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

Yukihiro Yoshida, MD, Nozomu Saeki, MPhil, Masaya Yotsukura, MD, Kazuo Nakagawa, MD, Hirokazu Watanabe, MD, Yasushi Yatabe, MD, Shun-ichi Watanabe, MD

PII: S2666-2736(22)00367-9
DOI: https://doi.org/10.1016/j.xjon.2022.10.003
Reference: XJON 667

To appear in: JTCVS Open

Received Date: 19 June 2022
Accepted Date: 7 October 2022

Please cite this article as: Yoshida Y, Saeki N, Yotsukura M, Nakagawa K, Watanabe H, Yatabe Y, Watanabe Si, Visualization of patterns of lymph node metastases in non-small cell lung cancer using network analysis, JTCVS Open (2022), doi: https://doi.org/10.1016/j.xjon.2022.10.003.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Copyright © 2022 The Authors. Published by Elsevier Inc. on behalf of The American Association for Thoracic Surgery
Visualization of patterns of lymph node metastases in non-small cell lung cancer using network analysis

Methods

- Jan 2010 to Dec 2018
- 783 patients with non-small cell lung cancer
- Lobectomy or pneumonectomy with systematic mediastinal lymph node dissection

Results

<table>
<thead>
<tr>
<th>Lobe</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUL</td>
<td>n=255</td>
</tr>
<tr>
<td>RML</td>
<td>n=115</td>
</tr>
<tr>
<td>RLL</td>
<td>n=112</td>
</tr>
<tr>
<td>LUL</td>
<td>n=188</td>
</tr>
<tr>
<td>LLL</td>
<td>n=103</td>
</tr>
</tbody>
</table>

LLL, left lower lobe; LUL, left upper lobe; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe.

Implications

Network analysis revealed different patterns of lymph node metastasis depending on the primary tumor location, indicating that selective mediastinal lymph node dissection can be a practical strategy to optimize efficiency.
Title: Visualization of patterns of lymph node metastases in non-small cell lung cancer using network analysis

Yukihiro Yoshida, MD¹, Nozomu Saeki, MPhil², Masaya Yotsukura, MD³, Kazuo Nakagawa, MD¹, Hirokazu Watanabe, MD³, Yasushi Yatabe, MD⁴, Shun-ichi Watanabe, MD¹

¹ National Cancer Center Hospital, Department of Thoracic Surgery, 5-1-1 Tsukiji, Chuo-ku, Tokyo, 104-0045 Japan
² Graduate School of Environmental and Life Science, Okayama University, 3-1-1 Tsushima-machi, Kita-ku, Okayama-shi, Okayama, 700-8530 Japan
³ National Cancer Center Hospital, Department of Diagnostic Radiology, 5-1-1 Tsukiji, Chuo-ku, Tokyo, 104-0045 Japan
⁴ National Cancer Center Hospital, Department of Diagnostic Pathology, 5-1-1 Tsukiji, Chuo-ku, Tokyo, 104-0045 Japan

Disclosure statement: The authors have no conflicts of interest associated with this manuscript to disclose.
Funding statement: The authors did not receive any grant from funding agencies in the public, commercial, or not-for-profit sectors for this research.

Corresponding author:
Yukihiro Yoshida, MD
Department of Thoracic Surgery, National Cancer Center Hospital, Tokyo, Japan
5-1-1 Tsukiji, Chuo-ku, Tokyo, 104-0045 Japan
TEL: +81-3-3542-2511
FAX: +81-3-3542-3815
E-mail: yukiyosh@ncc.go.jp

Date and number of IRB approval: June 11, 2020 and No. 2020-057
This was a retrospective study with no interventions or invasive procedures, and thus exempt from written consent.

Article word count: 3403
Glossary of Abbreviations

CT = computed tomography

Endobronchial ultrasound-guided transbronchial needle aspiration = EBUS-TBNA

IASLC = International Association for the Study of Lung Cancer

Non-small cell lung cancer = NSCLC

Positron emission tomography = PET

Central Message

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

Perspective Statement

Current guidelines recommend systematic mediastinal lymph node dissection for patients undergoing 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.
Structured 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 two 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 towards 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.

