Visualization of patterns of lymph node metastases in non–small cell lung cancer using network analysis

Check for updates

Yukihiro Yoshida, MD,[a] Nozomu Saeki, Mphil,[b] Masaya Yotsukura, MD,[a] Kazuo Nakagawa, MD,[a] Hirokazu Watanabe, MD,[c] Yasushi Yatabe, MD,[d] and Shun-ichi Watanabe, MD[a]

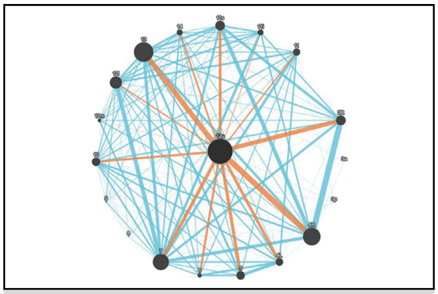
ABSTRACT

Objective: We aimed to visualize complicated patterns of lymph node metastases in surgically resected non–small cell lung cancer by applying a data mining technique.

Methods: In this retrospective study, 783 patients underwent lobectomy or pneumonectomy with systematic mediastinal lymph node dissection for non–small cell lung cancer between January 2010 and December 2018. Surgically resected lymph nodes were classified according to the International Association for the Study of Lung Cancer lymph node map. Network analysis generated patterns of lymph node metastases from stations 1 to 14, and the degree of connection between 2 lymph node stations was assessed.

Results: The median number of lymph nodes examined per patient was 20, and the pathological N category was pN0 in 428 cases, pN1 in 132, pN2 in 221, and pN3 in 2. N1 lymph node stations had strong associations with superior mediastinal lymph node stations for patients with primary tumors in the upper lobes and with station 7 for the lower lobes. There was also a connection from the N1 lymph node stations to superior mediastinal lymph node stations in the lower lobes. In the right middle lobe, an even distribution from station 12m toward stations 2R, 4R, and 7 was noted. We released an interactive web application to visualize these data: http://www.canexapp.com.

Conclusions: Lymph node metastasis patterns differed according to the lobe bearing the tumor. Our results support the need for clinical trials to further investigate selective mediastinal lymph node dissection. (JTCVS Open 2022;12:410-25)



Visualizing lymph node metastasis patterns in primary lung cancer using network analysis.

CENTRAL MESSAGE

Network analysis can help surgeons understand patterns of lymph node metastases, guiding them toward an optimal approach for lymph node dissection for primary lung cancer.

PERSPECTIVE

Current guidelines recommend systematic mediastinal lymph node dissection for patients who undergo surgery for primary lung cancer, regardless of which lobe bears the tumor. However, our network analysis revealed different patterns of lymph node metastases according to the location of the primary tumor, supporting a need for clinical trials to investigate selective mediastinal lymph node dissection.

Video clip is available online.

From the Departments of [a]Thoracic Surgery, [c]Diagnostic Radiology, and [d]Diagnostic Pathology, National Cancer Center Hospital, Tokyo, Japan; and [b]Graduate School of Environmental and Life Science, Okayama University, Okayama, Japan.
Received for publication June 19, 2022; revisions received Sept 21, 2022; accepted for publication Oct 6, 2022; available ahead of print Nov 5, 2022.
Address for reprints: Yukihiro Yoshida, MD, Department of Thoracic Surgery, National Cancer Center Hospital, 5-1-1 Tsukiji, Chuo-ku, Tokyo, 104-0045 Japan (E-mail: yukiyosh@ncc.go.jp).
2666-2736

https://doi.org/10.1016/j.xjon.2022.10.003

Lung cancer often metastasizes to regional lymph nodes such as to the parenchymal, hilar, or mediastinal lymph nodes. Accordingly, it has been considered important to resect not only the lobe bearing the tumor but also the regional lymph nodes, which is the concept of "radical lobectomy" proposed by Cahan in 1960.[1] Lymph node dissection contributes to: 1) accurate staging through histological confirmation, and 2)

Abbreviations and Acronyms

CT	= computed tomography
EBUS	= endobronchial ultrasound
IASLC	= International Association for the Study of Lung Cancer
NSCLC	= non–small cell lung cancer
PET	= positron emission tomography

improved local control and, consequently, improved prognosis. Several randomized studies on lymph node dissection and sampling in lung cancer have been published previously, but results were inconsistent; some studies[2] reported improved prognosis and others[3-5] did not see any effect on prognosis by removing lymph nodes alongside the primary tumor. In the American College of Surgeons Oncology Group (ACOSOG)-Z0030 study, approximately 5% more "unsuspected pN2" patients were detected in the systematic mediastinal lymph node dissection group than in the sampling group.[5] Indeed, accurate staging is considered one of the main reasons for performing lymph node dissection.

The International Association for the Study of Lung Cancer (IASLC) lymph node map defines 14 lymph node stations (stations 1-14) according to their anatomical locations.[6] The lymph nodes excised in surgery are categorized according to this map and evaluated for the presence or absence of metastasis. For thoracic surgeons, the anatomical lymph node map is easy to understand and offers critical information on which lymph node station on the map has metastasis, guiding decisions on surgical procedures. Detailed examination of lymph node metastasis patterns of the primary lung cancer according to site might elucidate complex patterns that can facilitate surgeons in optimizing lymph node dissection. However, no tool has yet been reported to clearly visualize the complex patterns of lymph node metastasis.

Network/graph theory is the study of relationships between discrete objects. Using this theory, factors comprising an entire system temporally and spatially can be reconstructed in silico to visualize how they interact with each other to form a complex system. This theory has been used to analyze communication or social networks as well as distant metastases of cancer.[7] We hypothesized that a detailed analysis of lymph node metastasis patterns might be possible by applying this theory in primary lung cancer. In this retrospective study we aimed to visualize the complex patterns of lymph node metastasis in patients who underwent surgery for primary lung cancer using a data mining technique.

METHODS

Patients

Patients were selected for this study using a database from the Department of Thoracic Surgery at the National Cancer Center Hospital. Of the 4133 patients with primary lung cancer who underwent surgery at our hospital between January 2010 and December 2018, those who had: 1) lobectomy or pneumonectomy, and 2) systematic mediastinal lymph node dissection were included. Patients were excluded if they had: 1) preoperative treatment, 2) bilateral simultaneous surgery, 3) small cell cancer, or 4) multilobe involvement or synchronous multiple lung cancer. Of the 2068 patients who underwent mediastinal lymph node dissection, 1285 underwent selective mediastinal lymph node dissection. Thus, the remaining 783 patients who underwent systematic mediastinal lymph node dissection were included in the analysis. A flow chart of the patient selection is shown in Figure 1.

