Variability in invasive mediastinal staging for lung cancer: A multicenter regional study

Lucas W. Thornblade, MD, MPH,[a] Douglas E. Wood, MD,[a] Michael S. Mulligan, MD,[a] Alexander S. Farivar, MD,[b] Michal Hubka, MD,[c] Kimberly E. Costas, MD,[d] Bahirathan Krishnadasan, MD,[e] and Farhood Farjah, MD, MPH[a] for The SCOAP-CERTAIN Collaborative

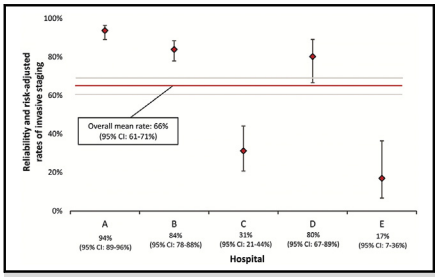
ABSTRACT

Objective: Prior studies have reported underuse of—but not variability in—invasive mediastinal staging in the pretreatment evaluation of patients with lung cancer. We sought to compare rates of invasive mediastinal staging for lung cancer across hospitals participating in a regional quality improvement and research collaborative.

Methods: We conducted a retrospective study (2011-2013) of patients undergoing resected lung cancer from the Surgical Clinical Outcomes and Assessment Program in Washington State. Invasive mediastinal staging included mediastinoscopy and/or endobronchial/esophageal ultrasound-guided nodal aspiration. We used a mixed-effects model to mitigate the influence of small sample sizes at any 1 hospital on rates of invasive staging and to adjust for hospital-level differences in the frequency of clinical stage IA disease.

Results: A total of 406 patients (mean age, 68 years; 69% clinical stage IA; and 67% lobectomy) underwent resection at 5 hospitals (4 community and 1 academic). Invasive staging occurred in 66% of patients (95% confidence interval [CI], 61%-71%). CI inspection revealed that 2 hospitals performed invasive staging significantly more often than the overall average (94%, [95% CI, 89%-96%] and 84% [95% CI, 78%-88%]), whereas 2 hospitals performed invasive staging significantly less often than overall average (31% [95% CI, 21%-44%] and 17% [95% CI, 7%-36%]).

Conclusions: Rates of invasive mediastinal staging varied significantly across hospitals providing surgical care for patients with lung cancer. Future studies that aim to understand the reasons underlying variability in care may inform quality improvement initiatives or lead to the development of novel staging algorithms. (J Thorac Cardiovasc Surg 2018;155:2658-71)



Variation in hospital-level rates of invasive mediastinal staging.

Central Message

Rates of invasive mediastinal staging vary significantly across hospitals that provide surgical care to patients with lung cancer.

Perspective

Our study demonstrates what many surgeons suspect—lung cancer staging varies significantly across hospitals. This finding provides the critical first step in a line of investigation that seeks to better understand the reasons underlying this highly variable care. Better understanding variability in lung cancer staging may inform quality improvement initiatives or lead to novel staging algorithms.

See Editorial Commentary page 2672.

An American College of Surgeons Commission on Cancer study described substantially lower than expected rates of invasive mediastinal staging (IMS) for resected lung cancer (27%) in 2001.[1] In addition, fewer than half of all patients who underwent mediastinoscopy had lymph node tissue submitted for pathologic evaluation. Reed[1] said in her discussion at an annual meeting of the Society of Thoracic Surgeons (STS) that, ''The results are an indictment of the present care of patients with non–small cell lung cancer.''

From the [a]Department of Surgery, University of Washington, Seattle, Wash; [b]Division of Thoracic Surgery, Swedish Cancer Institute, Seattle, Wash; [c]Department of Thoracic Surgery, Virginia Mason Medical Center, Seattle, Wash; [d]Division of Thoracic Surgery, Providence Regional Medical Center, Everett, Wash; and [e]General Thoracic Surgery, CHI Franciscan Health System, Tacoma, Wash.

Supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under award No. T32DK070555. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

University of Washington's Surgical Outcomes Research Center and Comparative Effectiveness Research Translational Network has received funding from the Agency for Healthcare Research and Quality, Life Sciences Discovery Fund, and National Institutes of Health.

Read at the 43rd Annual Meeting of The Western Thoracic Surgical Association, Colorado Springs, Colorado, June 21-24, 2017.

Received for publication May 17, 2017; revisions received Nov 28, 2017; accepted for publication Dec 14, 2017; available ahead of print March 11, 2018.

Address for reprints: Farhood Farjah, MD, MPH, Surgical Outcomes Research Center, 1107 NE 45th St, Ste 502, Box 354808, Seattle, WA 98105 (E-mail: ffarjah@uw.edu).

0022-5223/$36.00

https://doi.org/10.1016/j.jtcvs.2017.12.138

Abbreviations and Acronyms

EBUS = endobronchial ultrasound-guided nodal aspirate
IMS = invasive mediastinal staging
NSCLC = non–small cell lung cancer
PET = positron emission tomography
SCOAP = Surgical Care Outcomes Assessment Program
STS = Society of Thoracic Surgeons

Scanning this QR code will take you to a supplemental video and appendix for the article.

Since then, numerous studies have reported apparent underuse of IMS in other lung cancer populations using other data sources[2-8]; however, there are several challenges in drawing firm conclusions about the quality of care based on these studies. First, contemporary guidelines have shifted away from recommending routine IMS.[9-11] Second, there are no known datasets with sufficient granularity to directly assess adherence with guideline-recommended IMS. Variation in care delivery is a commonly used surrogate measure for suboptimal care. Nonetheless, there may be other reasons for variability in care, such as concerns over the level of evidence supporting guideline recommendations or providers who perceive IMS to be time-consuming, inconvenient, or even dangerous. These speculative reasons form the basis of a long-standing suspicion of variability in rates of IMS. Variability in hospital-level rates of IMS has never been documented.

Leveraging a unique data source of academic and community-based thoracic surgical practices in the Puget Sound region of Washington State, we sought to compare hospital-level rates of IMS among patients with resected non–small cell lung cancer (NSCLC). We hypothesized that there would be significant variation in rates of IMS across hospitals.

MATERIALS AND METHODS

Study Design and Population

We performed a retrospective cohort study of adult NSCLC patients who underwent resection between July 2011 and December 2013 (n = 514). Patients were excluded if they had a prior history of lung cancer (n = 43) or were treated with induction therapy (n = 59). Data were obtained from the medical record of hospitals/health systems in the Puget Sound region of Washington State participating in the Collaborative to Improve Native Cancer Outcomes study.[12] Of 13 Collaborative to Improve Native Cancer Outcomes study hospitals (Appendix E1), 7 contributed lung cancer cases. Two hospitals contributing fewer than 10 cases over the duration of the study were excluded from the analysis resulting in the exclusion of 6 additional patients. This study was not considered human subjects research by the University of Washington Institutional Review Board because both patient and hospital information were de-identified.

Data Source

Data were obtained from the Surgical Care and Outcomes Assessment Program (SCOAP). SCOAP is a physician-led quality improvement system in Washington State. The SCOAP database is a prospective, multicenter clinical registry that contains data on patients' characteristics, perioperative and surgical details, as well as clinical outcomes. SCOAP is administered by the Foundation for Health Care Quality. Research and development in SCOAP is funded in part by grants to the University of Washington Surgical Outcomes Research Center and the Comparative Effectiveness Research Translation Network.[13,14] In 2013, the Comparative Effectiveness Research Translation Network developed a thoracic surgery quality improvement module for use by hospitals that participate in SCOAP. Stakeholders include thoracic surgeons, pulmonologists, advanced practice providers, and hospital administrators.[12] This group determined the need to collect information about lung cancer staging and added disease-specific variables (eg, provider documented clinical stage, details of IMS [listed in the next paragraph], and use of noninvasive staging modalities) to the lung cancer module within SCOAP.

