Lung cancer is the most common cause of cancer-related mortality worldwide and in the US; non-small cell lung cancer (NSCLC) represents approximately 85% of all cases (1,2). Despite improvements in methods of diagnosis and treatment, the aggregate overall 5-year survival rate of all patients with lung cancer has only improved from 12% in the 1970s to 17% in contemporary times (2). This is largely because most patients present with advanced disease for which curative therapy is currently unavailable. However, patients with early stage disease who undergo definitive surgery or combined modality therapy may have long term survival. The most effective current prognostic tool is the Tumor, Node, and Metastasis (TNM) staging system which is currently in its 7th edition (3).

Staging, while of great prognostic value, is only as useful as the degree of thoroughness with which it is applied (4,5). Comparison of clinical and pathological (post-resection) staging survival curves on the same patients reveals greater separation between the pN0-3 subsets than the cN subsets, in part because pathologic staging defines a group of pN0 tumors with better survival and a group of pN3 tumors with worse survival than predicted by clinical staging.
alone (6). This reflects the fact that clinical staging tests have sensitivity and specificity limitations that impair their accuracy (7-9).

Pathologic nodal stage is the most important determinant of prognosis in patients who undergo resection for NSCLC, with survival ranging from 56% in patients with pN0 to 6% in pN3 (6). It is also the main driver of post-operative management. For example, patients with pN1-3 disease benefit from adjuvant chemotherapy (10-12), while those with mediastinal lymph node metastasis may benefit from radiation therapy in addition to chemotherapy (13). However, pN-stage is the TNM category most susceptible to variability in both surgical resection techniques and pathologic evaluation (14).

Examination of large databases, such as the California Cancer Registry, the Surveillance, Epidemiology, and End Results (SEER) database and the National Cancer Data Base (NCDB), reveals worrisome statistics about pathologic nodal staging of NSCLC: a median of five lymph nodes are examined in pN0 resection specimens (15); 12% of all resections (and 18% of all ‘node-negative resections’) have no lymph nodes examined (pathologic NX) (14-17); 12% of pN0 cases have no N1 lymph nodes examined (18); and 42% of resections (and 62% of ‘mediastinal lymph node negative’ cases) have no mediastinal lymph nodes examined (14,15,19,20).

Less than fastidious pathologic nodal staging has profound survival implications. For example, survival of patients with pN0 disease rises sequentially with the number of lymph nodes examined, until a maximal improvement is achieved at approximately 18-20 lymph nodes, suggesting the impact of sampling error when few lymph nodes are examined (21,22); patients who undergo pNX resections have a significantly worse survival than those with pN0, much more akin to the survival of those with pN1 disease (17); and failure to examine mediastinal lymph nodes is associated with a 14% survival deficit (20). Even in patients in whom lymph node metastasis is detected, examination of all available lymph nodes remains of prognostic value. As with many cancers, including colorectal, esophageal and gastric cancer, the prognosis of NSCLC worsens with increasing number of lymph node metastasis or a rising positive lymph node ratio (23-29). The number of N1 lymph node metastases is independently prognostic (30), but also correlates strongly with the likelihood of mediastinal lymph node metastasis (23).

Accurate pathologic lymph node staging involves three key processes: the intra-operative collection of the hilar (station 10) and mediastinal (stations 2-9) nodes; secure transfer, and accurate communication of the anatomic provenance, of all specimens between the operating room and pathology laboratory; and examination of all delivered specimens in the pathology laboratory, including the intrapulmonary lymph nodes (stations 11-14) retrieved by gross dissection of the lung resection specimen. The collection of hilar and mediastinal lymph nodes is the responsibility of the operating surgeon, without whose performance those specimens cannot be obtained; the extraction of intrapulmonary lymph nodes and the examination of all provided specimens is the responsibility of the pathologist; and the delivery of specimens in a secure, anatomically distinguishable fashion is the joint responsibility of the operating room and pathology teams.

