Thoracic oncology: early history of lung cancer staging

Between 1943 and 1952, Pierre Denoix, a surgeon at the Institute Gustave Roussy in Paris, developed, proposed and published a classification system for all solid tumors based on the size and extension of the primary tumor, its lymphatic involvement, and the presence/absence of metastases, and ever since, the resulting tumor-node-metastasis (TNM) staging system has been the international standard used by both the American Joint Commission on Cancer (AJCC) and the Union Internationale Contre le Cancer (UICC) (1). For lung cancer the proposed staging system was first published in the 1st edition of the UICC TNM pocket book, Le Livre de Poche, in 1968 and in the 1st edition of the AJCC Cancer Staging Manual in 1977. As a purely anatomically descriptive classification system, the development of accurate TNM descriptors and stage groupings depends on accurate and detailed patient clinical information and specific staging data. One of the earliest lung cancer staging champions was Clifford Mountain, who collected 2,155 lung cancer cases from MD Anderson Cancer Center in “computer-compatible format,” including 996 squamous cells, 521 adenocarcinomas, 195 undifferentiated large cell carcinomas, 368 small cell lung cancers (SCLC), and 75 without specified lung cancer cell type. Using this dataset, he analyzed 28 clinical variables and plotted over 300 survival curves; and subsequently proposed the first widely accepted lung cancer TNM staging system, which was accepted by the AJCC in 1973 and published in 1974 (2). He pointed out that such staging systems serve to aid in treatment planning, continuing self-assessment of outcomes, prognostication, and in the exchange of information between various centers. By 1997, Mountain had accumulated 4,351 primarily surgical lung cancer cases between 1975 and 1988, which were combined with 968 cases referred to MD Anderson from the National Cancer Institute’s cooperative Lung Cancer Study Group (LCSG) for confirmation of stage and histology (3). All but 66 were non-small cell lung cancers (NSCLC). The TNM descriptors and stage grouping were changed slightly to include definitions of satellite nodules as either T4 (within the primary-tumor lobe) or M1 [non-primary lobe(s) of the lung]. This iteration was accepted by as the AJCC and the UICC 5th edition of the TNM Classification of Malignant Tumors in 1997; however, there was no supporting data presented and the changes were simply included due to their inherent logic. With prior conflicting definitions and a perceived lack of clarity in the 4th TNM Classification System, oncologists then began to question the applicability of a lung cancer staging system devised from (I) almost exclusively surgical cases; (II) from essentially a single institution that perhaps was not entirely representative of the global problem; and (III) the apparent lack of validation, particularly with increasing data from other sources.
Initial international collaboration in lung cancer

With the shortcomings of a single institutional database being laid bare by the glaring deficiencies of the 5th edition of the TNM Classification of Malignant Tumors in 1997, the thoracic oncology community began to discuss alternatives. The International Association for the Study of Lung Cancer (IASLC) with its biennial meetings became the obvious group to lead the development of the next TNM staging system which was slated for 2007. After the IASLC invited and reviewed data presented from 20 separate databases throughout the world with approximately 80,000 patients in a meeting held in London in 2001, a partnership was developed between the IASLC and Cancer Research And Biostatistics (CRAB), a Seattle-based non-profit statistics and data management organization, with funding provided primarily by the global pharmaceutical industry. Although a plan for roughly 30,000 fully documented cases was adopted, over the ensuing years from 2001 to 2005, over 100,000 cases were submitted to the data center through international collaboration from 46 centers in over 19 countries, of which 81,495 were sufficient for analysis, including 68,463 cases of NSCLC and 13,032 cases of SCLC (4). Although cases were submitted from Europe (59%), North America (20%), Asia (13%), and Australia (8%), Africa and South America were not included and some large Asian countries, such as China and Russia were under-represented. Unfortunately, surgical cases still accounted for a disproportionate number (53% with surgery the only treatment modality in 41%) compared to the historical expected 20–25% of lung cancer cases. Furthermore, the routine use of positron emission tomography (PET) scans for the most accurate non-invasive staging was not widely employed and the biggest limitation was that the data was still collected retrospectively and many details were still missing. Despite these shortcomings, the IASLC lung cancer staging project demonstrated that even in a short period of time, broad international cooperation can provide amazingly useful information.

Mesothelioma staging: the early years

Malignant pleural mesothelioma (MPM) has always eluded ready staging descriptions and classification systems, a problem that predominantly emanates from its diffuse nature and the inherent difficulty in meaningful tumor imaging. In his original description of the use of pleuropneumonectomy in the treatment of diffuse MPM, Butchart and colleagues could only describe the preoperative imaging findings as a “pleural effusion” or “pleural thickening” noted on chest radiographs (5). Butchart’s staging system proposed in this publication identified four simple and relatively insightful stages of disease: (I) tumor confined to the ipsilateral pleura, lung, and pericardium; (II) tumor either invading mediastinal and/or mediastinal structures (e.g., esophagus, heart, and contralateral pleura) or tumor with thoracic lymph node involvement; (III) tumor either penetrating the diaphragm to directly involve the peritoneum or tumor with extra-thoracic lymph node involvement; and (IV) distant hematogenous metastases; however, it remained strictly a surgical-pathological system. Although relatively simplistic, the Butchart’s staging criteria were (and still are) associated with distinctly different prognostic groupings. Unlike lung cancer, mesothelioma staging does not allow easy applications of physical measurements as a basis for a T descriptor. Instead, Butchart created descriptions based on extent of invasion similar in concept to current T staging of gastrointestinal tumors. He chose essentially two simple T descriptors: (I) invasion limited to the ipsilateral pleura, lung, and pericardium (equivalent to T1) and (II) invasion into the chest wall and/or mediastinal structures (e.g., esophagus, heart, and contralateral pleura; equivalent to T2). This along with Butchart’s nodal staging, which consisted of no nodal disease (equivalent to N0), intra-thoracic lymph node involvement (equivalent to N1), and extra-thoracic lymph node involvement (equivalent to N2) remarkably parallels the current 8th edition AJCC and UICC TNM staging system just implemented on January 1, 2018.

