



Impact of a Lymph Node Specimen Collection Kit on the Distribution and Survival Implications of the Proposed Revised Lung Cancer Residual Disease Classification: A Propensity-Matched Analysis

Matthew P. Smeltzer, PhD,^a Nicholas R. Faris, MDiv,^b Carrie Fehnel, BBA,^b Olawale Akinbobola, MPH,^b Andrea Saulsberry, MBA,^b Meghan Meadows-Taylor, PhD,^b Alicia Pacheco, MHA,^b Meredith Ray, PhD,^a Raymond U. Osarogiagbon, M.B.B.S.^{b,*}

^aDivision of Epidemiology, Biostatistics, and Environmental Health, School of Public Health, University of Memphis, Memphis, Tennessee

^bThoracic Oncology Research Group, Baptist Memorial Health Care Corporation, Memphis, Tennessee

Received 21 February 2021; accepted 21 February 2021
Available online - 9 March 2021

ABSTRACT

Importance: The International Association for the Study of Lung Cancer (IASLC) has proposed a revision of the residual disease (R-factor) classification, to R0, 'R-uncertain', R1 and R2. We previously demonstrated longer survival after surgical resection with a lymph node specimen collection kit, and now evaluate R-factor redistribution as the mechanism of its survival benefit.

Objective: We retrospectively evaluated surgical resections for lung cancer in the population-based observational 'Mid-South Quality of Surgical Resection' cohort from 2009-2019, including a full-cohort and propensity-score matched analysis.

Results: Of 3,505 resections, 34% were R0, 60% R-uncertain, and 6% R1 or R2. The R0 percentage increased from 9% in 2009 to 56% in 2019 ($p < 0.0001$). Kit cases were 66% R0 and 29% R-uncertain, compared to 14% R0 and 79% R-uncertain in non-kit cases ($p < 0.0001$). Compared with non-kit resections, kit resections had 12.3 times the adjusted odds of R0 versus R-uncertainty.

Of 2,100 R-uncertain resections, kit cases had lower percentages of non-examination of lymph nodes, 1% vs. 14% ($p < 0.0001$) and non-examination of mediastinal lymph nodes, 8% vs. 35% ($p < 0.0001$). With the kit, more R-uncertain cases had examination of stations 7 (43% vs. 22%, $p < 0.0001$) and 10 (67% vs. 45%, $p < 0.0001$).

The adjusted hazard ratio (aHR) for kit cases versus non-kit cases was 0.75 (confidence interval [CI]: 0.66–0.85, $p < 0.0001$). In 2,100 subjects with R-uncertain resections, kit cases had an aHR of 0.79 versus non-kit cases ([CI: 0.64–0.99], $p=0.0384$); however, in the 1,199 R0 resections the

survival difference was not significant (aHR: 0.85[0.68–1.07], $p = 0.17$).

Conclusions and Relevance: A lymph node kit increased overall survival by increasing R0, reducing the probability of R-uncertain resections, and diminishing extreme R-uncertainty.

© 2021 The Authors. Published by Elsevier Inc. on behalf of the International Association for the Study of Lung Cancer. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

*Corresponding author.

Disclosure: Dr. Osarogiagbon owns patents for the lymph node specimen collection kit; owns stocks in Eli Lilly, Gilead Sciences, and Pfizer; has worked as a paid research consultant for the American Cancer Society, the Association of Community Cancer Centers, AstraZeneca, Biodesix, Eli Lilly, Triptych Healthcare Partners, and Genentech/Roche; and is founder of Oncobox Device, Inc. Dr. Smeltzer has worked as a paid research consultant for the Association of Community Cancer Centers. The remaining authors declare no conflict of interest.

Address for correspondence: Raymond U. Osarogiagbon, M.B.B.S., Thoracic Oncology Research Group, Baptist Memorial Health Care Corporation, 80 Humphreys Center Drive, Suite 330, Memphis, TN 38120. E-mail: rosarogi@bmhcc.org

Cite this article as: Smeltzer MP, Faris N, Fehnel C, et al. Impact of a lymph node specimen collection kit on the distribution and survival implications of the proposed revised lung cancer residual disease classification: a propensity-matched analysis. *JTO Clin Res Rep.* 2021;2:100161.

© 2021 The Authors. Published by Elsevier Inc. on behalf of the International Association for the Study of Lung Cancer. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

ISSN: 2666-3643

<https://doi.org/10.1016/j.jtocrr.2021.100161>

Keywords: Lymphadenectomy; Pathologic nodal staging; Quality of care; Surgical resection; Survival

Key Points

Question: How does a lymph node collection kit improve survival after lung cancer resection? We evaluated the completeness of resection as the mechanism.

Findings: In this population-based cohort, we found that using the lymph node collection kit increased the probability of achieving more stringently defined complete resection, reduced uncertainty, and skewed the severity of uncertainty on the completeness of resection away from the most extreme end of the spectrum.