IMS

Abstractors recorded information from patients' medical records on the type (eg, mediastinoscopy or mediastinotomy, endobronchial ultrasound-guided nodal aspirate [EBUS], or esophageal ultrasound-guided nodal aspirate) and timing (before resection, at the time of resection, or both before and at the time of resection) of IMS if it was performed. In addition, abstractors collected data on the number of unique mediastinal nodal stations sampled during IMS. Abstractors did not collect information on the indications for IMS, and therefore it is not possible to determine guideline adherence at the patient level.

Analysis

Patient characteristics were summarized by frequencies (categorical variables) and means ± standard deviation or medians (interquartile range) for normally and nonnormally distributed continuous variables, respectively. Univariate differences across hospitals are summarized by χ^2 tests (categorical variable) and analysis of variance tests or Kruskal-Wallis tests for normally and nonnormally distributed continuous variables, respectively. All analyses were performed using Stata version 14.2 (StataCorp LLC, College Station, Tex).

Confidence interval (CI) inspection was used to compare hospital-level rates of IMS to each other and the overall average rate for the study population. Because small sample sizes at any single hospital can influence the observed rate of IMS and erroneously inflate the magnitude of variation, we used an empirical Bayes shrinkage estimator.[15-17] This approach, using a mixed effects logistic regression model, also allowed for adjustment for clinical stage IA tumors. Guidelines allow for omission of IMS in a subset of these patients,[11,18] and the frequency of clinical stage IA disease may vary across hospitals; therefore, adjustment for clinical stage IA is imperative to avoid confounding bias in comparisons of hospital-level rates of IMS. Finally, this approach allowed for estimation of the proportion of variability explained by differences in the rates of clinical stage IA disease across hospitals (by calculating the ratio of random effects between models). CIs for point estimates following reliability adjustment were based on empirical Bayes variance estimates and on approximate normality of the posterior facility-level

TABLE 1. Characteristics of patients undergoing resection for lung cancer across hospitals

	All (n = 406)	Hospital A (n = 130)	Hospital B (n = 134)	Hospital C (n = 72)	Hospital D (n = 40)	Hospital E (n = 30)	*P* value
Mean age (y)	68.2 (10.0)	66.7 (10.6)	68.8 (9.9)	68.2 (10.0)	69.7 (8.8)	71.5 (8.5)	.12
Women	228 (57)	70 (54)	84 (64)	38 (53)	16 (40)	20 (67)	.05
Race							.79
White	354 (88)	109 (85)	116 (87)	64 (89)	37 (95)	28 (93)	
Black	7 (2)	4 (3)	2 (2)	1 (1)	0 (0)	0 (0)	
Asian	26 (6)	9 (7)	10 (8)	5 (7)	0 (0)	2 (7)	
Native American	4 (1)	3 (2)	0 (0)	0 (0)	1 (3)	0 (0)	
Pacific Islander	3 (1)	1 (1)	2 (2)	0 (0)	0 (0)	0 (0)	
Other	8 (2)	2 (2)	3 (2)	2 (3)	1 (3)	0 (0)	
Insurance							<.001
Commercial	145 (36)	39 (30)	49 (37)	44 (61)	9 (23)	4 (13)	
Medicare	77 (19)	9 (7)	16 (12)	27 (38)	15 (38)	10 (33)	
Medicaid	13 (3)	5 (4)	5 (4)	0 (0)	2 (5)	1 (3)	
Tricare (military)	2 (<1)	1 (1)	0 (0)	0 (0)	1 (3)	0 (0)	
Other government	3 (1)	2 (2)	0 (0)	0 (0)	0 (0)	1 (3)	
Uninsured	2 (<1)	1 (1)	0 (0)	1 (1)	0 (0)	0 (0)	
Government and commercial	126 (31)	63 (48)	44 (33)	0 (0)	7 (18)	12 (40)	
Mixed government	38 (9)	10 (8)	20 (15)	0 (0)	6 (15)	2 (7)	
Charlson comorbidity index							.34
0	227 (56)	73 (56)	69 (51)	47 (65)	21 (53)	17 (57)	
1	141 (35)	42 (32)	50 (37)	21 (29)	15 (38)	13 (43)	
2	28 (7)	12 (9)	11 (8)	1 (1)	4 (10)	0 (0)	
3+	10 (2)	3 (2)	4 (3)	3 (4)	0 (0)	0 (0)	
ASA class							.01
II	71 (17)	12 (9)	25 (19)	20 (28)	10 (25)	4 (13)	
III	312 (77)	107 (82)	99 (74)	51 (71)	29 (73)	26 (87)	
IV	23 (6)	11 (8)	10 (7)	1 (1)	1 (3)	0 (0)	
Smoking status							.13
Former	235 (60)	78 (63)	79 (60)	42 (63)	20 (53)	16 (53)	
Current	96 (25)	31 (25)	39 (30)	11 (16)	8 (21)	7 (23)	
Never	60 (16)	15 (12)	14 (11)	14 (21)	10 (26)	7 (23)	
% Predicted FEV_1	83 (69-96)	81 (66-95)	80 (66-94)	84 (75-100)	87 (73-100)	84 (68-103)	.16
% Predicted DLCO	64 (52-77)	64 (55-75)	63 (49-80)	66 (51-78)	66 (62-83)	60 (50-70)	.31
Clinical stage							<.001
IA	281 (69)	74 (57)	110 (82)	54 (75)	22 (55)	21 (70)	
IB	54 (13)	15 (12)	11 (8)	15 (21)	8 (20)	5 (17)	
IIA	33 (8)	17 (13)	6 (5)	2 (3)	4 (10)	4 (13)	
IIB	20 (5)	10 (8)	6 (5)	0 (0)	4 (10)	0 (0)	
IIIA	15 (4)	11 (8)	1 (1)	1 (1)	2 (5)	0 (0)	
IIIB	1 (<1)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	
IV	2 (<1)	2 (2)	0 (0)	0 (0)	0 (0)	0 (0)	
Preoperative PET scan	371 (92)	125 (96)	120 (90)	64 (89)	38 (97)	24 (83)	.049
Extent of resection							.001
Wedge	35 (9)	19 (15)	9 (7)	5 (7)	1 (3)	1 (3)	
Segmentectomy	7 (2)	3 (2)	2 (1)	0 (0)	2 (5)	0 (0)	
Lobectomy	272 (67)	71 (55)	102 (76)	46 (64)	28 (70)	25 (83)	
Sleeve lobectomy	2 (<1)	1 (1)	0 (0)	0 (0)	1 (3)	0 (0)	
Lobectomy and wedge	66 (16)	19 (15)	18 (13)	17 (24)	8 (20)	4 (13)	
Bilobectomy	14 (3)	9 (7)	1 (1)	4 (6)	0 (0)	0 (0)	
Pneumonectomy	10 (2)	8 (6)	2 (1)	0 (0)	0 (0)	0 (0)	

(Continued)

TABLE 1. Continued

	All (n = 406)	Hospital A (n = 130)	Hospital B (n = 134)	Hospital C (n = 72)	Hospital D (n = 40)	Hospital E (n = 30)	*P* value
MIS approach	292 (72)	84 (65)	129 (96)	41 (57)	33 (83)	5 (17)	<.001
Right sided resection	247 (61)	81 (62)	81 (60)	39 (54)	28 (70)	18 (60)	.75
Pathologic stage							.008
0	5 (1)	2 (2)	1 (1)	1 (1)	0 (0)	1 (3)	
IA	201 (50)	51 (39)	79 (59)	40 (56)	20 (50)	11 (37)	
IB	84 (21)	24 (18)	28 (21)	17 (24)	6 (15)	9 (30)	
IIA	38 (9)	15 (12)	11 (8)	5 (7)	1 (3)	6 (20)	
IIB	32 (8)	15 (12)	8 (6)	2 (3)	5 (13)	2 (7)	
IIIA	41 (10)	22 (17)	5 (4)	7 (10)	7 (18)	0 (0)	
IV	5 (1)	1 (1)	2 (1)	0 (0)	1 (3)	1 (3)	
Pathologic nodal stage							.053
pN0	317 (79)	93 (73)	110 (83)	55 (79)	32 (80)	27 (90)	
pN1	47 (12)	18 (14)	17 (13)	7 (10)	2 (5)	3 (10)	
pN2	36 (9)	17 (13)	5 (4)	8 (11)	6 (15)	0 (0)	
Histology							.95
Adenocarcinoma	270 (67)	82 (63)	87 (65)	49 (68)	30 (75)	22 (73)	
AIS	10 (2)	5 (4)	3 (2)	1 (1)	0 (0)	1 (3)	
Squamous cell carcinoma	86 (21)	30 (23)	30 (22)	12 (17)	9 (23)	5 (17)	
Large cell carcinoma	7 (2)	3 (2)	2 (1)	2 (3)	0 (0)	0 (0)	
NSCLC NOS	22 (5)	6 (5)	9 (7)	6 (8)	0 (0)	1 (3)	
Other	11 (3)	4 (3)	3 (2)	2 (3)	1 (3)	1 (3)	