Background patient clinical information included age, sex, smoking status, diameter of the entire lesion on a preoperative computed tomography (CT) scan, including ground-glass components, and the clinical N category. Furthermore, pathological findings included the histologic type, lesion site, pathological N category, and presence or absence of metastasis at lymph node stations 1 to 14. A patient who has never smoked was defined as a never-smoker and one who ceased smoking at least 1 year before surgery was considered a former smoker. The standard preoperative workup for staging included contrast-enhanced CT of the chest and upper abdomen and whole-body positron emission tomography (PET)/CT scan. Brain CT imaging or magnetic resonance imaging was also used if necessary. Endobronchial ultrasound-guided transbronchial needle aspiration or mediastinoscopy was recommended if there was a mediastinal lymph node >1 cm in diameter along the short axis on CT imaging or if there was increased uptake on PET/CT imaging. See Appendix 1 for more information.[8] Treatment strategies for each patient were discussed by the multidisciplinary team, and chemoradiation was recommended to patients with cN2 non–small cell lung cancer (NSCLC) as the standard of care. In some highly selected patients with cN2 NSCLC (n = 7) from the early period of our study (more than a decade ago), surgery was offered followed by adjuvant chemotherapy on the basis of the satisfactory surgical outcomes observed for cN2 NSCLC in the Japanese lung cancer registry study[9] and the European Society for Medical Oncology clinical practice guidelines.[10] See Supplementary Material for additional reasons that certain patients received surgery. The IASLC lymph node map was used to define the lymph node stations.[6] Each station was classified as either "positive" or "negative" according to the presence or absence of metastatic lymph nodes. The station was considered to be "negative" if there was no lymph node at the station during surgery. Briefly, station 11 is divided into 11s and 11i on the right. Station 11s includes the interlobar nodes between the upper lobe bronchus and bronchus intermedius. Station 11i includes the interlobar nodes between the middle and lower lobe bronchi. Station 12 is divided into 12u, 12m, and 12l. Station 12u includes the upper lobar lymph node, station 12m includes the middle lobar lymph node, and station 12l includes the lower lobar lymph node. According to the eighth edition of the TNM classification,[11] lymph node metastasis was further classified into N1a (N1 at a single station), N1b (N1 at multiple stations), N2a1 (N2 at a single station without N1 involvement), N2a2 (N2 at a single station with N1 involvement), and N2b (N2 at multiple stations).[12] The third[13] or fourth[14] editions of the World Health Organization classification or the 2011 IASLC/American Thoracic Society/European Respiratory Society classification[15] were used for pathological diagnosis depending on the time of diagnosis.

Systematic mediastinal lymph node dissection was defined according to the classification of lung cancer published by the Japan Lung Cancer Society.[16] Specifically, lymph node dissection was performed at stations 2R, 4R, and 7 for the right upper and middle lobes; stations 2R, 4R, 7, 8, and 9 for the right lower lobe; stations 4L, 5, 6, and 7 for the left upper lobe, and stations 4L, 5, 6, 7, 8, and 9 for the left lower lobe.

This study was conducted in compliance with the Declaration of Helsinki and the Ethical Guidelines for Medical and Health Research Involving Human Subjects, and approved by the Institutional Ethics Committee at National Cancer Center (approved on June 11, 2020; approval number: 2020-057). This was a retrospective study with no invasive

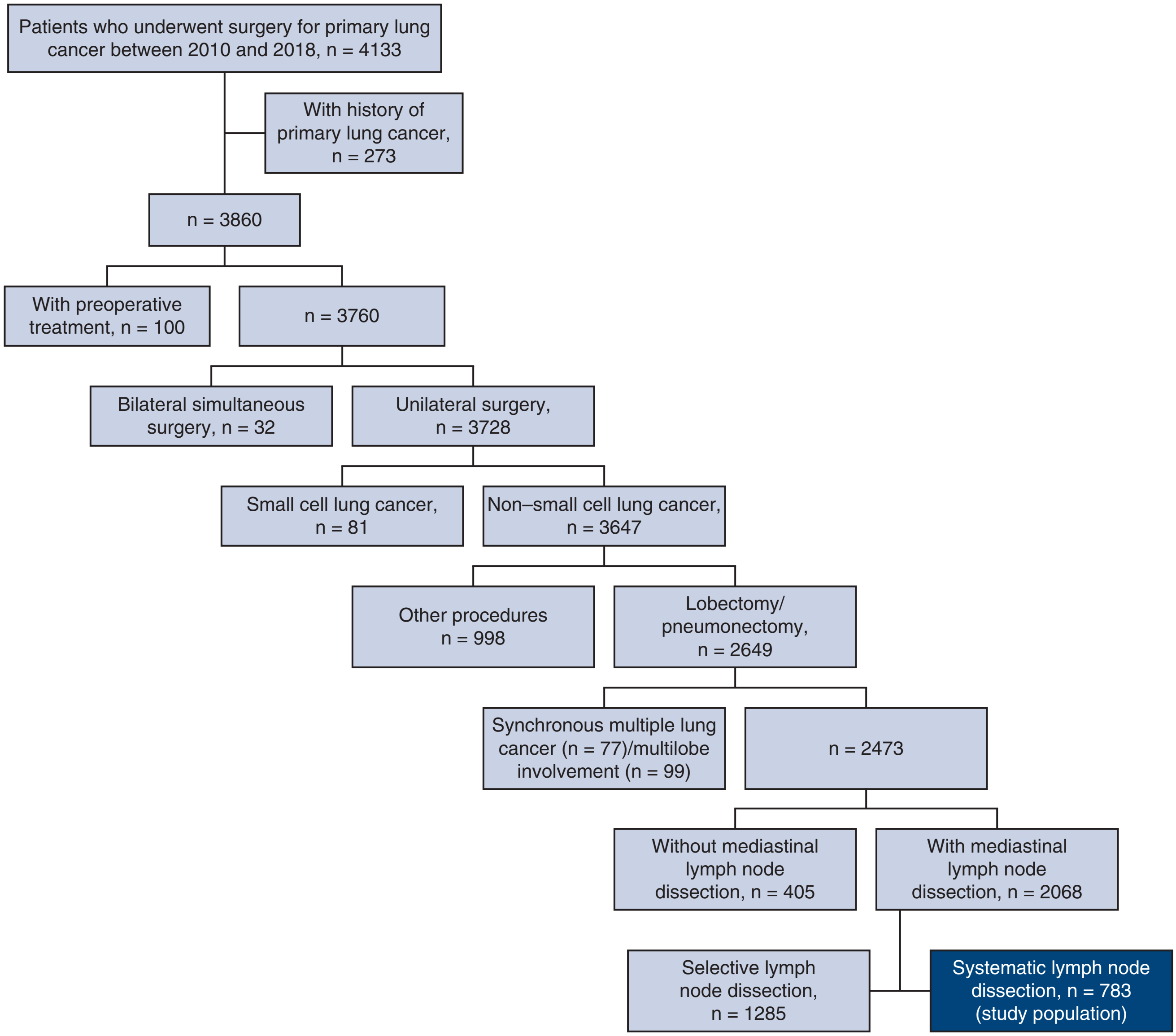


FIGURE 1. Flow chart of patient selection.

procedures or interventions, and thus exempt from requiring written consent from each patient.

Statistical Analysis

Continuous variables are expressed as means or medians with ranges. Categorical variables are expressed as numbers of patients and incidence.

Relationships Between Lymph Node Stations Using Graph Theory

The anonymized patient information and the pathological findings on lymph node metastasis at each lymph node station were described as a matrix B, represented as rows and columns, respectively (Figure E1). In matrix B, a lymph node station with metastasis was set to 1 and a lymph node station without metastasis was set to 0. Matrix B was represented as a bipartite graph, with the patient and each lymph node station as nodes. The transpose of matrix B is denoted by B^T. The relationship of each lymph node station with metastasis was represented as matrix P by projecting the pathological findings of lymph node metastasis at each lymph node station from matrix B ($P = B^T B$). The diagonal component of matrix P represented the number of patients who had a metastatic lymph node at each lymph node station, and the nondiagonal component represented the number of patients who had metastatic lymph nodes concurrently at 2 stations.