Meaning: A lymph node collection kit is an effective mechanism to aid surgeons in attaining complete resection, thereby improving long-term survival after curative-intent lung cancer surgery.

Introduction

Surgical resection is the most important curative-intent treatment modality for NSCLC. With successful implementation of lung cancer screening programs, the proportion of patients with NSCLC who will undergo curative-intent surgery is likely to increase considerably. Poor surgical quality reduces the survival benefit of curative-intent surgery.^{1,2} Suboptimal pathologic nodal evaluation is the most prevalent NSCLC surgery quality deficit.²⁻⁴ This problem is global and prevalent across institutions of different characteristics.⁵⁻⁸

In response to this, the International Association for the Study of Lung Cancer (IASLC) proposed a revision of the “residual disease” (R-factor) classification to provide a more stringent definition of “complete” resection (R0).⁹ The proposed revision includes a new category, “R-uncertain,” in which despite uninvolved resection margins, the risk of incomplete resection remains relatively high because of any combination of the following: positive pleural lavage cytology result, carcinoma in situ at the bronchial resection margin, involvement of the highest mediastinal lymph node, and suboptimal lymph node examination.⁹ The prognostic value of this proposed revised R-factor classification has been independently corroborated.¹⁰⁻¹² Suboptimal nodal examination was the cause of R-uncertainty in 58% to 98% of patients.¹⁰⁻¹²

Lung resection with a lymph node specimen collection kit improves pathologic nodal staging and has been associated with improved survival.^{13,14} We evaluated the

impact of surgery with the kit on the completeness of surgical resection under the proposed IASLC R-factor classification. We hypothesized that by improving the quality of pathologic node evaluation, use of the kit would result to the following: increase the probability of R0 resections and reduce the probability of R-uncertain resections; have limited survival benefits within the R-factor classifications, except for the R-uncertain category in which, by reducing the extremes of poor nodal evaluation, it would be associated with improved survival.

Materials and Methods

The population-based Mid-South Quality of Surgical Resection (MS-QSR) cohort is an observational cohort including detailed clinical and demographic information on greater than 95% of lung cancer resections within the four contiguous Hospital Referral Regions in Eastern Arkansas, Northern and Central Mississippi, and Western Tennessee from 2009 to 2019.^{4,15} This observational cohort includes curative-intent resections from 2009 to 2012 retrospectively abstracted in 2013 and prospective data abstraction from 2013 onward. The cohort is rigorously audited and actively maintained, with continual prospective observational updates, including survival updates for all living subjects every 6 months.^{4,15} Data used for this research have been compiled under an ongoing institutional review board–approved protocol at our institution with an appropriate waiver of consent.

We conducted a pilot study of a lymph node kit within this cohort,¹⁶ and subsequently, prospectively studied its implementation using a staggered implementation design.^{13,14} The kit includes 12 anatomically prelabeled specimen containers and a checklist to indicate certain lymph node stations mandated for examination.^{13,16} We designed it to improve the intraoperative retrieval of lymph nodes compatible with evidence-based guidelines, the secure transfer of lymph node specimens between surgery and pathology teams, and the accurate identification of the anatomical provenance of lymph node specimens to encourage thorough and accurate pathologic evaluation.^{13,16,17}

We reclassified all subjects in the MS-QSR cohort on the basis of the IASLC R-factor classification to evaluate the impact of the kit on the completeness of surgical resection, while controlling for potentially important factors including age, sex, histology, pathologic T-category (pT-category), pathologic M-category (pM-category), and number of comorbidities. We then evaluated the impact of kit use and IASLC R-classification on overall survival. We excluded subjects who had a previous lung cancer resection, preoperative chemotherapy, or preoperative radiation therapy, which could confound the retrieval of lymph nodes.

Statistical Methods

We used multinomial logistic regression (generalized logit model) to evaluate the crude and adjusted association between kit use (exposure variable) and R-classification (outcome variable).¹⁸ Model-based ORs and adjusted ORs are presented with 95% confidence intervals (CIs). We evaluated overall survival from the date of surgery until date of death or censoring using the Kaplan-Meier method and proportional hazards regression. The date of last follow-up for censored observations was August 31, 2019. Crude hazard ratios (HRs) and adjusted HRs (aHRs) are revealed with 95% CI. We evaluated the proportional hazards assumption using log(−log) plots. Multivariable models were adjusted for age, sex, histology, pT-category, pM-category, hospital Commission on Cancer structural category (“non-accredited,” “community,” “academic”), and number of comorbidities.