**Keywords:** network analysis; data mining; mediastinal lymph node dissection; systematic mediastinal lymph node dissection; selective mediastinal lymph node dissection; non-small cell lung cancer

**Abstract word count:** 239 words
Introduction

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) 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 impact on prognosis by removing lymph nodes alongside the primary tumor. In the 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 to 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 by site may 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. We hypothesized that a detailed analysis of lymph node metastasis patterns may be possible by applying this theory in primary lung cancer. This retrospective study 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 flowchart 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 one year prior to surgery was considered a former smoker. The standard preoperative workup for staging included contrast-enhanced computed tomography (CT) of the chest and upper abdomen and whole-body positron emission tomography (PET)/CT. Brain CT or magnetic resonance imaging was added if necessary. Endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA)
or mediastinoscopy was recommended if there was a mediastinal lymph node > 1 cm in
diameter along the short-axis on CT or if there was increased uptake on PET/CT
(Date/Number of IRB Approval: June 11, 2020 and No. 2020-057). This was a retrospective
study with no interventions or invasive procedures, and thus exempt from written consent.
See Supplementary Material for more information. 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 based on the
satisfactory surgical outcomes observed for cN2 NSCLC in the Japanese lung cancer registry
study and the European Society for Medical Oncology clinical practice guidelines. See
Supplementary Material for additional reasons that certain patients received surgery. The
IASLC lymph node map was used to define the lymph node stations. 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,\textsuperscript{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).\textsuperscript{12} The third\textsuperscript{13} or fourth\textsuperscript{14} editions of WHO classification or the 2011 IASLC/American Thoracic Society/European Respiratory Society classification\textsuperscript{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.\textsuperscript{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 No.: 2020-057). This was a retrospective study with no invasive procedures or interventions, and thus exempt from requiring written consent from each patient.

\textit{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 (Supplementary Figure 1). 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 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 non-diagonal component represented the number of patients who had metastatic lymph nodes concurrently at two stations.

Interactive visualization of the network between lymph node stations

Cytoscape (3.14.2) was used to visualize the lymph node metastases at each lymph node station. We also used Python (3.6.9), Numpy (1.18.2), and Pandas (1.0.3) for data
processing and matrix computation. MySQL (14.14), Django (3.0.4), Django-mysql (3.5.0), and mysqlclient (1.4.6) were used to visualize the data by 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 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 by 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, 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: one 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
medians (ranges) of the number of lymph node metastases were 1 (1–11) in 84 patients with N1a, 3 (2–9) in 48 with N1b, 1 (1–9) in 33 with N2a1, 3 (1–13) in 93 with N2a2, and 8 (2–52) in 95 with N2b. Pathological N category according to histologic subtype and clinical stage is shown in Supplementary Tables 1 and 2, respectively.

Visualization of lymph node metastasis patterns by network analysis

Results of the network analysis used to visualize lymph node metastasis patterns based on the lobe bearing the tumor are shown in Figure 2. Supplementary Tables 3-7 show the number of patients, by lobe, who had metastases concurrently at two lymph node stations, and Table 3 shows the top five combinations of two metastatic lymph node stations as determined by 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; 36 and 14.1% for stations 12u and 4R), but they had poor connections with the inferior mediastinal lymph nodes including subcarinal lymph nodes (i.e., only six patients [2.4%, 6/255] had concurrent metastases at stations 12u and 7).
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.

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 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).

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 (i.e., only seven patients had concurrent metastases at stations 12u and 7).

In the left lower lobe, connections were noted from stations 11 and 12l to station 7 (n = 16 and 14), while 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 Supplementary video) 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 by 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 demonstrated 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 may 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. 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. 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. 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
based on the lobe bearing the tumor. We eagerly await the results of an ongoing randomized
controlled study (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 non-small cell lung cancer compared to conventional lobectomy and systematic mediastinal lymph node dissection.\(^\text{23}\) 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 demonstrated the potential usefulness of network analysis for thoracic surgeons in determining the surgical approach for individual patients, though several challenges still remain. First, we classified tumor localization into five 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 by segments even in the same lobe.\(^\text{24}\) Analysis by lobe, rather than analysis by segment, was performed in the present study because the number of patients classified as N2 was relatively small: 222 (28.4%) out of total 783 patients. A larger number of patients is needed to analyze the detailed patterns of lymph node metastasis by segments. Similarly,
analysis according to histology (i.e., 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 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 employed to simplify the analysis while eliminating bias associated with surgical techniques, since the number of lymph nodes can increase if they are split or fragmented during dissection. However, one study reported that staging based on the number of positive lymph nodes more accurately reflects prognosis in primary lung cancer. Thus, network analysis based on the number of positive lymph nodes should be assessed in future.