Interactive Visualization of the Network Between Lymph Node Stations

Cytoscape (version 3.14.2; Institute for Systems Biology) was used to visualize the lymph node metastases at each lymph node station.[17] We also used Python (version 3.6.9), NumPy (version 1.18.2) package, and Pandas (version 1.0.3) software library for data processing and matrix computation. MySQL (version 14.14; Oracle Corp), Django (version 3.0.4) web framework, Django-MySQL (version 3.5.0) database management system, and mysqlclient (version 1.4.6) database client application were used to visualize the data in Cytoscape.

RESULTS

A total of 783 patients were included in the analysis, and patient background information is shown in Table 1. Forty-three patients received endobronchial ultrasound (EBUS) or mediastinoscopy, and 680 patients (86.8%) received a PET scan preoperatively. Lymph node metastasis was found in 355 patients (45.3%). Details of pathological N category according to lobe are shown in Table 2. Lymph node metastasis was present in 110/255 (43.1%) patients with a tumor in the right upper lobe, 25/115 (21.7%) in the right middle lobe, 65/122 (53.3%) in the right lower lobe, 96/188 (51.1%) in the left upper lobe, and 59/103 (57.3%) in the left lower lobe. There were no statistical differences in age, sex, smoking status, maximal standardized uptake value of the primary tumor on PET/CT imaging, clinical T category, clinical N category, or clinical stage between lobes, except for the right middle lobe (data not shown). Two patients were diagnosed as N3: 1 had adenocarcinoma in the right lower lobe with metastases at lymph node stations 2R, 4R, 7, 11s, 12u, and 4L, and another had adenocarcinoma in the right middle lobe with metastases at stations 2R, 4R, 7, 12m, and contralateral hilar station 10 in the hilar zone. See Supplementary Material for more information on these patients. The median of the number of lymph node metastases were 1 (range, 1-11) in 84 patients with N1a, 3 (range, 2-9) in 48 with N1b, 1 (range, 1-9) in 33 with N2a1, 3 (range, 1-13) in 93 with N2a2, and 8 (range, 2-52) in 95 with N2b. Pathological N category according to histologic subtype and clinical stage is shown in Tables E1 and E2, respectively.

Visualization of Lymph Node Metastasis Patterns Using Network Analysis

Results of the network analysis used to visualize lymph node metastasis patterns on the basis of the lobe bearing the tumor are shown in Figure 2. Tables E3-E7 show the number of patients, according to lobe, who had metastases concurrently at 2 lymph node stations, and Table 3 shows the top 5 combinations of 2 metastatic lymph node stations determined using the largest number of concurrent cases. In summary, the patterns of lymph node metastasis shown in Figure 2 and Table 3 demonstrated the following:

1. N1 lymph node stations 11s and 12u had strong associations with superior mediastinal lymph node stations 2R and 4R in the right upper lobe (number and percentage, respectively, of patients with concurrent metastases were 21% and 8.2% for stations 11s and 4R; 27% and 10.6% for stations 12u and 2R; and 36% and 14.1% for stations 12u and 4R), but they had poor connections

TABLE 1. Clinicopathological characteristics of patients who underwent systematic mediastinal lymph node dissection (N = 783)

Variable		
Age, y		
Median	66	
Range	26-84	
Sex, n		
Male	446	57.0%
Female	337	43.0%
Smoking status, n		
Current	208	26.6%
Former	321	41.0%
Never	254	32.4%
Entire lesion diameter on CT, cm		
Median	3.1	
Range	0.8-11.0	
Clinical N category, n		
cN0	611	78.0%
cN1	138	17.6%
cN2	34	4.3%
Clinical stage, n		
IA1	19	2.4%
IA2	157	20.1%
IA3	189	24.1%
IB	120	15.3%
IIA	56	7.2%
IIB	144	18.4%
IIIA	86	11.0%
IIIB	9	1.1%
IVA	3	0.4%
Lesion site, n		
RUL	255	32.6%
RML	115	14.7%
RLL	122	15.6%
LUL	188	24.0%
LLL	103	13.2%
Surgical procedure, n		
Pneumonectomy	28	3.6%
Lobectomy	755	96.4%
Lymph nodes dissected per patient, n		
Median	20	
Range	4-60	
Histology, n		
Adenocarcinoma	566	72.3%
Squamous cell carcinoma	131	16.7%
Other	86	11.0%
Pathological N category, n		
pN0	428	54.7%
pN1a	84	10.7%
pN1b	48	6.1%
pN2a1	33	4.2%
pN2a2	93	11.9%
pN2b	95	12.1%
pN3	2	0.3%

CT, Computed tomography; *RUL*, right upper lobe; *RML*, right middle lobe; *RLL*, right lower lobe; *LUL*, left upper lobe; *LLL*, left lower lobe.

TABLE 2. Pathological N category according to lobe of the primary tumor

	n	Pathological N category						
		N0	N1a	N1b	N2a1	N2a2	N2b	N3
RUL	255	145	21	13	14	28	34	0
RML	115	90	4	3	6	5	6	1
RLL	122	57	16	11	3	21	13	1
LUL	188	92	24	10	7	26	29	0
LLL	103	44	19	11	3	13	13	0

Data are shown as n. *RUL*, Right upper lobe; *RML*, right middle lobe; *RLL*, right lower lobe; *LUL*, left upper lobe; *LLL*, left lower lobe.

with the inferior mediastinal lymph nodes including subcarinal lymph nodes (ie, only 6 patients [2.4%; 6/255] had concurrent metastases at stations 12u and 7).

2. In the right middle lobe, an even distribution of metastases was noted around the N1 lymph node station 12m located in the peripheral zone, with the superior mediastinal lymph node stations 2R and 4R and the inferior mediastinal lymph node station 7 (n = 4, 5, and 8, respectively) showing metastases.
3. In the right lower lobe, N1 lymph node stations 11i, 11s, and 12l had strong associations with the subcarinal lymph node station 7 (n = 18, 14, and 16), and there