Values are presented as mean ± standard deviation, n (%), or median (interquartile range). *ASA*, American Society of Anesthesiologists; *FEV_1*, forced expiratory volume in 1 second; *DLCO*, diffusion capacity of carbon monoxide; *PET*, positron emission tomography; *MIS*, minimally invasive surgery (eg, video-assisted thoracoscopic surgery or robotic); *AIS*, adenocarcinoma in situ; *NSCLC*, non–small cell lung cancer; *NOS*, not otherwise specified.

effects. A post hoc analysis calculated the median odds ratio as another measure of across-hospital variation in IMS.[19] Missing data analysis revealed no missing values among variables included in risk and reliability adjustment.

A prespecified sensitivity analysis was conducted in 2 different subgroups in an attempt to evaluate hospital-level variability in IMS rates among populations more likely to be candidates for IMS. Specifically, 1 subgroup analysis was restricted to patients who underwent preoperative positron emission tomography (PET) on the presumption that the provider had a higher index of suspicion of NSCLC. Another subgroup analysis restricted to patients who underwent an anatomic resection on the presumption that surgeons preferentially perform an anatomic resection for cases that are certain to be NSCLC.

RESULTS

Between 2011 and 2013, 406 patients met criteria for this study from across 5 hospitals in the Puget Sound region of Washington State. Clinical and treatment variables are summarized in Table 1. Preoperative patient factors that varied across hospitals included insurance case-mix ($P < .001$), American Society of Anesthesiologists class ($P < .01$), and clinical stage IA disease ($P < .001$). Age, sex, race, Charlson comorbidity index, and lung function did not vary significantly across hospitals (all P values $> .05$). Operative factors (extent and approach to resection) varied significantly across hospitals ($P \leq .001$), as did pathologic stage ($P = .008$), but histology did not ($P = .95$). Nine percent of patients were ultimately diagnosed with pathologic N2 disease.

Overall, 268 patients underwent IMS (66% [95% CI, 61%-71%]). Mediastinoscopy accounted for a majority (85%) of IMS procedures, and a majority (64%) of patients underwent IMS on the day of resection. The median number of mediastinal lymph node stations sampled was 3 (range, 0-7). Seven percent of patients (95% CI, 4%-10%) underwent IMS but did not have lymph node tissue sampled based on pathologic evaluation. Of 18 patients who did not have lymph node tissue sampled, 11 underwent mediastinoscopy, 6 underwent EBUS, and 1 underwent both mediastinoscopy and EBUS. Table 2 summarizes patterns of IMS across hospitals in this study. The type of IMS modality varied across hospitals ($P < .001$), as did the timing of IMS ($P < 0.001$). Unadjusted rates of IMS varied significantly across hospitals ($P < .001$). Hospital B performed IMS without any lymph node tissue sampled 2 times more often (14%) than the overall average.

Adjusted rates of IMS varied widely ranging from 17% (95% CI, 7%-36%) at Hospital E to 94% (95% CI, 89%-96%) at Hospital A (Figure 1). Hospitals A, B, and D had markedly higher rates of IMS compared with Hospitals C and E. Hospital A and B performed IMS significantly more often than the overall average, whereas hospitals C and E performed IMS significantly less often than the overall average. Adjusting for clinical stage IA disease explained only 3% of the variability across hospitals, despite large variation in the frequency of clinical stage

TABLE 2. Patterns of invasive cancer staging for patients undergoing resection of lung cancer across hospitals

	All (n = 406)	Hospital A (n = 128)	Hospital B (n = 132)	Hospital C (n = 70)	Hospital D (n = 40)	Hospital E (n = 30)	*P* value
Invasive mediastinal staging							<.001
None	138 (34)	11 (8)	34 (25)	56 (78)	10 (25)	27 (90)	
Mediastinoscopy/mediastinotomy only	228 (56)	96 (74)	87 (65)	13 (18)	29 (73)	3 (10)	
EBUS only	9 (2)	2 (2)	5 (4)	1 (1)	1 (3)	0 (0)	
EUS only	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
EBUS and mediastinoscopy/mediastinotomy	30 (7)	21 (16)	7 (5)	2 (3)	0 (0)	0 (0)	
EBUS and EUS and mediastinoscopy/mediastinotomy	1 (<1)	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)	
Any invasive mediastinal staging	268 (66)	119 (92)	100 (75)	16 (22)	30 (75)	3 (10)	<.001
Timing of invasive staging*							<.001
Before resection	97 (36)	3 (3)	87 (87)	4 (25)	1 (3)	2 (67)	
At the time of resection	147 (55)	95 (80)	12 (12)	10 (63)	29 (97)	1 (33)	
Both before and at the time of resection	24 (9)	21 (18)	1 (1)	2 (13)	0 (0)	0 (0)	
Number of nodal stations sampled*							
Mean ± standard deviation	3.0 ± 1.2	3.2 ± 1.0	2.7 ± 1.4	2.9 ± 1.3	3.2 ± 1.4	2.7 ± 0.6	.05
Median (interquartile range)	3 (0-7)	3 (0-7)	3 (0-5)	3 (0-5)	3 (0-5)	3 (2-3)	.34
No nodes sampled	18 (7)	1 (1)	14 (14)	1 (6)	2 (6)	0 (0)	.004

Values are presented as n (%) unless otherwise indicated. *EBUS*, Endobronchial ultrasound-guided nodal aspirate; *EUS*, esophageal ultrasound-guided nodal aspirate. *Calculated among patients who underwent invasive mediastinal staging.

IA disease across hospitals. Our prespecified subgroup analyses revealed the same degree of variability in hospital-level rates of IMS among patients who underwent anatomic resection or underwent a PET scan (Figures 2 and 3).

Several post hoc analyses were conducted. As an alternative measure of variability in hospital-level rates of IMS, we estimated the median odds ratio across hospitals to be 4.99 (95% CI, 1.56-7.24), indicating significant across-hospital variance. This value is interpreted as follows: if a randomly selected patient from 1 center were to transfer care to a hospital with a higher probability of performing IMS, then his or her risk of undergoing IMS would increase by 499%. Table 3 summarizes a post hoc analysis exploring potential differences in patient characteristics between those who did and did not undergo IMS. Compared with those who did not receive IMS, patients who underwent IMS had a different payer mix ($P < .001$) and were more likely to have a higher American Society of Anesthesiologists class ($P = .002$), smoke ($P = .002$), have a lower predicted forced expiratory volume in 1 second ($P = .03$), have undergone a preoperative PET scan ($P < .001$), and have a higher clinical stage ($P = .001$). A fully adjusted model did not change the conclusions of our primary analysis, and it revealed that patient-level variables accounted for a total of 5% of the variation in rates of IMS across hospitals. Another patient-level stratified analysis showed that patients with clinical stage IB or higher disease had higher rates of IMS compared with those with clinical stage IA (78% vs 61%; $P < .001$). Finally, hospitals with an exclusively thoracic surgical-based practice model (A, B, and D) had higher rates of IMS compared with hospitals with a mixed-specialty practice model (82% vs 19%; $P < .001$).

DISCUSSION

We hypothesized that there would be significant variation in hospital-level rates of IMS and aimed to compare hospital-specific rates of IMS to each other and to the overall population mean. The key finding of our study is that there is significant variation in rates of IMS across hospitals providing surgical care to patients with NSCLC. Patient-level factors, including clinical stage IA disease, explain very little of the variability. Post hoc analyses confirm our key finding. Hospitals with a purely thoracic surgical practice (as opposed to a mixed-specialty practice) were associated with higher rates of IMS.

