.
15. Heiden BT, Eaton Jr DB, Chang SH, et al. Assessment of duration of smoking cessation prior to surgical treatment of non-small cell lung cancer [e-pub ahead of print]. *Ann Surg*. https://doi.org/10.1097/SLA.0000000000005312, accessed March 15, 2022.
16. Heiden BT, Eaton DBJ, Chang SH, et al. Comparison between veteran and non-veteran populations with clinical stage I non-small cell lung cancer undergoing surgery cancer [e-pub ahead of print]. *Ann Surg*. https://doi.org/10.1097/SLA.0000000000004928, accessed March 15, 2022.
17. Heiden BT, Eaton Jr DB, Chang SH, et al. The impact of persistent smoking after surgery on long-term outcomes following stage I non-small cell lung cancer resection. *Chest*. 2022;161:1687–1696.
18. Shinall Jr MC, Arya S, Youk A, et al. Association of preoperative patient frailty and operative stress with postoperative mortality. *JAMA Surg*. 2020;155:e194620.
19. Tarlov E, Lee TA, Weichle TW, et al. Reduced overall and event-free survival among colon cancer patients using dual system care. *Cancer Epidemiol Biomarkers Prev*. 2012;21:2231–2241.
20. Sohn M-W, Arnold N, Maynard C, Hynes DM. Accuracy and completeness of mortality data in the Department of Veterans Affairs. *Popul Health Metr*. 2006;4:2.
21. Ijsseldijk MA, Shoni M, Siegert C, et al. Oncologic outcomes of surgery versus SBRT for non-small-cell lung carcinoma: a systematic review and meta-analysis. *Clin Lung Cancer*. 2021;22:e235–e292.
22. American College of Surgeons. CoC quality of care measures, surveys. https://www.facs.org/Quality-Programs/Cancer/NCDB/qualitymeasurescocweb. Accessed December 21, 2020.
23. Brierley JD, Gospodarowicz MK, Wittekind C, et al. *TNM Classification of Malignant Tumours*. Chichester, United Kingdom: John Wiley & Sons, Inc.; 2017.
24. Lal N, Chan DKH, Ng ME, Vermeulen L, Buczacki SJA. Primary tumour immune response and lymph node yields in colon cancer. *Br J Cancer*. 2022;126:1178–1185.
25. Allen MS, Darling GE, Pechet TTV, et al. Morbidity and mortality of major pulmonary resections in patients with early-stage lung cancer: initial results of the randomized, prospective ACOSOG Z0030 trial. *Ann Thorac Surg*. 2006;81:1013–1020.
26. Pathak R, Goldberg SB, Canavan M, et al. Association of survival with adjuvant chemotherapy among patients with early-stage non-small cell lung cancer with vs without high-risk clinicopathologic features. *JAMA Oncol*. 2020;6:1741–1750.
27. Dai J, Liu M, Yang Y, et al. Optimal lymph node examination and adjuvant chemotherapy for stage I lung cancer. *J Thorac Oncol*. 2019;14:1277–1285.
28. Gregor A, Ujiie H, Yasufuku K. Sentinel lymph node biopsy for lung cancer. *Gen Thorac Cardiovasc Surg*. 2020;68:1061–1078.
29. Navani N, Fisher DJ, Tierney JF, Stephens RJ, Burdett S, NSCLC Meta-analysis Collaborative Group. The accuracy of clinical staging of stage I-IIIa non-small cell lung cancer: an analysis based on individual participant data. *Chest*. 2019;155:502–509.
30. Heineman DJ, ten Berge MG, Daniels JM, et al. The quality of staging non-small cell lung cancer in the Netherlands: data from the Dutch Lung surgery audit. *Ann Thorac Surg*. 2016;102:1622–1629.
31. Molena D. Shining light on sentinel node assessment for lung cancer: quality is superior to quantity. *J Thorac Cardiovasc Surg*. 2018;155:1292–1293.
32. Bille A, Woo KM, Ahmad U, Rizk NP, Jones DR. Incidence of occult pN2 disease following resection and

mediastinal lymph node dissection in clinical stage I lung cancer patients. *Eur J Cardiothorac Surg*. 2017;51:674–679.
33. Saji H, Okada M, Tsuboi M, et al. Segmentectomy versus lobectomy in small-sized peripheral non-small-cell lung cancer (JCOG0802/WJOG4607L): a multicentre, open-label, phase 3, randomised, controlled, non-inferiority trial. *Lancet*. 2022;399:1607–1617.
34. Merkow RP, Rademaker AW, Bilimoria KY. Practical guide to surgical data sets: National Cancer Database (NCDB). *JAMA Surg*. 2018;153:850–851.
35. Farjah F, Kaji AH, Chu D. Practical guide to surgical data sets: Society of Thoracic Surgeons (STS) national database. *JAMA Surg*. 2018;153:955–956.
36. Korevaar DA, Crombag LM, Cohen JF, Spijker R, Bossuyt PM, Annema JT. Added value of combined endobronchial and oesophageal endosonography for mediastinal nodal staging in lung cancer: a systematic review and meta-analysis. *Lancet Respir Med*. 2016;4:960–968.
37. Navani N, Nankivell M, Lawrence DR, et al. Lung cancer diagnosis and staging with endobronchial ultrasound-guided transbronchial needle aspiration compared with conventional approaches: an open-label, pragmatic, randomised controlled trial. *Lancet Respir Med*. 2015;3:282–289.
38. Subramanian MP, Hu Y, Puri V, Kozower BD. Administrative versus clinical databases. *J Thorac Cardiovasc Surg*. 2021;162:1173–1176.