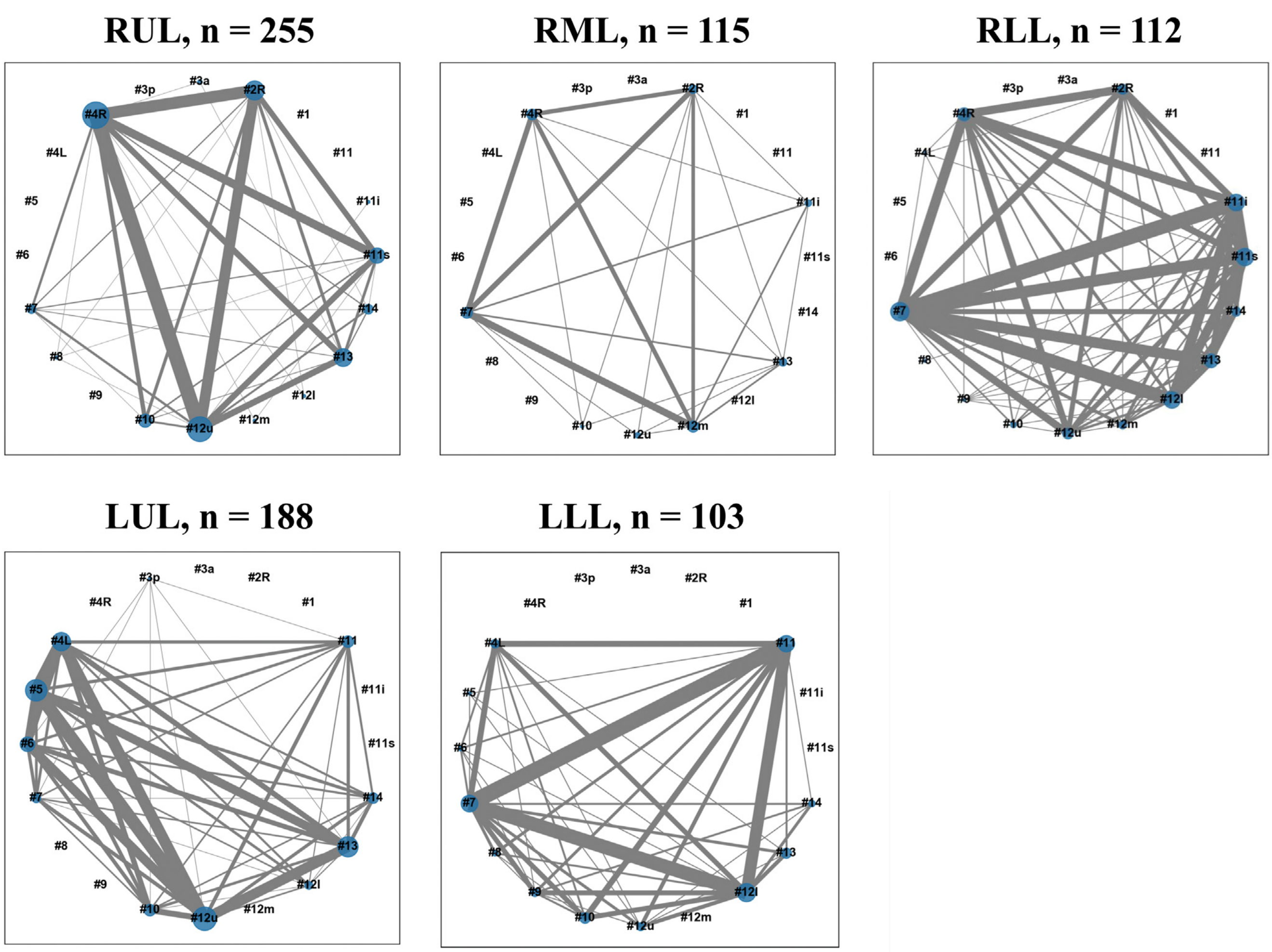


FIGURE 2. Network analysis-based visualization of lymph node metastasis patterns according to lobe of the primary tumor. Lymph node stations from stations 1 to 14 are arranged in a circumferential direction. Metastatic lymph node stations are indicated as *blue dots*. A *larger blue dot* represents a larger number of positive cases. Two stations are connected with a *line* if they share the same patients. A *thicker line* represents a larger number of patients shared. *RUL*, Right upper lobe; *RML*, right middle lobe; *RLL*, right lower lobe; *LUL*, left upper lobe; *LLL*, left lower lobe.

TABLE 3. The number of patients with concurrent metastases at 2 stations according to lobe of the primary tumor

RUL (n = 255)		RML (n = 115)		RLL (n = 122)		LUL (n = 188)		LLL (n = 103)	
1.	4R and 12u (n = 36)	1.	7 and 12m (n = 8)	1.	7 and 11i (n = 18)	1.	5 and 12u (n = 24)	1.	7 and 11 (n = 16)
2.	2R and 4R (n = 30)	2.	2R and 7 (n = 6)	2.	7 and 12l (n = 16)	2.	4L and 5 (n = 23)	2.	7 and 12l (n = 14)
3.	2R and 12u (n = 27)		4R and 7 (n = 6)	3.	7 and 13 (n = 13)	3.	4L and 12u (n = 21)	3.	11 and 12l (n = 12)
4.	4R and 11s (n = 21)	4.	2R and 4R (n = 5)	4.	4R and 7 (n = 12)	4.	12u and 13 (n = 19)	4.	10 and 11 (n = 6)
5.	12u and 13 (n = 16)	5.	4R and 12m (n = 5)	5.	11s and 11i (n = 12)	5.	6 and 12u (n = 17)		4L and 7 (n = 6)
					11i and 12l (n = 12)				4L and 11 (n = 6)
									10 and 11 (n = 6)

Top 5 lymph node station combinations ranked according to the largest number of patients are presented. *RUL*, Right upper lobe; *RML*, right middle lobe; *RLL*, right lower lobe; *LUL*, left upper lobe; *LLL*, left lower lobe.

was also a connection from stations 11s and 11i to station 4R (n = 8 or 10). A connection from station 12u to station 4R was also noted (n = 8).

4. In the left upper lobe, even connection was noted from the N1 lymph node station 12u to the superior mediastinal lymph node stations 4L, 5, and 6 (n = 21, 24, and 17). As with the right upper lobe, there were fewer connections to the inferior mediastinal lymph nodes including to the subcarinal lymph nodes (ie, only 7 patients had concurrent metastases at stations 12u and 7).
5. In the left lower lobe, connections were noted from stations 11 and 12l to station 7 (n = 16 and 14), whereas there were also connections from stations 11 and 12l to station 4L (n = 6 and 5). As with 4L, association with N1 lymph node stations 11 and 12l was weak for the superior mediastinal lymph node station 5 (n = 1 each) and station 6 (n = 2 and 1).

Web Application Development

We developed an interactive web application to visualize these lymph node metastasis patterns (Figure 3 and Video 1) and released it at http://www.canexapp.com. This intuitive web tool provides a clear visualization of the complex patterns of lymph node metastasis from real patients. See Supplementary Material for more information.

DISCUSSION

In this study, we performed a network analysis to visualize the incidence of lymph node metastasis according to