Variability in care is often considered a marker of poor quality care. Surgeons may not adhere to guidelines because they are unaware that they exist. Those with a practice solely dedicated to thoracic surgery may be more likely to know guideline recommended indications for IMS. Supporting this hypothesis is the observation that an exclusively thoracic surgical-based hospital practice model was associated with higher rates of IMS. However, a previous study examining the relationship between surgeon specialty and outcomes reported no differences in the use of mediastinoscopy among general, cardiothoracic, and thoracic surgeons.[20] Survey research may better elucidate the frequency of surgeon awareness of guideline recommendations and factors associated with both awareness and adherence.

Another possible explanation for variation in IMS is that surgeons may purposefully choose to not adhere to guidelines because they do not accept the recommendations to be valid. A previous study showed that providers may deviate from guidelines when either the level of

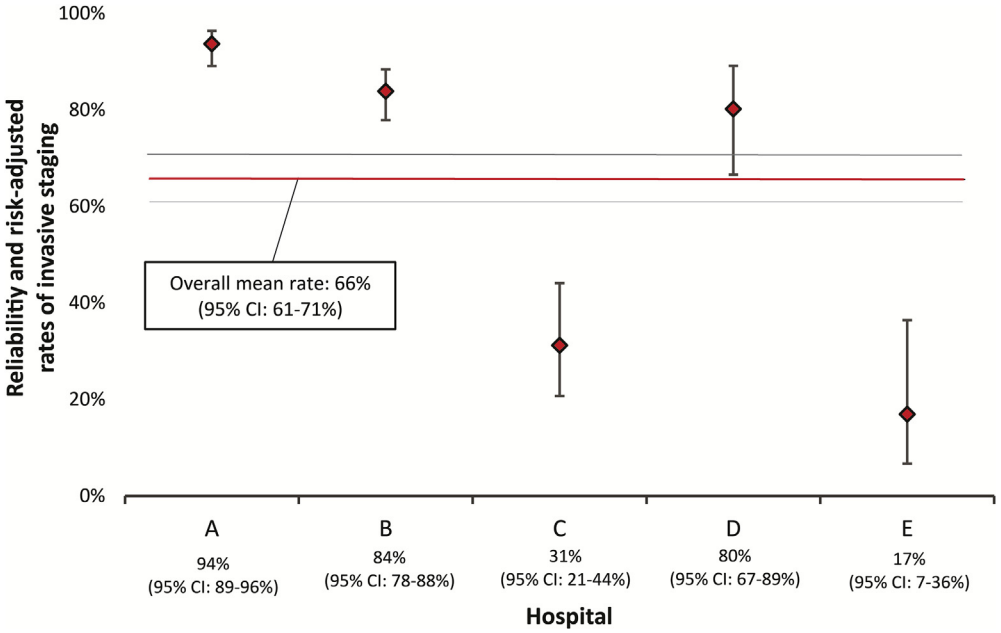


FIGURE 1. Rates of invasive mediastinal staging by hospital compared with population average. *CI*, Confidence interval.

supporting evidence for guidelines is low or when they perceive the supporting evidence to not be generalizable to their patient population or practice.[21] Indeed the level of evidence for guideline recommended indications for IMS is 2.[9-11] There are no randomized controlled trials comparing varying intensities of IMS (ie, routine vs selective vs highly selective IMS). It was recently shown that guideline-recommended IMS has a sensitivity and specificity of 100% and 35%, respectively.[22,23] Clinically, these estimates mean that existing guideline recommendations perfectly select all patients with true nodal disease to undergo IMS, but they also select many patients without true nodal disease to undergo IMS. The high rate of negative IMS may provide the rationale for why some surgeons do not accept guideline recommendations. A high true negative rate of IMS may be perceived as futile, inefficient, or cost-ineffective,[24] although recent evidence suggests these perceptions may not be correct.[25] However, underlying the rationale for guidelines directing a majority of lung cancer patients to IMS is the view that the identification of mediastinal nodal disease (N2 or N3, Stage IIIA or IIIB) is important from a patient and treatment selection perspective. Even if the multimodality options for stage IIIA are equivalent with regard to long-term survival,[26] many patients and providers would likely agree that

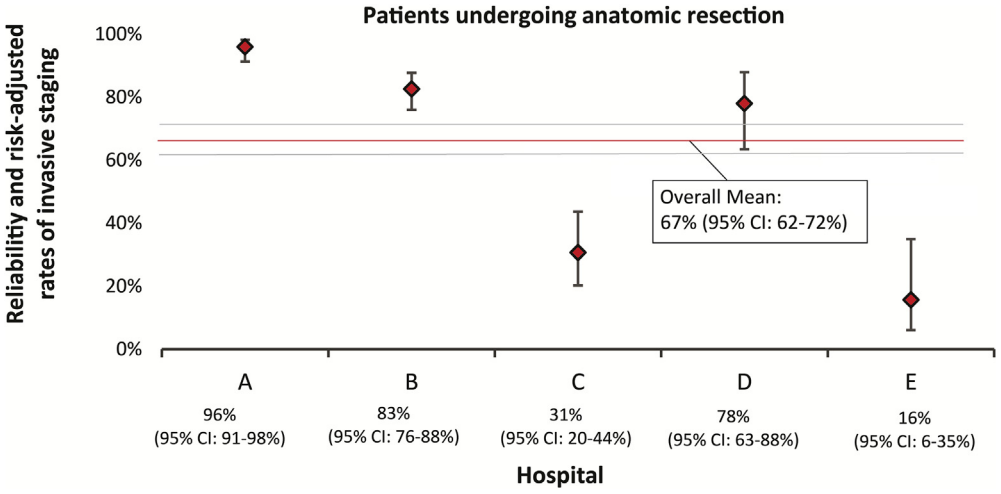


FIGURE 2. Rates of invasive mediastinal staging by hospital (among patients undergoing anatomic resection). *CI*, Confidence interval.

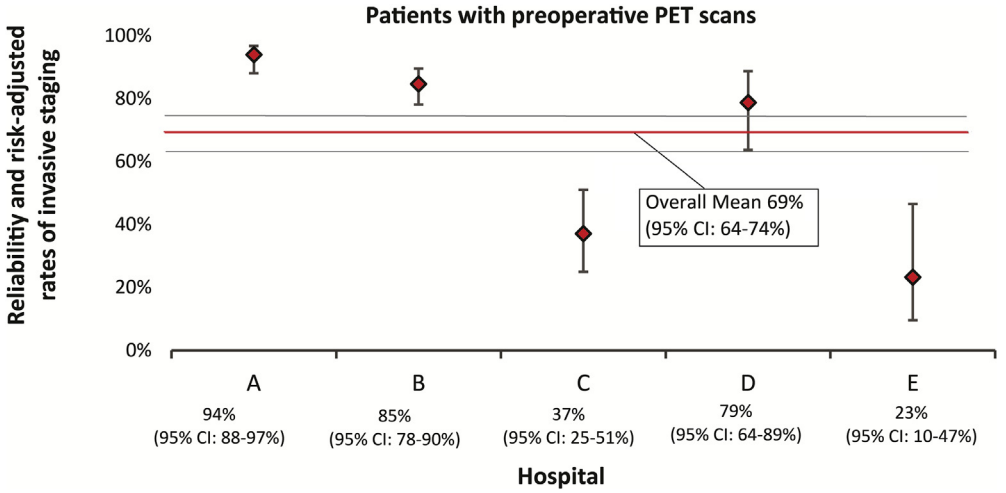


FIGURE 3. Rates of invasive mediastinal staging by hospital (among patients with preoperative positron emission tomography [PET] scans). *CI*, Confidence interval.

more accurate staging leads to more informed treatment decisions. Nonetheless, some providers may be skeptical of this view and believe that patients are better served by initial surgical resection.