Multiple efforts have been made to standardize the extent of the surgical lymph node harvest (31-35). Although the details differ slightly, it is generally agreed that a systematic collection of hilar and mediastinal lymph nodes should be attempted by the surgeon (Table 1). Some further advocates that the surgeon should collect stations 11 (interlobar) and 12 (lobar) lymph nodes (33). On the pathology side, standard recommendations call for examination of all lymph nodes in the resection specimen (36). While pathologists routinely examine all specifically identified specimens, retrieval of intrapulmonary nodes is dependent on the quality of the gross dissection of the resection specimen, warranting careful oversight of this aspect of the pathology examination.

Although techniques for gross dissection of intrapulmonary lymph nodes have been described, actual practice likely varies significantly, as evidenced by the fact that almost 50% of pNX cases are lobectomy or greater resections, suggesting that not only were hilar and mediastinal lymph nodes not provided from the operating room, but intrapulmonary lymph nodes were not retrieved during gross dissection of the resection specimen (17,37-40). More direct evidence comes from a study in which fastidious re-examination of discarded remnant lung resection specimens revealed a median of four N1 lymph nodes examined and a median of six discarded. Furthermore, 29% of patients in this study had discarded lymph nodes with metastasis, and 12% of pN0 specimens had discarded N1 lymph node metastasis (41).

The pathology examination ideally should indicate the anatomic source of each of the lymph nodes examined (lymph node mapping) in order to provide clinicians a clear idea of whether lymph nodes are from N1, N2 or N3...
stations. Because pathologists cannot identify the origin of lymph nodes provided by surgeons without accurate labeling, it is important for the communication between the operating room and the pathology laboratory to include unambiguous information on the source of all submitted specimens. The importance of this point is illustrated by an audit of mediastinal lymph node examination practices in a city-wide lung resection database, which revealed a 61% discordance between the procedure reported in operating surgeons’ notes and the procedure determined from objective review of the pathology report using pre-specified criteria. Whereas operating surgeons claimed a systematic nodal dissection in 45% of cases, only 8% met pathology criteria for systematic nodal dissection (42).

However, a blinded independent surgical review of the narrative description of the operation indicated that 30% of resections had adequately described a systematic nodal dissection (42). This sharp discordance between the operation narrative and the pathology report, which has been described as a ‘Tower of Babel’, suggests a multifaceted etiology of poor lymph node staging, encompassing actions both in the operating room and the pathology laboratory (43). This especially highlights the need for secure specimen delivery and better communication between the operating room and the pathology laboratory (43). A follow-up study in the same community institutions revealed marked improvement in concordance rate to 80% when surgeons used a lymph node specimen collection kit and checklist (45).

These observations suggest certain opportunities for intervention. The surgical hilar and mediastinal lymphadenectomy can be significantly improved with the use of pre-labelled surgical specimen collection kits, which help remind surgeons of the recommended lymph node collection procedure, provide a vehicle for the secure transfer of lymph node specimens, and with station-specific pre-labeling, eliminate all ambiguity about the anatomic source of each specimen. Use of such a specimen collection kit significantly improved hilar and mediastinal lymph node staging in a pilot study, with the ultimate result of an increase in the detection of pN2 disease from 8% of controls to 18% of cases (46). Routine use of kits such as these can address the operating room and communication aspects of the lymph node staging problem.

The finding of a high number of un-retrieved intrapulmonary lymph nodes has led to efforts to develop a more thorough standardized gross dissection method. Such a dissection protocol must be easy to learn, reproducible, quick to execute and feasible for use on fresh resection specimens in order not to interrupt the work flow in busy anatomic pathology laboratories. Such a protocol, in which blunt dissection of lymph nodes in the peri-bronchus is performed starting from the hilar surface of the resection specimen and working towards the periphery, has been shown to be feasible (47). This technique is easily taught, can be carried out on fresh specimens, requires a median of 9 minutes, and yields significantly more N1 lymph nodes than the current routine dissection protocol.