Third, this study was a retrospective analysis based on 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 Supplementary Figure 1). Looking at covariables was not possible in our study, and thus we were not able to account for other factors that may be driving the development of metastatic disease. In light of the advancements in precision medicine developed using artificial intelligence in recent years, it may be possible to predict the presence or absence of pathological lymph node metastasis for each station from preoperative information including clinical patient information and imaging data from CT or
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. As the purpose of this study was to elucidate the patterns of lymph node metastasis at all stations (stations 1 to 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 by 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 based on 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 US and the area of systematic mediastinal lymph node dissection did not include stations 8 and 9 for both right and left upper lobe tumors. In the ongoing JCOG 1413 prospective clinical trial, the same definition for systematic mediastinal lymph node dissection as in our study is adopted based on the definitions of the Classification of Lung Cancer published by the Japan Lung Cancer Society 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.

In conclusion, 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 this
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.

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
Acknowledgements

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


24. Watanabe S, Suzuki K, Asamura H. Superior and basal segment lung cancers in the


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

<table>
<thead>
<tr>
<th>Variables</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>26–84</td>
<td></td>
</tr>
<tr>
<td>Sex, n</td>
<td>Male</td>
<td>446</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>337</td>
</tr>
<tr>
<td>Smoking status, n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>208</td>
<td>26.6%</td>
</tr>
<tr>
<td>Former</td>
<td>321</td>
<td>41.0%</td>
</tr>
<tr>
<td>Never</td>
<td>254</td>
<td>32.4%</td>
</tr>
<tr>
<td>Entire lesion diameter on CT, cm</td>
<td>Median</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td>0.8–11.0</td>
</tr>
<tr>
<td>Clinical N category, n</td>
<td>cN0</td>
<td>611</td>
</tr>
<tr>
<td></td>
<td></td>
<td>78.0%</td>
</tr>
<tr>
<td></td>
<td>cN1</td>
<td>138</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17.6%</td>
</tr>
<tr>
<td></td>
<td>cN2</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.3%</td>
</tr>
<tr>
<td>Clinical stage, n</td>
<td>IA1</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.4%</td>
</tr>
<tr>
<td></td>
<td>IA2</td>
<td>157</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20.1%</td>
</tr>
<tr>
<td></td>
<td>IA3</td>
<td>189</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>IB</td>
<td>120</td>
<td>15.3%</td>
</tr>
<tr>
<td>IIA</td>
<td>56</td>
<td>7.2%</td>
</tr>
<tr>
<td>IIB</td>
<td>144</td>
<td>18.4%</td>
</tr>
<tr>
<td>IIIA</td>
<td>86</td>
<td>11.0%</td>
</tr>
<tr>
<td>IIIB</td>
<td>9</td>
<td>1.1%</td>
</tr>
<tr>
<td>IVA</td>
<td>3</td>
<td>0.4%</td>
</tr>
</tbody>
</table>

**Lesion site, n**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RUL</td>
<td>255</td>
<td>32.6%</td>
</tr>
<tr>
<td>RML</td>
<td>115</td>
<td>14.7%</td>
</tr>
<tr>
<td>RLL</td>
<td>122</td>
<td>15.6%</td>
</tr>
<tr>
<td>LUL</td>
<td>188</td>
<td>24.0%</td>
</tr>
<tr>
<td>LLL</td>
<td>103</td>
<td>13.2%</td>
</tr>
</tbody>
</table>

**Surgical procedure, n**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pneumonectomy</td>
<td>28</td>
<td>3.6%</td>
</tr>
<tr>
<td>Lobectomy</td>
<td>755</td>
<td>96.4%</td>
</tr>
</tbody>
</table>

**Lymph nodes dissected per patient, n**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td></td>
<td>20</td>
</tr>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td></td>
<td>4–60</td>
</tr>
</tbody>
</table>

**Histology, n**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenocarcinoma</td>
<td>566</td>
<td>72.3%</td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>131</td>
<td>16.7%</td>
</tr>
<tr>
<td>Pathological N category, n</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Other</td>
<td>86</td>
<td>11.0%</td>
</tr>
<tr>
<td>pN0</td>
<td>428</td>
<td>54.7%</td>
</tr>
<tr>
<td>pN1a</td>
<td>84</td>
<td>10.7%</td>
</tr>
<tr>
<td>pN1b</td>
<td>48</td>
<td>6.1%</td>
</tr>
<tr>
<td>pN2a1</td>
<td>33</td>
<td>4.2%</td>
</tr>
<tr>
<td>pN2a2</td>
<td>93</td>
<td>11.9%</td>
</tr>
<tr>
<td>pN2b</td>
<td>95</td>
<td>12.1%</td>
</tr>
<tr>
<td>pN3</td>
<td>2</td>
<td>0.3%</td>
</tr>
</tbody>
</table>

CT, computed tomography; LLL, left lower lobe; LUL, left upper lobe; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe.
Table 2. Pathological N category by lobe of the primary tumor.