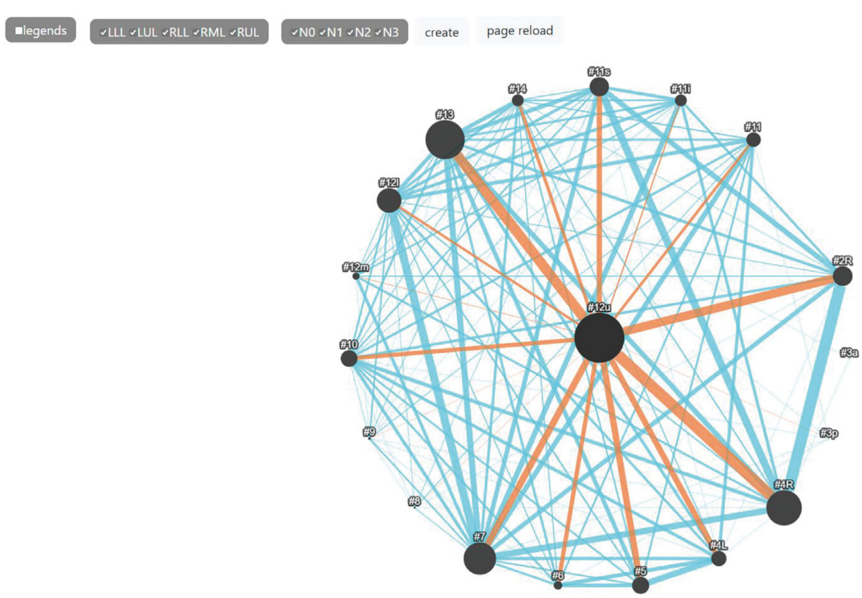
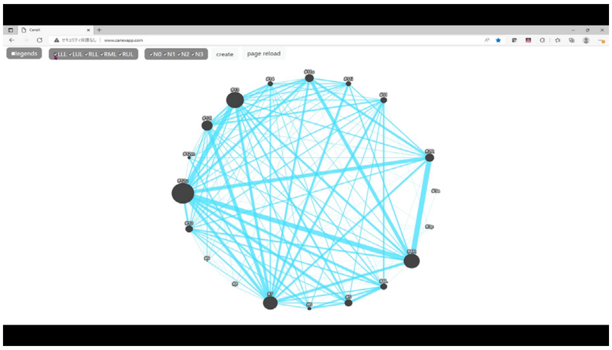


FIGURE 3. An interactive web application was developed to visualize patterns of lymph node metastasis. This interactive web application can be found at http://www.canexapp.com. Lymph node stations from stations 1 to 14 are arranged in a circumferential direction. Metastatic lymph node stations are indicated as *black dots*. A *larger black dot* represents a larger number of positive cases. Two stations are connected with a *line* if they share the same patients. A *thicker line* represents a larger number of patients shared. Clicking each lymph node station will display the number of cases with lymph node metastasis that occurred concurrently with the lymph node station in the selected station and other stations. In this example, the links between station 12u and other stations for all patients are shown. Network analysis will help thoracic surgeons understand patterns of lymph node metastases to perform patient-specific lymph node dissection. *LLL*, Left lower lobe; *LUL*, left upper lobe; *RLL*, right lower lobe; *RML*, right middle lobe; *RUL*, right upper lobe.



VIDEO 1. A tutorial of our interactive web application, which can be used to visualize patterns of lymph node metastasis. This application can be accessed at http://www.canexapp.com. Video available at: https://www.jtcvs.org/article/S2666-2736(22)00367-9/fulltext.

primary lung cancer tumor site and the degree of the association between lymph node stations. Several studies on lymph node metastasis patterns in primary lung cancer have been published.[18-20] Although these studies showed that patterns of lymph node metastasis could vary depending on primary tumor sites, they calculated the incidence of metastasis only between major lymph node stations. Results of our study were consistent with the previous reports, but we further expanded on them by applying a data mining technique to clearly visualize the patterns of more complex lymph node metastases that might be difficult to recognize by human processing. The connections between lymph node stations are displayed as lines, with the thickness corresponding to the degree of association. In addition, we developed an intuitive web application that can be used to visualize connections within each lobe or each pathological N category. Our results will help thoracic surgeons understand patterns of lymph node metastases.

We found that lymph node metastasis patterns differed depending on which lobe bore the tumor (Figure 2). The surgical procedure of mediastinal lymph node dissection recommended by the European Society of Thoracic Surgeons guidelines is an extensive and predetermined systematic mediastinal lymph node dissection from the superior to inferior mediastinum for the left or right side, regardless of which lobe is involved or any intraoperative findings.[21] Retrospective studies on lymph node metastasis patterns conducted in Japan since the 1990s suggest that some mediastinal lymph nodes are prone to developing metastasis and others rarely develop them, and this depends on which lobes have the primary tumor.[18-20] Accordingly, the concept of selective mediastinal lymph node dissection without dissecting the mediastinal lymph nodes where metastasis rarely occurs is being introduced into daily practice.[22] The results of our study clearly indicated that the patterns of lymph node metastases vary depending on which lobe contains the tumor—this supports the idea that the extent of dissection (and which lymph nodes are selected for dissection) should be determined on the basis of the lobe bearing the tumor. We eagerly await the results of an ongoing randomized controlled study (Japan Clinical

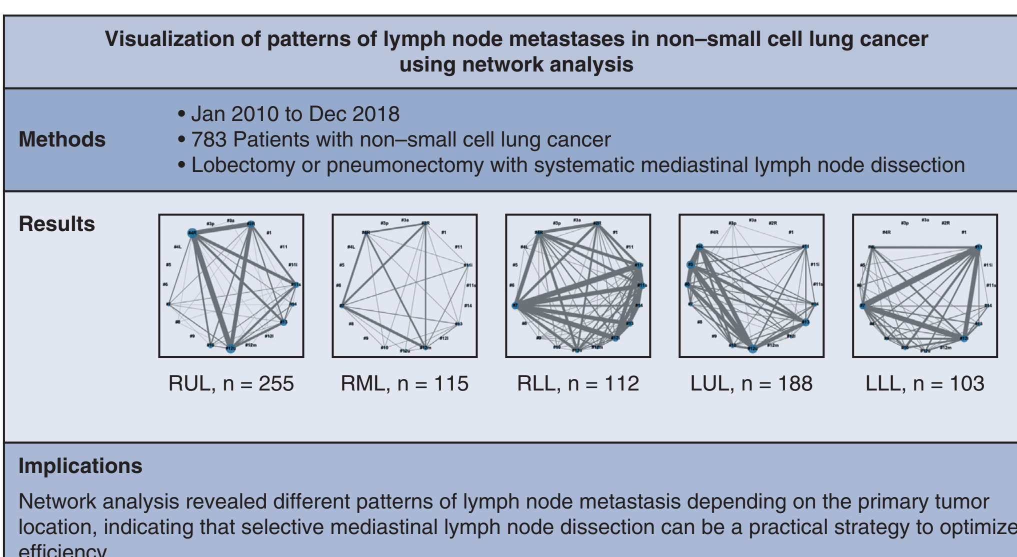


FIGURE 4. Network analysis applying graph theory was used to visualize lymph node metastasis patterns on the basis of the lobe bearing the primary cancer. Lymph node stations from stations 1 to 14 were arranged in a circumferential direction. Metastatic lymph node stations are indicated as *blue dots*. A *larger blue dot* represents a larger number of positive cases. Two stations are connected with a *line* if they share the same patients. A *thicker line* represents a larger number of patients shared. *RUL*, Right upper lobe; *RML*, right middle lobe; *RLL*, right lower lobe; *LUL*, left upper lobe; *LLL*, left lower lobe.

Oncology Group [JCOG] 1413) conducted by the Japan Clinical Oncology Group, which is examining the clinical efficacy of lobectomy and selective mediastinal lymph node dissection in clinical stage I/II NSCLC compared with conventional lobectomy and systematic mediastinal lymph node dissection.[23] In the JCOG 1413 trial, negative hilar (stations 10 and 11) node on frozen section is required for cN1 tumor suspected of hilar node metastasis preoperatively. Thus, in select cases, more time and resources are needed for selective mediastinal lymph node dissection than for systematic mediastinal lymph node dissection.