The combined use of the surgical specimen collection kit and thorough intrapulmonary lymph node retrieval protocols increased the number of lymph nodes examined in lung resection specimens from a median of 5 to 18, eliminated the pNX phenomenon, and, most importantly,

### Table 1 Minimum recommended surgical mediastinal lymph node staging quality parameters

<table>
<thead>
<tr>
<th>Tumor location</th>
<th>Guideline group and recommended surgical lymph node collection stations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACOSOG (33)</td>
</tr>
<tr>
<td>Right lung</td>
<td></td>
</tr>
<tr>
<td>Upper</td>
<td>2R, 4R, 7, 10R</td>
</tr>
<tr>
<td>Middle</td>
<td>Same</td>
</tr>
<tr>
<td>Lower</td>
<td>Same</td>
</tr>
<tr>
<td>Left lung</td>
<td></td>
</tr>
<tr>
<td>Upper</td>
<td>5, 6, 7, 10L</td>
</tr>
<tr>
<td>Lower</td>
<td>Same</td>
</tr>
</tbody>
</table>

*no nodal station specification. ACOSOG, American College of Surgeons Oncology Group; CoC, American College of Surgery Commission on Cancer; ESTS, European Society of Thoracic Surgeons; IASLC, International Association for the Study of Lung Cancer; NCCN, National Comprehensive Cancer Network; L, left; R, right.
increased the proportion of patients with detected node positive disease (and therefore potentially benefited by life-saving post-operative adjuvant therapy) from 30% to 45% (48). The potential survival impact of these combined interventions is large. Additionally, the improvement in lymph node mapping allows easy identification and correction of any errors in stage attribution.

In addition to identifying more patients with lymph node metastases, more lymph nodes with metastasis are found per patient with ‘node-positive disease’, potentially facilitating definitive examination of the prognostic impact of a higher number of lymph node metastasis in NSCLC. Ultimately, it is hoped that thorough lymph node retrieval will facilitate the search for other prognostic factors such as the real prevalence and prognostic value of micro-metastatic lymph node disease (detected by immunohistochemical analysis) and prognostic/predictive gene/protein expression patterns in primary tumors (49-54).

The combination of these two interventions, the surgical specimen collection kit and the standardized lung specimen dissection protocol, will be the subject of the ‘Strategies to Improve Lymph node Examination in Non-small cell lung cancer Trial’, an institutional randomization study with the acronym ‘SILENT’, which is currently in development. Objectives of this study are to test the impact of improved lymph node examination on stage distribution and survival, as well as the economic value associated with these corrective interventions.

Looking ahead, it is ultimately expected that molecular predictors of response to adjuvant therapy and independent molecular prognosticators can be identified from gene and protein expression profiles of primary tumors. However, optimal development and testing of such molecular markers will need accurately staged groups of patients (55,56). This will require marked improvement in the routine pathologic staging of resected lung cancer, to minimize the confounding of results caused by suboptimal use of the TNM staging system.

Major questions remain. How can we equitably compensate pathologists for any additional time, manpower, equipment and supply costs required to achieve more thorough examination? How can we successfully implement better pathology practices across the spectrum of practice environments? The first steps, possibly, are to universally acknowledge the existence of the gap in quality of pathologic staging, recognize the impact on survival, and commit to implementing corrective measures. Some measures, such as routine use of specimen collection kits, may be relatively easy to implement, while others might seem less so. Although improving the dissection and retrieval of intrapulmonary lymph nodes may require a bit more time and effort from pathologists, doing so will allow for more accurate identification of high-risk patients who will benefit from intensive post-operative intervention. This, in turn, is likely to provide a population-wide improvement in outcomes of resected early stage NSCLC.

Acknowledgements

Funding: Dr. Osarogiagbon was supported by a grant from the US National Institutes of Health: RO1 CA172253.

Footnote

Conflicts of Interest: Dr. Osarogiagbon has a patent application for a surgical specimen collection kit under review. All other authors have no conflicts of interest to declare.

References

34. CoC Quality of Care Measures. Available online: https://www.facs.org/quality%20programs/cancer/ncdb/