We created a propensity score-matched cohort, to account for potential differences in subjects for whose surgery the kit was, and was not, used. This one-to-one (kit versus nonkit)-matched propensity analysis was balanced on the following: age at surgery, sex, race, insurance, smoking status, histologic category, extent of resection, pathologic grade, number of comorbidities, type of surgeon (dedicated thoracic: yes or no), use of invasive staging, pT-category, pM-category, and the use of preoperative positron emission tomography-computed tomography scans.¹⁸⁻²¹ This analysis matched on the basis of the Mahalanobis distance with the threshold distance for a match set to 0.1.²¹

Propensity score matching was performed in R Gui using the software package “Matching” that automates balance optimization.²²⁻²⁵ All other analyses were performed in Statistical Analysis System version 9.4 (Cary, NC). The *p* values less than 0.05 were considered statistically significant with no adjustment for multiple comparisons.²⁶

Results

We evaluated a total of 3505 subjects who had surgical resection, with a median age of 68 years (interquartile range: 61–74) (Table 1 and Supplementary Table 1). The cohort included 1600 female patients (46%), with 2744 white (78%) versus 718 black (20%) patients. The subjects most frequently had clinical stage I or II, adenocarcinoma or squamous cell histology, and underwent lobectomy (Table 1). The surgical kit was used in 1356 cases (39%), increasing from 0 (0%) in 2009 to 199 (88%) in 2019 ($p < 0.0001$, Fig. 1A). Patients for whom the kit was used were more frequently of female sex and had Medicare insurance (Table 1). Other significant differences were observed according to

histology, final extent of resection, and pathologic nodal stage (Table 1). On the basis of the IASLC R-factor classification, we observed 1199 R0 (34%), 2100 R-uncertain (60%), and 206 microscopic incomplete resection (R1) or grossly incomplete resection (R2) (6%). The percentage of R0 resections increased from 9% (33 of 349) in 2009 to 56% (127 of 227) in 2019 ($p < 0.0001$, Fig. 1B). The propensity score-matched cohort achieved high-level matches of 1172 kit cases with 1172 nonkit cases on the basis of all matching criteria.

Kit Impact on IASLC R-Classification

Cases in which the kit was used were 66% R0, 29% R-uncertain, and 5% R1/R2 compared with 14% R0, 79% R-uncertain, and 7% R1 or R2 in nonkit cases ($p < 0.0001$, Table 1). Compared with nonkit resections, kit resections had 12.3 (10.4–14.7) times the adjusted odds of R0 versus R-uncertain ($p < 0.0001$, Table 2). Significant differences in R-factor distribution persisted when evaluated by year with and without kit use (Fig. 1C and D and Supplementary Table 2). In the propensity score-matched cohort, the kit had 9.4 (7.7–11.4) times the odds of R0 compared with R-uncertain resection (Table 2).

Quality of R-Uncertain Resections

R-uncertainty includes extremes of poor nodal examination such as nonexamination of lymph nodes (pNX) and nonexamination of mediastinal lymph nodes (pNX^{med}), which have even more adverse survival. We evaluated the “degree of uncertainty” between kit and nonkit cases among the 2100 R-uncertain resections. There were significantly fewer kit cases with pNX, pNX^{med}, missing station seven lymph node, and missing station 10 lymph node (all $p < 0.0001$, Table 3). When these metrics were compared in the 1261 R-uncertain cases from the propensity score-matched cohort, all differences remained statistically significant (all $p < 0.0001$, Table 3).

Overall Survival

We evaluated the associations of kit use and R-factor on overall survival in the full MS-QSR cohort and propensity score-matched cohort. Surgical resection with the kit was associated with a lower hazard of death compared with resection without the kit in unadjusted analyses (HR: 0.73 [0.65–0.83], $p < 0.0001$) and after adjusting for age, sex, histology, pT-category, pM-category, hospital Commission on Cancer structural category, and number of comorbidities (aHR: 0.75 [CI: 0.66–0.85], $p < 0.0001$). In the propensity score-matched cohort, we found a 21% lower hazard of death in resections with kit use compared with those without kit use (propensity