Through our study, we are pioneering research in a novel technique, network analysis, to elucidate lymph node metastasis patterns. We visualized distinct patterns of lymph node metastasis and showed the potential usefulness of network analysis for thoracic surgeons in determining the surgical approach for individual patients, although several challenges still remain. First, we classified tumor localization into 5 lobes. However, another study reported that the patterns of lymph node metastasis were different between a lung cancer located at the superior segment (S6) and one located in the basal segments of the lower lobe, suggesting that more detailed patterns of lymph node metastasis should be examined according to segment even in the same lobe.[24] Analysis according to lobe, rather than analysis according to segment, was performed in the present study because the number of patients classified as N2 was relatively small: 222 (28.4%) of a total of 783 patients. A larger number of patients is needed to analyze the detailed patterns of lymph node metastasis according to segment. Similarly, analysis according to histology (ie, adenocarcinoma and squamous cell carcinoma) will be a future plan, when larger sample numbers become available.

Second, the analysis in this study was performed using a dichotomous variable for the presence of metastasis—positive or negative—instead of the number of metastatic lymph nodes at each lymph node station as recorded in the pathology reports. This dichotomous variable was used to simplify the analysis while eliminating bias associated with surgical techniques, because the number of lymph nodes can increase if they are split or fragmented during dissection. However, one study reported that staging on the basis of the number of positive lymph nodes more accurately reflects prognosis in primary lung cancer.[25] Thus, network analysis on the basis of the number of positive lymph nodes should be assessed in the future.

Third, this study was a retrospective analysis on the basis of the metastatic status of lymph nodes confirmed after dissection. Network analysis generated patterns of lymph node metastases by defining any lymph node station with metastasis as 1 and lymph node stations without metastasis as 0 (see Figure E1). Examination of covariables was not possible in our study, and thus we were not able to account for other factors that might be driving the development of metastatic disease. In light of the advancements in precision medicine developed using artificial intelligence in recent years,[26] it might be possible to predict the presence or absence of pathological lymph node metastasis for each station from preoperative information including clinical patient information and data from CT or PET/CT imaging.

This study contains several limitations. First, this study was a single-center retrospective study. Second, sample size was limited, particularly the number of patients positive for lymph node metastasis. Because the purpose of this study was to elucidate the patterns of lymph node metastasis at all stations (stations 1-14), only patients who underwent systematic mediastinal lymph node dissection were included in the analysis. We also excluded patients who underwent segmentectomy because complete assessment of N1 nodes was difficult in these cases, leading to a lack of information, and we excluded patients who received preoperative treatments for cN2 NSCLC because pathological findings are influenced by preoperative treatments. Thus, our results might have been biased because of these selection criteria. A larger number of patients need to be accumulated for artificial intelligence-based analysis. To overcome these limitations, we are currently planning a multi-institutional study on the basis of a large-scale nationwide database. Third, the definition of systematic mediastinal lymph node dissection in our study is different from the one used in the United States and the area of systematic mediastinal lymph node dissection did not include stations 8 and 9 for right and left upper lobe tumors. In the ongoing JCOG 1413 prospective clinical trial,[23] the same definition for systematic mediastinal lymph node dissection as in our study is adopted on the basis of the definitions of the classification of lung cancer published by the Japan Lung Cancer Society[16] and the usual procedure in Japanese clinical practice. Lymph nodes at stations 8 and 9 were removed and submitted to pathological examination if metastasis was suspected at the time of intrathoracic evaluation during the surgery, and we believe the possibility that certain positive N2 lymph nodes were missed was extremely rare and did not influence results of our study.

CONCLUSIONS

We applied a data mining technique in this retrospective study to clarify the patterns of lymph node metastasis of primary lung cancer. We further used these data to develop an intuitive web tool that enables users to clearly visualize the complex patterns of lymph node metastasis from real patients. Our results showed that patterns of lymph node metastases differ according to the lobe bearing the tumor; these data form the basis for clinical trials to further evaluate selective mediastinal lymph node dissection (Figure 4). However, before the standard of practice can be changed to encompass selective mediastinal lymph node dissection, we have to wait for the results of clinical trials such as JCOG 1413.[23]

Conflict of Interest Statement

The authors reported no conflicts of interest.

The *Journal* policy requires editors and reviewers to disclose conflicts of interest and to decline handling or reviewing manuscripts for which they may have a conflict of interest. The editors and reviewers of this article have no conflicts of interest.

The authors acknowledge Ms Kumiko Nakatsuji for the English language editing and Dr Katherine Halievski for providing scientific writing support.

References

1. Cahan WG. Radical lobectomy. *J Thorac Cardiovasc Surg*. 1960;39:555-72.
2. Yl Wu, Huang ZF, Wang SY, Yang XN, Ou W. A randomized trial of systematic nodal dissection in resectable non–small cell lung cancer. *Lung Cancer*. 2002;36:1-6.
3. Sugi K, Nawata K, Fujita N, Ueda K, Tanaka T, Matsuoka T, et al. Systematic lymph node dissection for clinically diagnosed peripheral non–small-cell lung cancer less than 2 cm in diameter. *World J Surg*. 1998;22:290-4.
4. Izbicki JR, Passlick B, Pantel K, Pichlmeier U, Hosch SB, Karg O, et al. Effectiveness of radical systematic mediastinal lymphadenectomy in patients with resectable non–small cell lung cancer: results of a prospective randomized trial. *Ann Surg*. 1998;227:138-44.
5. Darling GE, Allen MS, Decker PA, Ballman K, Malthaner RA, Inculet RI, 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;141:662-70.
6. Rusch VW, Asamura H, Watanabe H, Giroux DJ, Rami-Porta R, Goldstraw P. The IASLC lung cancer staging project: a proposal for a new international lymph node map in the forthcoming seventh edition of the TNM classification for lung cancer. *J Thorac Oncol*. 2009;4:568-77.
7. Chen LL, Blumm N, Christakis NA, Barabási AL, Deisboeck TS. Cancer metastasis networks and the prediction of progression patterns. *Br J Cancer*. 2009;101:749-58.
8. Okubo Y, Matsumoto Y, Tanaka M, Imabayashi T, Uezono Y, Watanabe SI, et al. Clinical validity of 25-gauge endobronchial ultrasound-guided transbronchial needle in lymph node staging of lung cancer. *J Thorac Dis*. 2021;13:3033-41.
9. Yoshino I, Yoshida S, Miyaoka E, Asamura H, Nomori H, Fujii Y, et al. Surgical outcome of stage IIIA- cN2/pN2 non–small-cell lung cancer patients in Japanese lung cancer registry study in 2004. *J Thorac Oncol*. 2012;7:850-5.
10. Postmus PE, Kerr KM, Oudkerk M, Senan S, Waller DA, Vansteenkiste J, et al. Early and locally advanced non–small-cell lung cancer (NSCLC): ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2017;28:iv1-21.
11. Union for International Cancer Control (UICC), Brierley JD, Gospodarowicz MK, Wittekind C. *TNM Classification of Malignant Tumours*. 8th ed. Wiley-Blackwell; 2017.
12. Asamura H, Chansky K, Crowley J, Goldstraw P, Rusch VW, Vansteenkiste JF, et al. The International Association for the Study of Lung Cancer Lung Cancer Staging Project: proposals for the revision of the N descriptors in the forthcoming 8th edition of the TNM classification for lung cancer. *J Thorac Oncol*. 2015;10:1675-84.
13. Travis WD, Brambilla E, Müller-Hermelink HK, Harris CC, eds. *Pathology and Genetics of Tumours of the Lung, Thymus and Heart (World Health Organization Classification of Tumours)*. 3rd ed. IARC Press; 2004.
14. Travis WD, Brambilla E, Burke AP, Marx A, Nicholson AG. *WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart*. 4th ed. International Agency for Research on Cancer; 2015.
15. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, et al. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol*. 2011;6:244-85.
16. Ichinose J, Murakawa T, Hino H, Konoeda C, Inoue Y, Kitano K, et al. Prognostic impact of the current Japanese nodal classification on outcomes in resected non–small cell lung cancer. *Chest*. 2014;146:644-9.
17. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*. 2003;13:2498-504.
18. Okada M, Tsubota N, Yoshimura M, Miyamoto Y. Proposal for reasonable mediastinal lymphadenectomy in bronchogenic carcinomas: role of subcarinal nodes in selective dissection. *J Thorac Cardiovasc Surg*. 1998;116:949-53.
19. Asamura H, Nakayama H, Kondo H, Tsuchiya R, Naruke T. Lobe-specific extent of systematic lymph node dissection for non–small cell lung carcinomas according to a retrospective study of metastasis and prognosis. *J Thorac Cardiovasc Surg*. 1999;117:1102-11.
20. Watanabe S, Asamura H, Suzuki K, Tsuchiya R. The new strategy of selective nodal dissection for lung cancer based on segment-specific patterns of nodal spread. *Interact Cardiovasc Thorac Surg*. 2005;4:106-9.
21. Lardinois D, De Leyn P, Van Schil P, Porta RR, Waller D, Passlick B, et al. ESTS guidelines for intraoperative lymph node staging in non–small cell lung cancer. *Eur J Cardiothorac Surg*. 2006;30:787-92.
22. Hishida T, Miyaoka E, Yokoi K, Tsuboi M, Asamura H, Kiura K, et al. Lobe-specific nodal dissection for clinical stage I and II NSCLC: Japanese multi-institutional retrospective study using a propensity score analysis. *J Thorac Oncol*. 2016;11:1529-37.
23. Hishida T, Saji H, Watanabe SI, Asamura H, Aokage K, Mizutani T, et al. A randomized phase III trial of lobe-specific vs. systematic nodal dissection for clinical stage I-II non–small cell lung cancer (JCOG1413). *Jpn J Clin Oncol*. 2018;48:190-4.
24. Watanabe S, Suzuki K, Asamura H. Superior and basal segment lung cancers in the lower lobe have different lymph node metastatic pathways and prognosis. *Ann Thorac Surg*. 2008;85:1026-31.
25. Wei S, Asamura H, Kawachi R, Sakurai H, Watanabe S. Which is the better prognostic factor for resected non–small cell lung cancer: the number of metastatic lymph nodes or the currently used nodal stage classification? *J Thorac Oncol*. 2011;6:310-8.
26. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018;18:500-10.