Better understanding the reasons underlying hospital-level variability in IMS is an opportunity to improve care. Qualitative research (eg, focus groups and key informant surgeon interviews) can help identify the reasons for variability in practice patterns. To the extent that a knowledge gap explains variability in IMS, hospital credentialing and privileging policies could be an avenue for evaluating and increasing the knowledge base of surgeons who do not have a dedicated oncologic practice and/or board certification in thoracic surgery.[27] Greater access to continuing medical education—such as the GAIN (Engaging an Interdisciplinary Team for NSCLC Diagnosis, Personalized Assessment, and Treatment) curriculum sponsored by the American College of Chest Physicians, National Comprehensive Cancer Network programs, and others—is another opportunity to address knowledge gaps.[28] Finally, clinical registries (eg, STS and National Surgical Quality Improvement Program) can facilitate performance feedback in terms of process compliance (ie, benchmarked rates of IMS).[29-31]

On the other hand, if the variability in IMS is due to widespread disagreement about the optimal indications for IMS, then more research will be needed to elucidate the optimal IMS strategy. One example of an alternative staging strategy is selective IMS using prediction models. Preliminary research suggests that prediction models can, similar to existing practice guidelines, maintain 100% sensitivity while achieving higher specificity.[22,32] Clinically, this means that the use of a prediction model is expected to lead to fewer IMS procedures without compromising the ability to detect true nodal disease before first treatment. This hypothesis could also be tested through a randomized controlled trial. Still other alternative staging strategies may arise from qualitative interviews with thoracic surgeons and pulmonologists.

One unexpected finding was that the rate of IMS in this study (66%) was higher than previously reported (27%).[1] Differences across studies may reflect differences in the quality of data collection or delivery of care over time. Another possibility is the Hawthorne effect, given that all hospitals agreed to participate in a regional quality improvement collaborative. However, this explanation is unlikely because the decision to measure use of IMS was made by our regional collaborative in 2013, the study period spans 2011 to 2013, and there were no interventions intended to influence rates of IMS. Furthermore, rates of IMS for the entire study population were lower than expected population rates of IMS derived from prior studies (77%-79%).[22,23,33] In addition, although rates of failed lymph node sampling (7%) were much lower than that previously reported ($\sim$ 47%),[1] opportunities for quality improvement remain given that 1 hospital failed sample lymph node tissue in up to 14% of patients undergoing IMS—a rate double that of other hospitals. Failing to submit or obtain adequate lymph node tissue is an outcome measure that could be monitored by a clinical registry such as the STS General Thoracic Surgery Database.[30,31]

There are important limitations to this study. There may be concerns about generalizing regional results to the rest of the nation. By design, this multicenter study included mostly community hospitals and only a single academic center in an attempt to maximize generalizability. We expect hospital-level variability in IMS to be found across the nation, and several members of our study team are currently testing this hypothesis using the STS General Thoracic Surgery Database (although this database also

TABLE 3. Characteristics of patients receiving and not receiving invasive staging

	All (n = 406)	Invasive staging (n = 265)	No invasive staging (n = 137)	*P* value
Mean age (y)	68.2 ± 10.0	68.0 ± 10.4	68.8 ± 9.3	.48
Women	228 (57)	150 (56)	78 (57)	.95
Race				.33
White	354 (88)	231 (87)	123 (90)	
Black	7 (2)	6 (2)	1 (1)	
Asian	26 (6)	17 (6)	9 (7)	
Native American	4 (1)	4 (2)	0 (0)	
Pacific Islander	3 (1)	3 (1)	0 (0)	
Other	8 (2)	4 (2)	4 (3)	
Insurance				<.001
Commercial	145 (36)	89 (33)	46 (41)	
Medicare	77 (19)	35 (13)	42 (30)	
Medicaid	13 (3)	11 (4)	2 (1)	
Tricare (military)	2 (<1)	2 (1)	0 (0)	
Other government	3 (1)	3 (1)	0 (0)	
Uninsured	2 (<1)	1 (<1)	1 (1)	
Government and commercial	126 (31)	97 (36)	29 (21)	
Mixed government	38 (9)	30 (11)	8 (6)	
Charlson comorbidity index				.36
0	227 (56)	148 (55)	79 (57)	
1	141 (35)	90 (34)	51 (37)	
2	28 (7)	22 (8)	6 (4)	
3+	10 (2)	8 (3)	2 (1)	
ASA class				.002
II	71 (17)	34 (13)	37 (27)	
III	312 (77)	217 (81)	95 (69)	
IV	23 (6)	17 (6)	6 (4)	
Smoking status				.002
Former	235 (60)	159 (62)	76 (57)	
Current	96 (25)	70 (27)	26 (19)	
Never	60 (16)	28 (11)	32 (24)	
% Predicted FEV_1	83 (69-96)	80 (69-94)	85 (72-100)	.03
% Predicted DLCO	64 (52-77)	63 (51-76)	66 (56-80)	.15
Clinical stage				.001
IA	281 (69)	170 (63)	111 (80)	
IB	54 (13)	36 (13)	18 (13)	
IIA	33 (8)	25 (9)	8 (6)	
IIB	20 (5)	20 (7)	0 (0)	
IIIA	15 (4)	14 (5)	1 (1)	
IIIB	1 (<1)	1 (<1)	0 (0)	
IV	2 (<1)	2 (1)	0 (0)	
Preoperative PET scan	371 (92)	255 (96)	116 (85)	<.001
Extent of resection				.02
Wedge	35 (9)	20 (7)	15 (11)	
Segmentectomy	7 (2)	6 (2)	1 (1)	
Lobectomy	272 (67)	180 (67)	92 (67)	
Sleeve lobectomy	2 (<1)	0 (0)	2 (1)	
Lobectomy and wedge	66 (16)	40 (15)	26 (19)	
Bilobectomy	14 (3)	12 (4)	2 (1)	
Pneumonectomy	10 (2)	10 (4)	0 (0)	
MIS approach	292 (72)	201 (75)	91 (66)	.05

(Continued)

TABLE 3. Continued

	All (n = 406)	Invasive staging (n = 265)	No invasive staging (n = 137)	*P* value
Right-sided resection	247 (61)	163 (61)	84 (61)	.77
Pathologic stage				.01
0	5 (1)	1 (<1)	4 (3)	
IA	201 (50)	121 (45)	80 (58)	
IB	84 (21)	56 (21)	28 (20)	
IIA	38 (9)	26 (10)	12 (9)	
IIB	32 (8)	26 (10)	6 (4)	
IIIA	41 (10)	34 (13)	7 (5)	
IV	5 (1)	4 (1)	1 (1)	
Pathologic nodal stage				.003
pN0	317 (79)	198 (74)	119 (89)	
pN1	47 (12)	37 (14)	10 (7)	
pN2	36 (9)	31 (12)	5 (4)	
Histology				.049
Adenocarcinoma	270 (67)	174 (65)	96 (70)	
AIS	10 (2)	5 (2)	5 (4)	
Squamous cell carcinoma	86 (21)	67 (25)	19 (14)	
Large cell carcinoma	7 (2)	5 (2)	2 (1)	
NSCLC NOS	22 (5)	10 (4)	12 (9)	
Other	11 (3)	7 (3)	4 (3)	

Values are presented as mean ± standard deviation or n (%). *ASA*, American Society of Anesthesiologists; *FEV_1*, forced expiratory volume in 1 second; *DLCO*, diffusion capacity of carbon monoxide; *PET*, positron emission tomography; *MIS*, minimally invasive surgery (eg, video-assisted thoracoscopic surgery or robotic); *AIS*, adenocarcinoma in situ; *NSCLC*, non–small cell lung cancer; *NOS*, not otherwise specified.

VIDEO 1. Dr Farhood Farjah explains the significance of the research findings. Video available at: http://www.jtcvsonline.org/article/S0022-5223(18)30311-8/fulltext.

has generalizability concerns). Because we were unable to directly access medical records to determine the indications for or against IMS, we could not measure adherence with guideline recommendations at the patient level. Further, because we were unable to measure central versus peripheral tumors, some of the variation in IMS may be explained by differences in case-mix (defined by central tumors) across hospitals. Neither did we have information on the provider who performed or decided against performing IMS, and therefore we cannot better understand the contribution of individual providers or provider-specialty to variation in care. We may have underestimated rates of IMS. Data abstractors may not have been able to record IMS occurring at a referring hospital/health system (ie, not participating in our collaborative). However, given the fact that rates of IMS are higher than previously described in community practice, it is unlikely that we significantly underestimated rates of IMS. Finally, we were unable to elucidate the reasons for variable rates of IMS across hospitals. As discussed earlier, we plan to conduct qualitative and mixed-methods research to gain a better understanding of the determinants of variability in IMS.