Table 1. Demographic and Disease Characteristics

Characteristics	Kit Used n = 1356 n (%)	No Kit n = 2149 n (%)	Total N = 3505 N (%)	p Value
Sex				
Male	679 (50)	1226 (57)	1905 (54)	<0.0001
Female	677 (50)	923 (43)	1600 (46)	
Race				
White	1087 (80)	1657 (77)	2744 (78)	0.1016
Black or African American	254 (19)	464 (22)	718 (20)	
Asian	7 (1)	10 (0.5)	17 (0.5)	
Other or unknown	8 (1)	18 (1)	26 (1)	
PET/CT				
No	172 (13)	421 (20)	593 (17)	<0.0001
Yes	1184 (87)	1728 (80)	2912 (83)	
Insurance				
Medicare	656 (48)	969 (45)	1625 (46)	0.0003
Medicaid	232 (17)	293 (14)	525 (15)	
Commercial	430 (32)	810 (38)	1240 (35)	
Self-insured or none	38 (3)	77 (4)	115 (3)	
Final extent of resection				
Pneumonectomy	38 (3)	133 (6)	171 (5)	<0.0001
Bilobectomy	55 (4)	106 (5)	161 (5)	
Lobectomy	1182 (87)	1514 (70)	2696 (77)	
Segmentectomy	39 (3)	117 (5)	156 (4)	
Wedge	42 (3)	278 (13)	320 (9)	
Other or unknown	0 (0)	1 (0.05)	1 (0)	
Histology				
Adenocarcinoma	787 (58)	1111 (52)	1898 (54)	<0.0001
Squamous cell carcinoma	428 (32)	739 (34)	1167 (33)	
Large cell carcinoma	1 (0.07)	1 (0.05)	2 (0.06)	
Adenosquamous carcinoma	26 (2)	50 (2)	76 (2)	
Other	114 (8)	248 (12)	362 (10)	
Pathologic N-category				
pNX	1 (0.07)	227 (11)	228 (7)	<0.0001
pN0	1081 (80)	1490 (69)	2571 (73)	
pN1	156 (12)	258 (12)	414 (12)	
pN2	118 (9)	173 (8)	291 (8)	
pN3	0 (0)	1 (0.05)	1 (0.03)	
Pathologic T-category				
T0 or TX or T1	647 (48)	946 (44)	1593 (45)	0.0163
T2	471 (35)	741 (34)	1212 (35)	
T3	168 (12)	306 (14)	474 (14)	
T4	70 (5)	156 (7)	226 (6)	
Pathologic M-category				
M0	1344 (99)	2120 (99)	3464 (99)	0.2129
M1	12 (1)	29 (1)	41 (1)	

M, T, and N are the formal category names of pathologic M-categories per the IASLC staging manual. CT, computed tomography; PET, positron emission tomography; pNx, nonexamination of lymph node.

score-matched HR: 0.79 [CI: 0.68–0.91], $p = 0.0009$). The estimated kit effect on overall survival was consistent when evaluated by year, although these analyses did not have adequate statistical power (Supplementary Table 2).

In 2100 subjects with R-uncertain resections, kit cases had better crude (HR: 0.76 [0.61–0.94], $p = 0.0133$) and adjusted overall survival (aHR: 0.79 [CI: 0.64–0.99], $p = 0.0433$) versus nonkit cases (Fig. 2). However, in the 1199 R0 resections, the survival

difference between the kit and nonkit cases was smaller and did not reach statistical significance (HR: 0.87 [0.69–1.09], $p = 0.21$; aHR: 0.86 [0.68–1.08], $p = 0.22$; Fig. 2). We found no difference in overall survival with kit use on the 206 subjects with incomplete resection in crude models (HR: 0.80 [0.53–1.21]), but we found a difference in adjusted models (0.64 [0.41, 0.995; $p = 0.0475$], Fig. 2). These results were consistent in the propensity score-matched cohort (Fig. 2).