Key Words: network analysis, data mining, mediastinal lymph node dissection, systematic mediastinal lymph node dissection, selective mediastinal lymph node dissection, non–small cell lung cancer

APPENDIX 1

Methods

Patients. We performed endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA) or mediastinoscopy preoperatively in select cases. EBUS procedures have been described elsewhere.[E1] A visual survey of lymph nodes through EBUS-TBNA was performed in the order of N1, N2, and then N3 stations. Subsequently, lymph nodes with echoic findings suggestive of malignancy or at a location important for treatment protocols, depending on their metastatic status, were punctured in the order of N3, N2, and then N1 stations.

In addition to the highly selected patients with cN2 NSCLC (n = 7) who received surgery, others received surgery for the following reasons: underlying pulmonary disease (n = 15) in which radiation was contraindicated (i.e., idiopathic pulmonary fibrosis or infection), medically uncontrollable bloody sputum (n = 2), bulky primary tumor of 7 cm or greater (n = 2) in which surgery offered better local control, patient preference (n = 4), renal insufficiency in which platinum-based chemotherapy was contraindicated (n = 1), atypical carcinoid (n = 1), synchronous esophageal carcinoma (n = 1), and suspected lymph node metastasis at station 5 that was confirmed negative on frozen section (n = 1).

Results

There were two patients with pN3 in our study. One patient had idiopathic pulmonary fibrosis, and chemoradiation was contraindicated; this patient was referred to the surgical department, although preoperative EBUS identified a metastatic lymph node at station 4R. During the surgery, the lymph node at station 4L was enlarged and metastasis was suspected, and thus it was resected as much as possible from the right side. Pathological examination revealed a metastatic lymph node at station 4L. Another patient with lymph node metastasis at contralateral hilar station 10 was diagnosed as cN1 preoperatively. The lymph node at station 12 was positive on frozen section, and the lymph node at contralateral hilar station 10 along with station 7 was resected. Pathological examination revealed metastatic lymph nodes at both station 7 and contralateral hilar station 10. Apart from these two exceptions, resections of N3 nodes occur rarely at our institution.

Web application development. When the user first visits the website, the default setting is to display all cases. To display select cases, users can tick checkboxes marking the right upper, middle, and lower and left upper and lower lobes as well as pathological N0, N1, N2, and N3 cases, and then click the "create" button. Ticking the "legends" checkbox will display the number of cases who had concurrent lymph node metastasis in two lymph node stations. Clicking each lymph node station will display the number of cases with lymph node metastasis that occurred concurrently at the lymph node station in question and other stations. Clicking the "page reload" button will bring the user back to default settings.

E-Reference

E1. Okubo Y, Matsumoto Y, Tanaka M, Imabayashi T, Uezono Y, Watanabe SI, et al. Clinical validity of 25-gauge endobronchial ultrasound-guided transbronchial needle in lymph node staging of lung cancer. *J Thorac Dis*. 2021;13:3033-41.

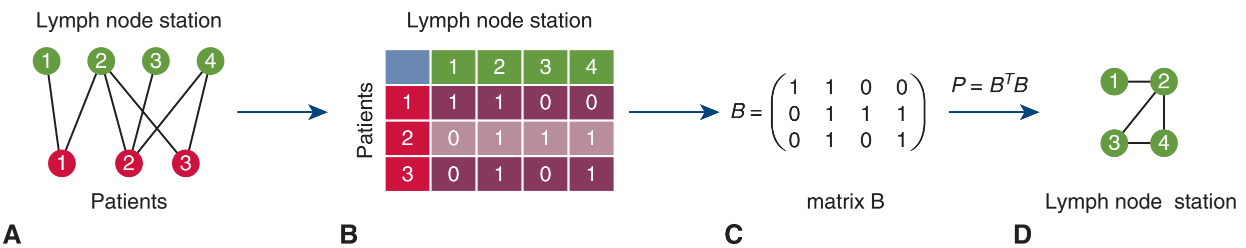


FIGURE E1. Assessing relationships between lymph node stations using graph theory. First, a bipartite graph was generated, with the patient (*red circle*) and each lymph node station (*green circle*) as nodes (A). In this example bipartite graph, patient 1 has lymph node metastases at station 1 and 2; patient 2 has lymph node metastases at station 2, 3, and 4; and patient 3 has lymph node metastases at stations 2 and 4. Next, the anonymized patient information and the pathological findings on lymph node metastasis at each lymph node station were represented as rows and columns, respectively (B). A lymph node station with metastasis was set to 1 and a lymph node station without metastasis was set to 0 (B, C). Matrix B (C) was then generated to show the relationship of each lymph node station with metastasis, which was finally represented as a weighted graph (D). Two stations are connected with a *line* if they share the same patients. A *thicker line* represents a larger number of patients shared (D).

