The Quality of Staging Non-Small Cell Lung Cancer in the Netherlands: Data From the Dutch Lung Surgery Audit

David Jonathan Heineman, MD, Martijn Geert ten Berge, MD, Johannes Marlene Daniels, MD, PhD, Michaël Ignatius Versteegh, MD, Perla Jacqueline Marang-van de Mheen, PhD, Michael Wilhelmus Wouters, MD, PhD, and Wilhelmina Hendrika Schreurs, MD, PhD

Department of Surgery, Medisch Centrum Alkmaar, Alkmaar; Departments of Surgery, Cardiothoracic Surgery, and Medical Decision Making, Leiden University Medical Center, Leiden; Department of Pulmonary Diseases, VU University Medical Centre, Amsterdam; Department of Surgical Oncology, Antoni van Leeuwenhoek Hospital/Netherlands Cancer Institute, Amsterdam; and Scientific Bureau, Dutch Institute for Clinical Auditing, Leiden, the Netherlands

Background. **Clinical staging of non-small cell lung cancer (NSCLC) determines the initial treatment offered to a patient. The similarity between clinical and pathologic staging in some studies is as low as 50%, and others publish results as high as 91%. The Dutch Lung Surgery Audit is a clinical database that registers the clinical and pathologic TNM of almost all NSCLC patients who undergo operations in the Netherlands. The objective of this study was to determine the accuracy of clinical staging of NSCLC.**

Methods. **Prospective data were derived from the Dutch Lung Surgery Audit in 2013 and 2014. Patients were included if they had undergone a surgical resection for stage IA to IIIB NSCLC without neoadjuvant treatment and had a positron emission tomography–computed tomography scan as part of the clinical workup. Clinical (c)TNM and pathologic (p)TNM were compared, and whether discrepancy was based on tumor or nodal staging was determined.**

Results. **From 2,834 patients identified, 2,336 (82.4%) fulfilled the inclusion criteria and had complete data. Of these 2,336, 1,276 (54.6%) were staged accurately, 707 (30.3%) were clinically understaged, and 353 (15.1%) were clinically overstaged. In the understaged group, 346 patients had a higher pN stage (14.8%), of which 148 patients had unforeseen N2 disease (6.3%). In the overstaged group, 133 patients had a cN that was higher than the pN (5.7%).**

Conclusions. **Accuracy of NSCLC staging in the Netherlands is low (54.6%), even in the era of positron emission tomography–computed tomography. Especially accurate nodal staging remains challenging. Future efforts should include the identification of specific pitfalls in NSCLC staging.**

(Ann Thorac Surg 2016;■:■–■)

Lung cancer staging involves a complex multidisciplinary process in which a specific combination of imaging modalities, minimally invasive staging procedures, or invasive staging procedures are selected for the individual patient to achieve accurate stage information with the lowest possible patient burden [1, 2]. In May 2013 the journal *Chest* published a supplement, "Methods for staging non-small cell lung cancer," an evidence-based clinical practice guideline by the American College of Chest Physicians. The authors of this guideline conducted a thorough review of the literature, and from that review they proposed a diagnostic work-up for patients suspected of non-small cell lung cancer (NSCLC) to assure accurate clinical (c)TNM staging. This staging routinely comprises a positron emission tomography (PET)–computed tomography (CT) scan, if available, and otherwise a CT scan. In case of a suspicious mediastinal node on the PET-CT, minimally invasive techniques, such as endoscopic ultrasound (EUS) and endobronchial ultrasound (EBUS), or a surgical biopsy (mediastinoscopy) are advised to stage the mediastinum. In patients with an intermediate risk for N2 or N3 involvement, with a central tumor or N1 lymph node involvement, invasive staging of the mediastinum is also recommended. Magnetic resonance imaging of the brain is recommended in patients with clinical stage III disease. Although data are published regarding the sensitivity and specificity of the different diagnostic modalities, the accuracy of the diagnostic process as a whole was not described in this publication of the American College of Chest Physicians [2].

The available studies on the individual staging techniques often had a retrospective design and a small sample size, which precludes robust conclusions. The accuracy of clinical staging is generally low, at approximately 50% to 60% [3–8]. A recent study from Denmark

Accepted for publication June 22, 2016.

Address correspondence to Dr Heineman, Department of Surgery, Medisch Centrum Alkmaar, Wilhelminalaan 12, 1815 JD Alkmaar, the Netherlands; email: david.heineman@gmail.com.