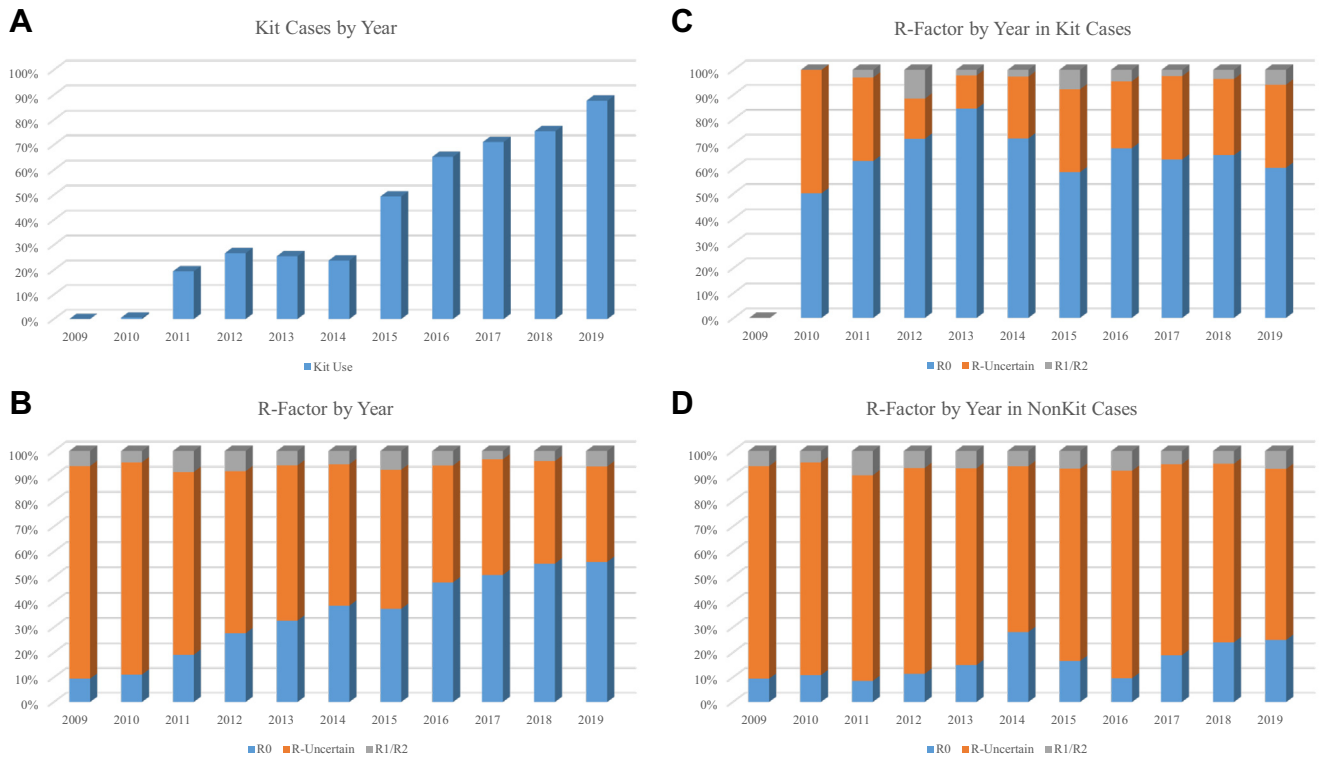


Figure 1. (A) Use of a lymph node collection kit by year from 2009 to 2019 in a population-based cohort. (B) Distribution of IASLC R-factor by year from 2009 to 2019. (C) Distribution of IASLC R-factor by year in cases in which a lymph node collection kit was used. (D) Distribution of IASLC R-factor by year in cases in which a lymph node collection kit was not used. IASLC, International Association for the Study of Lung Cancer; R0, complete resection; R1, microscopic incomplete resection; R2, grossly incomplete resection; R-factor, residual disease.

Discussion

The IASLC has proposed a revised R-factor classification for enactment with the ninth edition of the International Lung Cancer Staging System.^{9,11} The main effect is to more stringently define “complete resection” beyond the absence of R-factor at the resection margin.⁹ The primary innovation was the creation of the “R-uncertain” category, in which

despite absence of invasive cancer at the resection margins, there remains significant risk of R-factor because of the presence of carcinoma in situ at the margin, positive pleural lavage cytology result, metastasis at the highest mediastinal lymph node, or suboptimal pathologic nodal evaluation indicated by failure to perform either a systematic or lobe-specific nodal evaluation.⁹

Table 2. R-Factor Distribution by Use of a Lymph Node Collection Kit

R-Factor	Total N	Kit Used n (%)	No Kit n (%)
Full cohort			
R0	1199	894 (66)	305 (14)
R-uncertain	2100	397 (29)	1703 (79)
R1 or R2	206	65 (5)	141 (7)
R0 vs. R-uncertain, aOR (95% CI)	12.3 (10.4-14.7)		
R0 vs. R1 or R2, aOR (95% CI)	5.5 (4.0-7.7)		
Propensity-matched cohort			
R0	961	760 (65)	201 (17)
R-uncertain	1261	362 (31)	899 (77)
R1 or R2	122	50 (4)	72 (6)
R0 vs. R-uncertain, pmOR (95% CI)	9.39 (7.71-11.44)		
R0 vs. R1 or R2, pmOR (95% CI)	5.44 (3.68-8.06)		

aOR, adjusted OR; CI, confidence interval; pmOR, propensity-matched OR; R0, complete resection; R1, microscopic incomplete resection; R2, grossly incomplete resection; R-factor, residual disease.

Table 3. Quality Metrics in R-Uncertain Resections

Sub-Set of R-Uncertainty	Kit Used n (%)	No Kit n (%)	Total	p Value
R-uncertain resections in full cohort	397	1703	2100	
pNX	3 (1)	225 (13)	228	<0.0001
No mediastinal lymph nodes examined	32 (8)	599 (35)	631	<0.0001
Station 7 lymph nodes not examined	226 (57)	1331 (78)	1557	<0.0001
Station 10 lymph nodes not examined	132 (33)	936 (55)	1068	<0.0001
R-uncertain resections in propensity-matched cohort (n)	362	899	1261	
pNX	3 (1)	63 (7)	66	<0.0001
No mediastinal lymph nodes examined	31 (9)	279 (31)	310	<0.0001
Station 7 lymph nodes not examined	210 (58)	679 (76)	889	<0.0001
Station 10 lymph nodes not examined	123 (34)	439 (49)	562	<0.0001

pNX, nonexamination of lymph node.