CONCLUSIONS

We describe—for the first time—that there is significant variability in rates of IMS across hospitals (Video 1). We speculate that this variability might be due to gaps in the quality of care, disagreement over the best indications for IMS, or both. Planned qualitative and mixed-methods

studies will likely reveal the reasons underlying variability and lead to either quality improvement and educational interventions and/or trials comparing novel IMS strategies to guide recommended staging.

Conflict of Interest Statement

Authors have nothing to disclose with regard to commercial support.

The authors thank the Surgical Care and Outcomes Assessment Program for providing data for this research. The authors also thank Rebecca G. Symons; Megan Zadworny; Hao He, PhD; and Katherine Odem-Davis, PhD, at the Surgical Outcomes Research Center for helping with data acquisition and statistical consultation.

References

1. Little AG, Rusch VW, Bonner JA, Gaspar LE, Green MR, Webb WR, et al. Patterns of surgical care of lung cancer patients. *Ann Thorac Surg*. 2005;80:2051-6.
2. Little AG, Gay EG, Gaspar LE, Stewart AK. National survey of non–small cell lung cancer in the United States: epidemiology, pathology and patterns of care. *Lung Cancer*. 2007;57:253-60.
3. Farjah F, Flum DR, Ramsey SD, Heagerty PJ, Symons RG, Wood DE. Multi-modality mediastinal staging for lung cancer among medicare beneficiaries. *J Thorac Oncol*. 2009;4:355-63.
4. Lathan CS, Neville BA, Earle CC. The effect of race on invasive staging and surgery in non-small-cell lung cancer. *J Clin Oncol*. 2006;24:413-8.
5. Vest MT, Tanoue L, Soulos PR, Kim AW, Detterbeck F, Morgensztern D, et al. Thoroughness of mediastinal staging in Stage IIIA non–small cell lung cancer. *J Thorac Oncol*. 2012;7:188-95.
6. 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;145:1097-113.
7. Ost DE, Niu J, Elting LS, Buchholz TA, Giordano SH. Quality gaps and comparative effectiveness in lung cancer staging and diagnosis. *Chest*. 2014;145:331-45.
8. Faris N, Yu X, Sareen S, Signore RS, McHugh LM, Roark K, et al. Preoperative evaluation of lung cancer in a community health care setting. *Ann Thorac Surg*. 2015;100:394-400.
9. National Comprehensive Cancer Network. Non–small cell lung cancer (Version 5.2017). Available at: https://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf. Accessed March 18, 2017.
10. Silvestri GA, Gonzalez AV, Jantz MA, Margolis ML, Gould MK, Tanoue LT, et al. Methods for staging non-small cell lung cancer. *Chest*. 2013;143:e211S-50S.
11. Ettinger DS, Akerley W, Borghaei H, Chang AC, Cheney RT, Chirieac LR, et al. Non–small cell lung cancer, version 2.2013. *J Natl Compr Cancer Netw*. 2013;11:645-53.
12. Comparative Effectiveness Research Translation Network. Lung cancer quality improvement collaborative. Available at: http://www.becertain.org/projects/cancer-care/lung-cancer. Accessed March 14, 2017.
13. Devine EB, Capurro D, van Eaton E, Alfonso-Cristancho R, Devlin A, Yanez ND, et al. Preparing electronic clinical data for quality improvement and comparative effectiveness research: The SCOAP CERTAIN automation and validation project. *EGEMS (Wash, DC)*. 2013;1:1025.
14. Flum DR, Alfonso-Cristancho R, Devine EB, Devlin A, Farrokhi E, Tarczy-Hornoch P, et al. Implementation of a "real-world" learning health care system: Washington state's comparative effectiveness research translation network (CERTAIN). *Surgery*. 2014;155:860-6.
15. Dimick JB, Staiger DO, Birkmeyer JD. Ranking hospitals on surgical mortality: the importance of reliability adjustment. *Health Serv Res*. 2010;45(6 Part 1):1614-29.
16. Dimick JB, Ghaferi A, Osborne NH, Ko CY, Hall BL. Reliability adjustment for reporting hospital outcomes with surgery. *Ann Surg*. 2012;255:703-7.
17. MacKenzie TA, Grunkemeier GL, Grunwald GK, O'Malley AJ, Bohn C, Wu YX, et al. A primer on using shrinkage to compare in-hospital mortality between centers. *Ann Thorac Surg*. 2015;99:757-61.
18. Detterbeck FC, Jantz MA, Wallace M, Vansteenkiste J, Silvestri GA. Invasive mediastinal staging of lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest*. 2007;132(3 Suppl):202S-20S.
19. Larsen K, Merlo J. Appropriate assessment of neighborhood effects on individual health: Integrating random and fixed effects in multilevel logistic regression. *Am J Epidemiol*. 2005;161:81-8.
20. Farjah F, Flum DR, Varghese TK, Symons RG, Wood DE. Surgeon specialty and long-term survival after pulmonary resection for lung cancer. *Ann Thorac Surg*. 2009;87:995-1006.
21. Brouwers MC, Makarski J, Garcia K, Akram S, Darling GE, Ellis PM, et al. A mixed methods approach to understand variation in lung cancer practice and the role of guidelines. *Implement Sci*. 2014;9:36.
22. Farjah F, Backhus LM, Varghese TK, Manning JP, Cheng AM, Mulligan MS, et al. External validation of a prediction model for pathologic N2 among patients with a negative mediastinum by positron emission tomography. *J Thorac Dis*. 2015;7:576-84.
23. Farjah F, Madtes DK, Wood DE, Flum DR, Zadworny ME, Waworuntu R, et al. Vascular endothelial growth factor C complements the ability of positron emission tomography to predict nodal disease in lung cancer. *J Thorac Cardiovasc Surg*. 2015;150:796-803.
24. Meyers BF, Haddad F, Siegel BA, Zoole JB, Battafarano RJ, Veeramachaneni N, et al. Cost-effectiveness of routine mediastinoscopy in computed tomography- and positron emission tomography-screened patients with stage I lung cancer. *J Thorac Cardiovasc Surg*. 2006;131:822-9.
25. Czarnecka-Kujawa K, Rochau U, Siebert U, Atenafu E, Darling G, Waddell TK, et al. Cost-effectiveness of mediastinal lymph node staging in non–small cell lung cancer. *J Thorac Cardiovasc Surg*. 2017;153:1567-78.
26. Martins RG, D'Amico TA, Loo BW Jr, Pinder-Schenck M, Borghaei H, Chaft JE, et al. The management of patients with stage IIIA non-small cell lung cancer with N2 mediastinal node involvement. *J Natl Compr Canc Netw*. 2012;10:599-613.
27. Blackmon SH, Cooke DT, Whyte R, Miller D, Cerfolio R, Farjah F, et al. The Society of Thoracic Surgeons expert consensus statement: a tool kit to assist thoracic surgeons seeking privileging to use new technology and perform advanced procedures in general thoracic surgery. *Ann Thorac Surg*. 2016;101:1230-7.
28. CHEST GAIN. Regional non–small cell lung cancer summits. Available at: http://www.chestnet.org/Education/Advanced-Clinical-Training/CHEST-GAIN-NSCLC. Accessed November 4, 2017.
29. American College of Surgeons. NSQIP data collection, analysis, and reporting. Available at: https://www.facs.org/quality-programs/acs-nsqip? Accessed September 29, 2016.
30. Grover FL, Shahian DM, Clark RE, Edwards FH. The STS National Database. *Ann Thorac Surg*. 2014;97:S48-54.
31. Jacobs JP, Shahian DM, Prager RL, Edwards FH, McDonald D, Han JM, et al. Introduction to the STS National Database Series. *Ann Thorac Surg*. 2015;100:1992-2000.
32. Farjah F, Lou F, Sima C, Rusch VW, Rizk NP. A prediction model for pathologic N2 disease in lung cancer patients with a negative mediastinum by positron emission tomography. *J Thorac Oncol*. 2013;8:1170-80.
33. Mazzone PJ, Vachani A, Chang A, Detterbeck F, Cooke D, Howington J, et al. Quality indicators for the evaluation of patients with lung cancer. *Chest*. 2014;146:659-69.