<table>
<thead>
<tr>
<th>Pathological N category</th>
<th>N0</th>
<th>N1a</th>
<th>N1b</th>
<th>N2a1</th>
<th>N2a2</th>
<th>N2b</th>
<th>N3</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUL</td>
<td>255</td>
<td>145</td>
<td>21</td>
<td>13</td>
<td>14</td>
<td>28</td>
<td>34</td>
</tr>
<tr>
<td>RML</td>
<td>115</td>
<td>90</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>RLL</td>
<td>122</td>
<td>57</td>
<td>16</td>
<td>11</td>
<td>3</td>
<td>21</td>
<td>13</td>
</tr>
<tr>
<td>LUL</td>
<td>188</td>
<td>92</td>
<td>24</td>
<td>10</td>
<td>7</td>
<td>26</td>
<td>29</td>
</tr>
<tr>
<td>LLL</td>
<td>103</td>
<td>44</td>
<td>19</td>
<td>11</td>
<td>3</td>
<td>13</td>
<td>13</td>
</tr>
</tbody>
</table>

Data are shown as n.

LLL, left lower lobe; LUL, left upper lobe; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe.
Table 3. The number of patients with concurrent metastases at two stations by lobe of the primary tumor.

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

Top five lymph node station combinations ranked by the largest number of patients are presented.

LLL, left lower lobe; LUL, left upper lobe; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe.
Figure Legends

Figure 1. Flowchart of patient selection.

Figure 2. Network analysis–based visualization of lymph node metastasis patterns by 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. LLL, left lower lobe; LUL, left upper lobe; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe.

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 in order to perform patient-specific lymph node dissection.

Figure 4. Graphical Abstract. Network analysis applying graph theory was used to visualize lymph node metastasis patterns based on 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. LLL, left lower lobe; LUL, left upper lobe; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe.

Central picture legend.

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

Supplementary figure legend
Supplementary Figure 1. Assessing relationships between lymph node stations using graph theory. First, a bipartite graph was generated, with the patient (orange 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).

Supplementary video legend. 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/.
Patients who underwent surgery for primary lung cancer between 2010 and 2018, n = 4133

With history of primary lung cancer, n = 273

n = 3860

With preoperative treatment, n = 100

Bilateral simultaneous surgery, n = 32

Unilateral surgery, n = 3728

Small cell lung cancer, n = 81

Non-small cell lung cancer, n = 3647

Other procedures, n = 998

Lobectomy/pneumonectomy, n = 2649

Synchronous multiple lung cancer (n = 77)/multilobe involvement (n = 99)

n = 2473

Without mediastinal lymph node dissection, n = 405

With mediastinal lymph node dissection, n = 2068

Selective lymph node dissection, n = 1285

Systematic lymph node dissection, n = 783 (study population)
Visualization of patterns of lymph node metastases in non-small cell lung cancer using network analysis

**Methods**
- Jan 2010 to Dec 2018
- 783 patients with non-small cell lung cancer
- Lobectomy or pneumonectomy with systematic mediastinal lymph node dissection

**Results**

<table>
<thead>
<tr>
<th>Lobe</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUL</td>
<td>n=255</td>
</tr>
<tr>
<td>RML</td>
<td>n=115</td>
</tr>
<tr>
<td>RLL</td>
<td>n=112</td>
</tr>
<tr>
<td>LUL</td>
<td>n=188</td>
</tr>
<tr>
<td>LLL</td>
<td>n=103</td>
</tr>
</tbody>
</table>

LLL, left lower lobe; LUL, left upper lobe; RLL, right lower lobe; RML, right middle lobe; RUL, right upper lobe.

**Implications**

Network analysis revealed different patterns of lymph node metastasis depending on the primary tumor location, indicating that selective mediastinal lymph node dissection can be a practical strategy to optimize efficiency.