Most R-uncertain resections would have previously been categorized as R0. Inadequate lymph node examination is the reason for recategorization from R0 to R-uncertain in the overwhelming majority of cases.¹⁰⁻¹² The survival impact of this revised R-factor classification has been independently validated in a single institutional Italian cohort, a large international data set predominantly from Japan, and our population-based

multi-institutional U.S. cohort.¹⁰⁻¹² Adoption of the proposed revised R-factor classification will heighten global awareness of the survival impact of the pathologic nodal staging quality gap, potentially stimulating worldwide efforts at quality improvement.²⁷

In pilot studies, we previously described the effectiveness of a lymph node kit in improving the quality of pathologic nodal staging and overcoming the sharp discordance between surgeons and pathologists in identifying the lymph node evaluation procedure performed.^{14,16,17,28} We subsequently revealed strong early evidence of a survival impact in a staggered implementation study.¹³ In another report, we quantified the effect size of using the kit versus not using it in pre- and postkit implementation cohorts, on multiple quality benchmarks in an effort to estimate sample size and statistical power for a proposed institutional cluster-randomized comparative effectiveness clinical trial.¹⁴ We conducted that analysis with the assumption that postimplementation resections without the kit were done in a state of “heightened awareness” of the pathologic nodal staging quality gap, its adverse effects, and the possibility of overcoming it.¹⁴ The kit was designed with the idea that optimal pathologic nodal staging is a team-based activity, involving surgical retrieval of lymph nodes, correct labeling, secure transfer of specimens between the surgery and pathology teams, and thorough examination and accurate reporting by the pathologist. In the “chain of responsibility” conceptual model, a breakdown in performance by team members at any of these steps weakens the overall performance of the group.^{13,14}

In this report, we evaluated the potential mechanism of the survival benefit of the kit, by testing its impact on the distribution of completeness of resection under the more stringent IASLC R-factor definition. We hypothesized that use of the lymph node collection kit would promote achievement of the more stringently defined R0 resection while diminishing the frequency and severity of R-uncertainty. We further hypothesized that survival

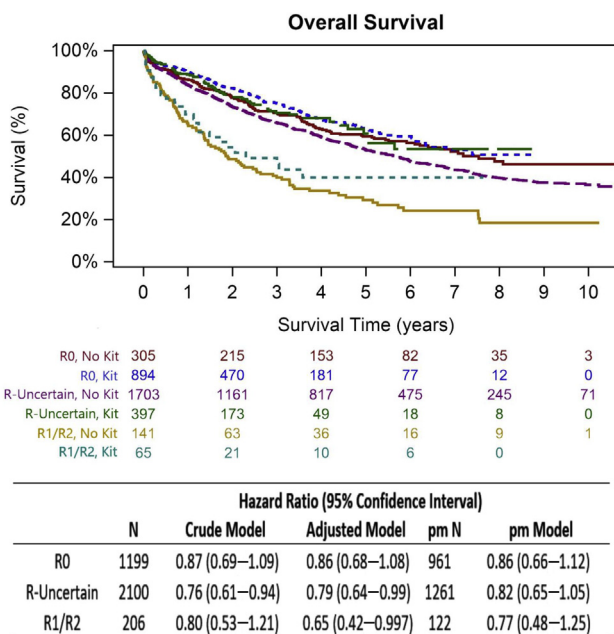


Figure 2. Overall survival based on use of a lymph node collection kit use within IASLC R-factor categories evaluated by Kaplan-Meier analysis and proportional hazards models. Adjusted models control for age, sex, histology, pT-category, pM-category, and number of comorbidities. pm analyses control for age, sex, race, insurance, smoking status, histologic category, extent of resection, pathologic grade, number or comorbidities, type of surgeon, use of invasive staging, pT-category, pM-category, and the use of preoperative PET-CT. PET-CT, positron emission tomography-computed tomography; pm, propensity matched; pM-category; pathologic M-category; pT-category, pathologic T-category; R-factor, residual disease.

of patients within the R0 and R1/R2 categories would be equivalent between kit and nonkit cases, but there would be some improvement in survival within the R-uncertain kit cases because of a shift away from the most dire end of the uncertainty spectrum induced by pNX and pNX^{med}.

Both hypotheses are supported by analysis of this expanded population-based cohort. Kit and nonkit R0 resections had no significant survival differences; kit R-uncertain resections had more thorough lymph node evaluation and slightly better overall survival than nonkit resections. This suggests that the kit both increases the R0 resection rate and mitigates the adverse impact of suboptimal nodal evaluation by reducing the extremes of uncertainty. We speculate that extremes of R-uncertainty increase the risk that occult, oligometastatic lymph node disease might have been left behind. Previous analyses of the U.S. Surveillance, Epidemiology, and End Results database revealed that the extremes of poor lymph node evaluation, pNX, and pNX^{med} are associated with significantly worse survival.^{2,3} Unexpectedly, we also found some indication of better survival with kit use in the 206 R1 or R2 cases, although this was not consistent across analyses.