Published by Elsevier

0003-4975/$36.00
http://dx.doi.org/10.1016/j.athoracsur.2016.06.071

Abbreviations and Acronyms

ASA = American Society of Anesthesiologists
CI = confidence interval
CT = computed tomography
cTNM = clinical TNM
DLCO = diffusion capacity of the lung for carbon monoxide
DLSA = Dutch Lung Surgery Audit
EBUS = endobronchial ultrasound
ECOG = Eastern Cooperative Oncology Group
EUS = endoscopic ultrasound
FEV_1 = forced expiratory volume in 1 second
LLL = left lower lobe
LUL = left upper lobe
NSCLC = non-small cell lung cancer
PET = positron emission tomography
pTNM = pathologic TNM
RLL = right lower lobe
RML = right middle lobe
RUL = right upper lobe

showed an increase in the accuracy of the staging process from 68% to 91% in the last 10 years, possibly due to centralization of lung cancer care during this period [9].

The introduction of the PET-CT scan in the diagnostic work-up is also thought to benefit the accuracy of the full staging process, although little is known about this effect because most studies looking at the accuracy of staging date from the era before PET-CT. PET was introduced in the Netherlands in 1991 but was only widely used after 2007 [10]. Part of the improvement seen in the Danish study might also be due to introduction of the PET-CT scan. A high correlation between the cTNM and pathologic (p)TNM staging is considered very important, because inaccurate staging may cause undertreatment or overtreatment of patients, especially with the recent introduction of stereotactic ablative radiotherapy in early-stage lung cancer and induction therapy for stage IIIA tumors. Furthermore, the correlation between cTNM and pTNM is also an important indicator of the quality of the total diagnostic setup [9].

The main objective of this study was to assess the real-world accuracy of the staging process by evaluating discrepancies between cTNM and pTNM staging in a national database including patients who underwent surgical resection for NSCLC in the Netherlands.

Material and Methods

Data Source

The study used data from the Dutch Lung Surgery Audit (DLSA), a nationwide clinical registry used for evaluation of quality of care for benign and malignant lung operations. Information on patient characteristics, diagnostics, tumor characteristics, treatment, and outcomes has been recorded prospectively since 2012. The quality of this database is regularly checked, for example, by comparing the data with the Netherlands Cancer Registry, a database with data on the incidence, prevalence, survival, and death of all cancer types [11]. Completeness and data consistency of the DLSA is checked through queries, with results given as feedback to individual hospitals and requests to check any inconsistencies that are identified by these queries.

Patients

The study evaluated data for all patients who underwent anatomical parenchymal (pneumonectomy, [bi]lobectomy, or sublobar) lung resections between January 1, 2013, and December 31, 2014, and were registered in the DLSA. Minimal data requirements for inclusion in the analysis were information on cTNM and pTNM stage, type of parenchymal resection, and the histopathologic determination. Patients who presented with acute symptoms, other histopathology than NSCLC, stage IV lung cancer, neoadjuvant treatment, or no PET-CT scan were excluded.

Outcome

The primary outcome was accurate clinical staging, using the pTNM stage (The TNM Classification of Malignant Tumours, 7th Edition) as the gold standard. When cTNM stage was lower than pTNM stage this was considered as understaging, regardless of the extent of understaging (eg, 1A vs 1B or 1A vs 2A). Secondary outcomes were the number of misdiagnosed patients based on N and T stage. Patients misdiagnosed on both the N and T stage were placed in the N stage group because clinical consequences for inaccurate N staging were generally more important than for inaccurate T staging.

The use of invasive diagnostics, such as EUS, EBUS, and (video) mediastinoscopy, was investigated in suspicious nodes (enlarged [short axis of diameter >1 cm] or PET-positive) to analyze the adherence to guidelines on staging mediastinal lymph nodes. This was done by analyzing the results of negative invasive diagnostic studies in suspicious nodes that were pathologically reviewed. Because an accurate lymph node dissection or sampling is mandatory to provide correct pTNM staging, we analyzed the lymph node stations that were dissected or sampled.

Statistical Analysis

As a first step, we compared included with excluded patients on a number of preoperative patient characteristics and the clinical stage to be able to assess the generalizability of our results (selection bias). This was done using χ^2 tests and the Fisher exact test when expected counts were less than 5.

Secondly, we compared the clinical and pathologic stage of the included patients and estimated the accuracy for each clinical stage. Then we compared patients with accurate staging with patients with inaccurate staging on age, sex, performance score, comorbidities, previous thoracic operations, clinical stage, and tumor side using