Key Words: lung cancer staging, lymph nodes, diagnostics

Discussion

Moderator: David T. Cooke

THOR

Moderator: Murray H. Kwon

Dr Farjah. Good morning and thank you for the opportunity to present our work. Dr. Thornblade very much wanted to be here to share these results with you but he and his wife may be having their first child a few days early so he's in Seattle and I'm very happy to present on his behalf. We have no disclosures or conflicts of interest to disclose.

Staging is among the most important aspects of cancer care, and that's because accurate staging leads to appropriate treatment selection and appropriate treatment selection leads to optimal patient outcomes. Invasive mediastinal staging is important because of the well-characterized false negative rate of imaging and because the results of invasive staging have a profound influence on treatment recommendations, which can vary from surgery alone to definitive chemoradiation therapy. In 2005, this study presented at the Society of Thoracic Surgeons annual meeting reported a 27% rate of invasive mediastinal staging, a rate far below expectations, and this key finding of an apparent underuse of invasive staging has been observed over and over again in different populations and using different data sources, and it has led to substantial concerns over the quality of thoracic surgical care in the community at large. Variability in invasive staging has not been adequately described, but there is reason to suspect that it exists. Some surgeons and pulmonologists may not know about the guidelines that advise us how to stage patients. Others may disagree with those guidelines. Some might find invasive staging to be inconvenient or an inefficient use of time, or time-consuming, and yet others may not know how to do the broad spectrum of invasive staging procedures or even fear doing some of them. So, leveraging a unique data source on thoracic surgical practice patterns in Washington State, and specifically in the Puget Sound region, we hypothesized that there is significant variation in the rate of invasive staging across hospitals. We conducted a retrospective cohort study of adult patients with non–small cell lung cancer who underwent resection between July 2011 and December 2013. Data were ascertained from medical records maintained by hospitals participating in a National Cancer Institute-funded collaborative that aimed to improve Native American cancer outcomes within the region. The clinical registry that resulted from that collaborative collected information on all patients with lung cancer regardless of race, and included information on demographic characteristics, cancer care, tumor characteristics, and outcomes. Patients were excluded if they had a prior history of lung cancer, or if they were treated with induction therapy or if they were cared for at a hospital that performed < 10 cases over the study period. A total of 406 patients were eligible for the study from 5 hospitals and all 5 hospitals included in this study were staffed by at least 1 Board-certified thoracic surgeon with a noncardiac practice. So, we used the mixed-effects model to obtain estimates and confidence intervals for hospital-level rates of invasive staging.

This approach allowed us to safeguard against inflated estimates of variability that could arise from small sample sizes at any 1 hospital. This methodology also allowed us to adjust for variation in the prevalence of stage IA disease across hospitals and this is important because guidelines allow for omission of invasive staging among patients with peripheral stage IA tumors. So, performing the adjustment safeguards against potential confounding. The study cohort had features typical of a lung cancer population with the mean age 68 years; 57% were women; 88% were white; 77% of people were American Society of Anesthesiologists class 3; 84% were former or current smokers; and a majority of patients had clinical stage IA disease, adenocarcinoma, and underwent a lobectomy. Overall, 66% of patients underwent any invasive mediastinal staging and among the 268 who did undergo staging, the most common procedure performed was mediastinoscopy alone, followed by endobronchial ultrasound guided nodal aspirate and mediastinoscopy. Invasive staging was performed at the time of resection in 2 out of 3 cases. The median number of mediastinal nodal stations sampled was 3, and 7% of patients who underwent invasive staging did not have any lymph node tissue sent to pathology.

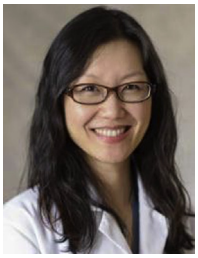
Going back to the cohort of 406 patients, the prevalence of clinical stage IA disease at presentation did vary across hospitals, as we suspected, and significantly so. The only other variables that differed across hospitals was insurance and American Society of Anesthesiologists class. We found evidence of significant variation in hospital-level rates of invasive staging. The *x*-axis shows the de-identified hospitals, the *y*-axis shows the adjusted rates of invasive staging. The horizontal line shows the overall mean of 66% with the 95% confidence intervals associated with that mean, and the vertical bars are the 95% confidence intervals for the hospital-specific rates. And you can see that relative to the mean, 2 hospitals performed invasive staging more often than the mean, 2 hospitals performed it less often than the mean, and 1 hospital performed it no different than the mean. Another way to look at variability with respect to some sort of a reference point or a referent, is to compare it with expected

rates of invasive staging. Three studies recently provided estimates of the expected rate of invasive staging in various populations of patients with lung cancer, and amazingly, the estimates are very tight. They range from 77% to 79%. So, if we use these values as our referent, you can see that 1 hospital performed invasive staging more often than expected, 2 performed it less often than expected, and 2 performed it no different than expected.

We found evidence of 5-fold variation in hospital level rates of invasive mediastinal staging that was not explained by chance or case mix. Our study had several limitations, and the first is generalizability. Care within our region may not represent care across the nation. However, with 1 academic hospital and 4 community sites, we believe our study is likely to be generalizable to the community at large as long as there is 1 Board-certified thoracic surgeon with a noncardiac practice at that hospital. Another limitation is that the clinical registry did not collect enough information for us to directly calculate expected rates of invasive staging because there was no information on imaging results such as lymphadenopathy, fluorodeoxyglucose-uptake in lymph nodes, tumor size, or central or peripheral tumors. Lastly to the extent that invasive mediastinal staging occurred before referral to 1 of the 5 study hospitals, we may have underestimated the rate of invasive staging, but because our rate of invasive staging was 2 times higher than previous reports, we think that is unlikely but a possibility.

In conclusion, our understanding of the apparent underuse of invasive staging is further complicated by the fact that there are highly variable patterns of invasive staging, and notable at hospitals staffed by at least 1 Board-certified thoracic surgeon with a noncardiac practice. This variability could be a marker of poor quality care because guidelines are not supported by level-1 evidence, it is equally plausible that this variability might represent uncertainty or even disagreement with the practice guidelines, and specifically about the appropriate indications for invasive staging. So, several members of our team have taken steps to further this line of investigation. Another was completed at the national level using Society of Thoracic Surgeons data and that has been submitted to a national meeting for presentation. We are planning provider interviews and focus groups to try to understand the drivers of variability in invasive staging among surgeons. Finally, our group recently developed a risk-based approach to selective invasive staging using a prediction model that predicts the probability of nodal disease, and some of our modeled analyses show that this strategy is as good a guideline in identifying people who truly have nodal disease, but its use could lead to a reduction in the use of invasive staging procedures. Accordingly, we are now planning a multicenter validation study of this alternative strategy.

I would like to acknowledge the considerable efforts of Dr Thornblade, a general surgery resident at our institution and a recipient of a National Institutes of Health T32 Post-Doctoral Award. We would also like to acknowledge the University of Washington Surgical Outcomes Research Center and the Washington State Surgical Care and Outcomes Assessment Program for their contributions in terms of data acquisition.



Dr Jane Yanagawa . It is an interesting study because historically, lower rates of mediastinoscopy are automatically assumed to be a reflection of low-quality care and you suggest that it might not be the case, that it might be more complicated than that. One situation I thought of that might lead to a lower rate of mediastinoscopy is a patient who by all preoperative imaging is believed to have an early stage lung cancer and you might choose to avoid the morbidity, time, and cost of a needle biopsy, to go instead straight to the operating room for a wedge resection for diagnosis and then perhaps go straight to a completion lobectomy based on frozen section results. In that case, hopefully they would at least get a thorough ipsilateral lymph node staging. That patient would obviously not get a mediastinoscopy. Do you know if scenarios like that or similar to that might have influenced the results in your study?