TABLE E1. Pathological N category according to histologic subtype

Histologic subtype	Number	pN0	pN1	pN2	pN3
Adenocarcinoma	566	320	69	175	2
Squamous cell carcinoma	131	66	39	26	0
Pleomorphic carcinoma	32	15	7	10	0
Large cell neuroendocrine carcinoma	24	13	7	4	0
Adenosquamous carcinoma	14	8	3	3	0
Large cell carcinoma	5	2	3	0	0
Unclassified	4	1	2	1	0
Atypical carcinoid	2	0	0	2	0
Carcinosarcoma	2	2	0	0	0
Typical carcinoid	2	1	1	0	0
Adenoid cystic carcinoma	1	0	1	0	0

TABLE E2. Pathological N category according to clinical stage

Clinical stage	Number	pN0	pN1	pN2	pN3
IA1	19	17	1	1	0
IA2	157	107	15	35	0
IA3	189	122	16	51	0
IB	120	76	16	28	0
IIA	56	33	9	14	0
IIB	144	46	54	43	1
IIIA	86	26	19	40	1
IIIB	9	0	2	7	0
IVA	3	1	0	2	0

TABLE E3. The number of patients with concurrent metastases at 2 lymph node stations among those with right upper lobe primary lesions (n = 255)

	1	2R	3a	3p	4R	4L	5	6	7	8	9	10	11s	11i	11	12u	12m	12l	13	14
1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2R			0	0	30	0	0	0	3	1	0	8	14	0	0	27	0	1	8	1
3a				0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
3p					0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4R						0	0	0	6	1	0	11	21	0	0	36	1	1	15	3
4L							0	0	0	0	0	0	0	0	0	0	0	0	0	0
5								0	0	0	0	0	0	0	0	0	0	0	0	0
6									0	0	0	0	0	0	0	0	0	0	0	0
7										0	0	1	3	0	0	6	0	0	2	1
8											0	0	1	0	0	1	0	0	0	0
9												0	0	0	0	0	0	0	0	0
10													4	0	0	5	0	0	2	0
11s														0	0	15	1	0	7	1
11i															0	1	0	0	0	0
11																0	0	0	0	0
12u																	0	0	16	5
12m																		0	0	0
12l																			0	0
13																				3
14																				

The number in each cell represents the number of patients with concurrent metastases at 2 lymph node stations indicated in the corresponding row and column.

TABLE E4. The number of patients with concurrent metastases at 2 lymph node stations among those with right middle lobe primary lesions (n = 115)

	1	2R	3a	3p	4R	4L	5	6	7	8	9	10	11s	11i	11	12u	12m	12l	13	14
1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2R			0	0	5	0	0	0	6	0	0	1	0	1	0	1	4	0	1	0
3a				0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3p					0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4R						0	0	0	6	0	0	1	0	1	0	0	5	0	1	0
4L							0	0	0	0	0	0	0	0	0	0	0	0	0	0
5								0	0	0	0	0	0	0	0	0	0	0	0	0
6									0	0	0	0	0	0	0	0	0	0	0	0
7										0	0	1	0	2	0	1	8	0	2	0
8											0	0	0	0	0	0	0	0	0	0
9												0	0	0	0	0	0	0	0	0
10													0	0	0	0	0	0	1	0
11s														0	0	0	0	0	0	0
11i															0	0	2	0	1	0
11																0	0	0	0	0
12u																	1	0	1	0
12m																		0	2	0
12l																			0	0
13																				0
14																				

The number in each cell represents the number of patients with concurrent metastases at 2 lymph node stations indicated in the corresponding row and column.

TABLE E5. The number of patients with concurrent metastases at 2 lymph node stations among those with right lower lobe primary lesions (n = 122)

	1	2R	3a	3p	4R	4L	5	6	7	8	9	10	11s	11i	11	12u	12m	12l	13	14
1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2R			0	0	10	1	0	0	8	0	0	1	6	8	0	6	1	2	2	1
3a				0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3p					0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4R						1	0	0	12	0	1	2	8	10	0	8	2	4	3	2
4L							0	0	1	0	0	0	1	0	0	1	0	0	0	0
5								0	0	0	0	0	0	0	0	0	0	0	0	0
6									0	0	0	0	0	0	0	0	0	0	0	0
7										0	1	2	14	18	0	9	4	16	13	6
8											0	0	1	0	0	0	0	0	0	0
9												0	1	1	0	0	1	1	1	1
10													2	2	0	1	0	1	0	0
11s														12	0	6	2	11	9	3
11i															0	4	4	12	7	6
11																0	0	0	0	0
12u																	1	3	3	1
12m																		3	2	2
12l																			9	5
13																				5
14																				

The number in each cell represents the number of patients with concurrent metastases at 2 lymph node stations indicated in the corresponding row and column.

TABLE E6. The number of patients with concurrent metastases at 2 lymph node stations among those with left upper lobe primary lesions (n = 188)

	1	2R	3a	3p	4R	4L	5	6	7	8	9	10	11s	11i	11	12u	12m	12l	13	14
1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2R			0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3a				0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3p					0	0	0	1	1	0	0	1	0	0	1	1	0	0	1	0
4R						0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4L							23	13	5	0	0	10	0	0	6	21	0	4	10	4
5								14	5	0	0	7	0	0	6	24	0	5	16	6
6									6	0	0	7	0	0	5	17	0	3	9	4
7										0	0	3	0	0	3	7	0	1	3	1
8											0	0	0	0	0	0	0	0	0	0
9												0	0	0	0	0	0	0	0	0
10													0	0	3	11	0	1	5	3
11s														0	0	0	0	0	0	0
11i															0	0	0	0	0	0
11																7	0	2	6	4
12u																	0	3	19	6
12m																		0	0	0
12l																			2	1
13																				7
14																				

The number in each cell represents the number of patients with concurrent metastases at 2 lymph node stations indicated in the corresponding row and column.

TABLE E7. The number of patients with concurrent metastases at 2 lymph node stations among those with left lower lobe primary lesions (n = 103)

	1	2R	3a	3p	4R	4L	5	6	7	8	9	10	11s	11i	11	12u	12m	12l	13	14
1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2R			0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3a				0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3p					0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4R						0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4L							0	2	6	0	1	2	0	0	6	1	0	5	1	0
5								0	1	0	0	0	0	0	1	1	0	1	0	0
6									2	0	1	1	0	0	2	0	0	1	0	0
7										3	5	5	0	0	16	3	0	14	3	2
8											1	1	0	0	3	0	0	2	0	0
9												2	0	0	3	1	0	5	0	2
10													0	0	6	0	0	5	0	0
11s														0	0	0	0	0	0	0
11i															0	0	0	0	0	0
11																4	0	12	2	1
12u																	0	4	1	1
12m																		0	0	0
12l																			4	3
13																				0
14																				

The number in each cell represents the number of patients with concurrent metastases at 2 lymph node stations indicated in the corresponding row and column.