The “oligo-metastatic lymph node” hypothesis may seem to have been refuted by the failure of the American College of Surgeons Oncology Group Z0030 trial to reveal a survival difference between recipients of a stringently defined systematic sampling or mediastinal nodal dissection procedure, despite the finding of unexpected mediastinal nodal metastasis in 4% of the more extensive dissection arm.²⁹ However, technical limitations of that trial have left the debate open.^{15,30} More importantly, resections that fall significantly below the systematic sampling quality threshold have an adverse association with survival and the effect of the kit may be in limiting the severity of the deviation from the minimum protective threshold, below which R-factor becomes much more likely. Translational studies, such as those measuring circulating tumor DNA as a surrogate for minimal R-factor, might provide a means of testing this hypothesis in the future.³¹ Furthermore, improved lymph node evaluation, by increasing the likelihood of detecting nodal metastasis, also increases the likelihood of correctly identifying candidates for beneficial adjuvant treatment.⁴ Therefore, increased appropriate use of adjuvant therapy may account for some of the survival improvements, but those benefits are modest (only 4%–5%).^{32,33}

Potential limitations of this study include the nonrandom assignment of lymph node kit use, which may lead to confounding-by-indication or other imbalances in the data. To address this, we evaluated multiple-variable statistical models and constructed a propensity

score-matched cohort to balance the potential confounding factors between cases in which the lymph node collection kit was and was not used. The propensity score-matched analysis was consistent with the full-cohort analyses. In addition, although we hypothesize that R-factor mediates the effect of the lymph node kit on overall survival, methodologic techniques are not currently available to estimate the mediated effect in a causal mediation analysis.³⁴ Although the MS-QSR is population based, including greater than 95% of surgical resections in a defined geographic region, the magnitude of the lymph node kit impact may not be generalizable to all situations. Specifically, lung cancer surgery quality in the MS-QSR may be significantly worse than in other regions. Pathologic nodal staging quality and long-term postoperative survival vary significantly between categories of institutions, with academic programs faring much better than community programs.⁸ However, we have previously revealed that many of the surgical quality deficits observed in the MS-QSR are consistent with the U.S. national and international experiences.^{2,3,35,36} In any case, most lung cancer care in the United States is performed in community-based programs.³⁷ Ultimately, the main limitation of this study is that it was not a randomized controlled trial. We propose to conduct such a trial to further evaluate the lymph node kit.

In this population-based cohort, we found that a lymph node collection kit used at surgery increased the probability of achieving a stringently defined complete resection, minimized uncertainty on the completeness of resection, and skewed R-uncertain resections away from the worst end of the spectrum of uncertainty. A lymph node kit seems to provide lung cancer surgery teams with a mechanism to achieve more complete resections, and by that means, may improve the long-term survival of the patients with lung cancer.

Acknowledgments

This work was supported by the National Cancer Institute (R01 CA172253). The authors thank all past and present members of the Thoracic Oncology Research Group for their contributions. Dr. Osarogiagbon had full access to all study data, vouches for the accuracy of the report, and bears full responsibility for the decision to submit for publication.

Supplementary Data

Note: To access the supplementary material accompanying this article, visit the online version of the *JTO Clinical and Research Reports* at www.jtocrr.org and at <https://doi.org/10.1016/j.jtocrr.2021.100161>.