Dr Farjah. Thanks, Jane. I think it could have. I think it could explain this variability and the reason why I say that is not because we can actually measure who did and did not get preoperative histology with this data set, but the patient that you described would not have undergone invasive staging, and it sounds like that is what you would do at your institution. It is interesting that you bring it up because at our institution if we strongly suspect lung cancer, and we do not have preoperative tissue confirmation, we will actually do the mediastinoscopy before the lung resection. What you are describing, and what I know at least at our own institution, I would say that may be a driver of variability**.**

Dr Yanagawa. Another consideration you brought up is that your study was performed at a hospital level, but I know at our institution, every surgeon practices some things differently. Do you think that if people are not following the guidelines, that it would be at the hospital level? When someone needs preoperative mediastinoscopy, is it something that is discussed at tumor boards?

Dr Farjah. As we start to move toward reimbursement for disease-based care, the looking at it from the perspective of an institution and a patient makes sense. That being said, I'm sure there is variability at the individual provider level and I think that it could contribute to the variability seen in this study. At our institution, there is probably some variability, but more or less we definitely have a pathway that

is closely aligned with National Comprehensive Cancer Network guidelines. If you looked at our intersurgeon practice patterns, we probably would not be too different. Most surgeons are more or less institutionally aligned with the National Comprehensive Cancer Network guidelines. I cannot speak for the other centers in the study.

Dr Yanagawa. The last thing I would like you to comment about is the nodes harvested. One of the hospitals had the highest rate of mediastinoscopy, but a center with the highest rate of the mediastinoscopy had no lymph nodes actually harvested during the procedure. Why might that be the case?

Dr Farjah. I think that was hospital B, and I think their invasive staging rate was something like 75%. I do not think it was the highest, but it was high. And the rate of no lymph nodes from the staging procedure was something like 14% or 15%. I think this highlights a problem that is well-known in the world of quality measurement. If you look at rates of invasive staging, which is a process measure, and looking at process measures makes a lot of sense when you have very high-level data. It is akin to an internal mammary artery for coronary bypass. If you do not have the level-1 evidence and there is a lot of consensus, then it probably makes sense. We see a lot of variability, so maybe there is not a lot of consensus, and when you look at a process measure and you do not look at outcomes, you can get a full signal about quality and that is an unintended consequence of any quality measurement or quality improvement program that only looks at 1 metric. If we only looked at rates of invasive staging, we would say that hospital did great, but obviously, if 15% did not have lymph nodes harvested, then is doing procedures way better? Research by Litle found 50% of lymph nodes harvested may not be sufficient.

Dr Paul Shipper. Two quick questions. The first is a follow-up. You said that you corrected for the frequency of stage IA and I would have approached a 1A nodule the way that Dr Yanagawa described, with a wedge and then a thoracoscopic invasive staging. But then you would have counted that against me. That would be counted as no invasive staging procedure. Is that correct?

Dr Farjah. Can you repeat that last part again, I want to make sure I understand it.

Dr Shipper. When you say that you looked at the frequency of stage IA and corrected for it, I interpreted that as meaning that you corrected for the situation that Dr Yanagawa described where you operate knowing that it is a small tumor and you wedge it out. Then you do your staging once you know you have cancer. But your answer made me believe that that actually isn't the case, that if we did that, those would have been nonstaged patients?

Dr Farjah. Your latter statement is correct. But we adjusted for the fact that some hospitals may see more stage I cancers than others. This is akin to case mix and risk adjustment with outcomes. As surgeons, we always say, well my patients are different, so that's why we did that adjustment. But our measurement is still about the rates of invasive staging, and it is not viewed as counting against or for you, it is simply, what are the rates? All we did was adjust for the difference in case mix for a very reasonable reason why someone would not do invasive staging.

Dr Shipper. Second quick question. Endobronchial-ultrasound guided nodal aspirate seemed very low in frequency. Any idea as to why that is?

Dr Farjah. I do not know. I also do not know which hospitals are actually in this study. I think almost all or all of them have an endobronchial-ultrasound guided nodal aspirate capability and either pulmonologists or surgeons are doing endobronchial-ultrasound guided nodal aspirate, so if it is true and these hospitals are actually in the study, I suspect what you are seeing is just a preference among the providers for that invasive staging modality, but at this point I am speculating. We don't know why.

Dr Richard Whyte. Great presentation. A quick addition to that last question: Did you see endobronchial-ultrasound guided nodal aspirate use increase in frequency over the time of your study? But more importantly, I want to applaud you on this whole concept of this collaborative. My question is how do you get this information out amongst the people who participate in this collaborative? We have struggled with doing a surgical collaborative in Massachusetts, but Rich Prager invited me to the Michigan collaborative a year ago and it is remarkable how much cross-fertilization there is in that collaborative. You presented de-identified data, it was not just organization A, B, C, and D. It was hospitals and you know who they are, and peer pressure is a huge thing, so if you are the guy down on the lower right-hand corner of your figure, I guarantee people will go, ''Holy smoke, what are you doing? Why are you not doing more invasive staging,'' and you got to get up there and sort of justify what you do, but I think this is terrific, thanks.

Dr Farjah. Thank you.

Dr Matt Blum *(Colorado Springs, Colo)*. This a neat study and a neat way of kind of looking at things through several different institutions and we did this with the University of Colorado because we now have 3 hospitals that do a substantial amount of pulmonary surgery and the way things were staged were quite different across the hospitals. They all do appropriate staging, but they are quite different, and I was wondering if you could eliminate some of the question or variability for stage I lung cancers, certainly

people use different criteria. You can use the National Comprehensive Cancer Network criteria where it is a 1 cm solid nodule, or if it is smaller than that, maybe you do not need to do mediastinal staging if the computed tomography scan is negative. I mean could you just look at stage II or III patients and then eliminate the whole question of how many stage I patients were staged? Because there is a lot of variability. Some people would not stage less than a 3-cm peripheral nodule invasively if everything else looked clean. I wonder if a lot of the variability is in that judgment of those stage I patients because you had a lot of stage I patients in that group.

Dr Farjah. We actually had planned some subgroup analyses, as well as some postop analyses, but they became challenging because of the small sample sizes. As you pointed out, clinical stage IA, was 69% to 70% of the cohort, so removing those patients would have left us with only 30% of the original population. We could do it, but our confidence limits would increase and our signal-to-noise ratio would go down. In theory, the adjustment for clinical stage IA should account for that? I think a lot of people, including some of our coauthors, were unhappy because it was not particularly transparent. So performing a subgroup analysis is more transparent, and it's something that we could certainly do.

Dr L____. Did you look at the variability between the hospitals in the final pathologic staging? So perhaps hospital E maybe did not have many lymph nodes, but if they did complete pathologic staging, they may have had more accurate staging. The variability in the surgeons at the different hospitals may reflect how thoroughly they provide their thoracic care.

Dr Farjah. The number of mediastinal lymph node stations, the median number, did not vary across hospitals. It was 3. The pathologic staging was also not different. The distribution of nodal disease was not significantly different. The distribution of pathologic stage was not different. The frequency of clinical stage IA was different at presentation. The pathologic N0, N1, and N2 did not vary across hospitals.

Dr L____. Were all endobronchial-ultrasound guided nodal aspirate procedures captured? Even if it was done by a pulmonologist?

Dr Farjah. Yes it was a hospital-based perspective, so whether it was surgeon or pulmonologist, it was captured.

APPENDIX E1. PARTICIPATING HOSPITALS

MultiCare Allenmore Hospital
MultiCare Good Samaritan Hospital
Harborview Medical Center
Peace Health St Joseph Medical Center
Providence Regional Medical Center Everett
St Joseph Medical Center
St Francis Hospital
St Anthony Hospital
St Clare Hospital
St Elizabeth Hospital
MultiCare Tacoma General Hospital
University of Washington Medical Center
Virginia Mason Medical Center