References

- Osarogiagbon RU, Lin CC, Smeltzer MP, Jemal A. Prevalence, prognostic implications, and survival modulators of incompletely resected non-small cell lung cancer in the U.S. National Cancer Data Base. *J Thorac Oncol.* 2016;11:e5-e16.
- Osarogiagbon RU, Yu X. Mediastinal lymph node examination and survival in resected early-stage non-small-cell lung cancer in the Surveillance, Epidemiology, and End Results database. *J Thorac Oncol.* 2012;7:1798-1806.
- Osarogiagbon RU, Yu X. Nonexamination of lymph nodes and survival after resection of non-small cell lung cancer. *Ann Thorac Surg.* 2013;96:1178-1189.
- 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.
- Verhagen AF, Schoenmakers MC, Barendregt W, et al. Completeness of lung cancer surgery: is mediastinal dissection common practice? *Eur J Cardiothorac Surg.* 2012;41:834-838.
- 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.
- Liang W, He J, Shen Y, et al. Impact of examined lymph node count on precise staging and long-term survival of resected non-small-cell lung cancer: a population study of the US SEER database and a Chinese multi-institutional registry. *J Clin Oncol.* 2017;35:1162-1170.
- Osarogiagbon RU, Sineshaw HM, Lin CC, Jemal A. Institution-level differences in quality and outcomes of lung cancer resections in the United States [e-pub ahead of print]. *Chest.* 10.1016/j.chest.2020.10.075, accessed February 19, 2021.
- Rami-Porta R, Wittekind C, Goldstraw P, International Association for the Study of Lung Cancer (IASLC) Staging Committee. Complete resection in lung cancer surgery: proposed definition. *Lung Cancer.* 2005;49:25-33.
- Gagliasso M, Migliaretti G, Ardisson F. Assessing the prognostic impact of the International Association for the Study of Lung Cancer proposed definitions of complete, uncertain, and incomplete resection in non-small cell lung cancer surgery. *Lung Cancer.* 2017;111:124-130.
- Edwards JG, Chansky K, Van Schil P, et al. The IASLC Lung Cancer Staging Project: analysis of resection margin status and proposals for residual tumor descriptors for non-small cell lung cancer. *J Thorac Oncol.* 2020;15:344-359.
- Osarogiagbon RU, Faris NR, Stevens W, et al. Beyond margin status: population-based validation of the proposed International Association for the Study of Lung Cancer residual tumor classification recategorization. *J Thorac Oncol.* 2020;15:371-382.
- Osarogiagbon RU, Smeltzer MP, Faris NR, et al. Outcomes following use of a lymph node collection kit for lung cancer surgery: a pragmatic, population-based, multi-institutional, staggered implementation study [e-pub ahead of print]. *J Thorac Oncol.* <https://doi.org/10.1016/j.jtho.2020.12.025>, accessed March 27, 2021.
- Ray MA, Fehnel C, Akinbobola O, et al. Comparative effectiveness of a lymph node collection kit versus 'heightened awareness' on lung cancer surgery quality and outcomes [e-pub ahead of print]. *J Thorac Oncol.* <https://doi.org/10.1016/j.jtho.2021.01.1618>, accessed March 27, 2021.
- Ray MA, Smeltzer MP, Faris NR, Osarogiagbon RU. Survival after mediastinal node dissection, systematic sampling, or neither for early stage NSCLC. *J Thorac Oncol.* 2020;15:1670-1681.
- Osarogiagbon RU, Miller LE, Ramirez RA, et al. Use of a surgical specimen-collection kit to improve mediastinal lymph-node examination of resectable lung cancer. *J Thorac Oncol.* 2012;7:1276-1282.
- Osarogiagbon RU, Sareen S, Eke R, et al. Audit of lymphadenectomy in lung cancer resections using a specimen collection kit and checklist. *Ann Thorac Surg.* 2015;99:421-427.
- Agresti A. *Categorical data analysis.* 3rd ed. Hoboken, NJ: John Wiley & Sons; 2013.
- Cochran WG, Rubin DB. Controlling bias in observational studies: a review. *Sankhya A.* 1973;35:417-446.
- Rubin DB. Using multivariate sampling and regression adjustment to control bias in observational studies. *J Am Stat Assoc.* 1979;74:318-328.
- Rubin DB. Bias reduction using Mahalanobis metric matching. *Biometrics.* 1980;36:293-298.
- R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.R-project.org/>. Accessed March 27, 2021.
- Sekhon JS. Multivariate and propensity score matching software with automated balance optimization. *J Stat Softw.* 2011;42:1-52.
- Diamond A, Sekhon JS. Genetic matching for estimating causal effects: a general multivariate matching method for achieving balance in observational studies. *Rev Econ Stat.* 2013;95:932-945.
- Sekhon JS, Grieve RD. A matching method for improving covariate balance in cost-effectiveness analyses. *Health Econ.* 2012;21:695-714.
- Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology.* 1990;1:43-46.
- Rami-Porta R, Wittekind C, Goldstraw P. Complete resection in lung cancer surgery: from definition to validation and beyond. *J Thorac Oncol.* 2020;15:1815-1818.
- Osarogiagbon RU, Allen JW, Farooq A, Wu JT. Objective review of mediastinal lymph node examination in a lung cancer resection cohort. *J Thorac Oncol.* 2012;7:390-396.
- 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.

30. Mokhles S, Macbeth F, Treasure T, et al. Systematic lymphadenectomy versus sampling of ipsilateral mediastinal lymph-nodes during lobectomy for non-small-cell lung cancer: a systematic review of randomized trials and a meta-analysis. *Eur J Cardiothorac Surg*. 2017;51:1149-1156.
31. Chaudhury AA, Chabon JJ, Lovejoy AF, et al. Early detection of molecular residual disease in localized lung cancer by circulating tumor DNA profiling. *Cancer Discov*. 2017;7:1394-1403.
32. Kris MG, Gaspar LE, Chaft JE, et al. Adjuvant systemic therapy and adjuvant radiation therapy for stage I to IIIA completely resected non-small-cell lung cancers. *J Clin Oncol*. 2017;35:2960-2974.
33. Smeltzer MP, Lin CC, Kong FM, Jemal A, Osarogiagbon RU. Survival impact of post-operative therapy modalities according to margin status in non-small cell lung cancer patients in the United States. *J Thorac Cardiovasc Surg*. 2017;154:661-672.e10.
34. Lange T, Hansen JV. Direct and indirect effects in a survival context. *Epidemiology*. 2011;22:575-581.
35. Osarogiagbon RU, Allen JW, Farooq A, Berry A, Spencer D, O'Brien T. Outcome of surgical resection for pathologic N0 and Nx non-small cell lung cancer. *J Thorac Oncol*. 2010;5:191-196.
36. Allen JW, Farooq A, O'Brien TF, Osarogiagbon RU. Quality of surgical resection for nonsmall cell lung cancer in a US metropolitan area. *Cancer*. 2011;117:134-142.
37. Little AG, Rusch VW, Bonner JA, et al. Patterns of surgical care of lung cancer patients. *Ann Thorac Surg*. 2005;80:2051-2056.