Initial mesothelioma TNM staging system

In 1994 the International Mesothelioma Interest Group (IMIG) and the IASLC co-sponsored a workshop to analyze known staging and prognostic data in an effort to create a true TNM staging system for MPM that would be accepted by the AJCC and UICC staging committees (6). Evidence from reported surgical databases and available small clinical trials was scrutinized for data that could be applied to the radiographic, surgical, and pathologic staging of MPM. One of the key publications by Boutin and colleagues involved thoracoscopic assessment of the pleural space in patients with mostly early MPM (7). In this study, the French authors described their observations made during 188 thoracoscopic procedures done primarily for diagnostic purposes. Of the 188 procedures, 173 (92%) had a known
pleural effusion and only 13 (7%) had visible tumor on simple chest radiography (one had an empyema and one had a spontaneous pneumothorax). During thoracoscopy, distribution of pleural-based nodules was noted and recorded; furthermore, in each case, 10–20 large biopsies were obtained of the parietal pleural (the diaphragm, the parietal pleura, the costovertebral gutter, and the posterior costophrenic angle) to confirm the diagnosis but only “suspicious zones” of visceral pleura were actually ever biopsied. This led to speculation that the progression of mesothelioma was from parietal to visceral pleura. The data strongly suggested that patients without obvious visual involvement of the visceral pleura had a more favorable prognosis, but routine pathological involvement of the visceral pleura (similar to that performed for the parietal pleura) actually was never assessed. Nevertheless, this observation was incorporated formally into an IMIG staging system as T1a (parietal pleural involvement only) and T1b (visceral pleural involvement) without further validation and into the AJCC and UICC TNM staging systems in 1997.

In nearly 200 patients at UCLA we have found that regardless of the gross appearance of the visceral pleural, pathological involvement is diffuse and universal with no demonstrable areas of “normal” pleura except in two patients who had known prior fusion of the pleural space from extensive inflammation secondary to TB, pleurisy, etc. Thus, the distinction between visual visceral and parietal pleural involvement has not been confirmed pathologically and is not actually an objective measurement relevant to staging. The nature of the observations made by Boutin and colleagues were clearly not a distinctive pathological staging factor but likely a crude qualitative volumetric tumor measurement which Pass and colleagues as well as others have correlated with prognosis (8).

A second issue with the IMIG staging system was the N stage. Although nodal involvement was broadly accepted as a significant prognostic factor, the specific nodal staging that was adopted for mesothelioma essentially was borrowed from and identical to that of lung cancer. Yet the lymphatic drainage from a pleural-based disease (an extra-pulmonary disease) clearly is different from that of lung cancer, a pulmonary parenchymal disease. Subsequently, it was quickly appreciated that the relatively arbitrary distinction between hilar (N1) and mediastinal (N2) nodes in the IMIG mesothelioma staging system was clinically meretricious. Although data from small retrospective surgical case series were used to support the IMIG staging system, the number of T1 tumors was tiny, distinction between T1a and T1b tumors even in surgical case series was non-existent, and no reported data supported a sequential and progressive spread from lung parenchymal and hilar to mediastinal lymph nodes correlating with distinct prognostic groups (8,9). Consequently, almost immediately following its adoption by IMIG in 1996, widespread concerns were expressed regarding the validity of the IMIG staging system and its lack of true validation. A third staging system later proposed by Sugarbaker and colleagues (the Brigham staging system) was never widely accepted primarily because it departed substantially from standard TNM anatomic staging parameters (10). Furthermore, it relied heavily on surgical margins in patients undergoing extra-pleural pneumonectomy, which from an oncologic standpoint in mesothelioma are problematic and dubious at best and are not available in patients not undergoing surgery.

**IASLC-IMIG mesothelioma staging project**

Bolstered by the success of the IASLC lung cancer staging project, IMIG and IASLC agreed to embark on a similar project in mesothelioma. An international database was initiated for this purpose at the IASLC Workshop in London in 2009 utilizing data on 3,101 patients treated between 1995 and 2009 and submitted from 15 high volume centers on four continents with most (2,843; 91.7%) from the US (1,151; 37.1%), Italy (549; 17.7%), Australia (392; 12.6%), Turkey (236; 7.6%), Japan (180; 5.8%), Great Britain (177; 5.7%), and Switzerland (158; 5.1%) (11). Interestingly, 1,586 (51.1%) of patients had known or probable asbestos exposure but as many as 957 (30.9%) had no information despite receiving care in experienced high-volume mesothelioma centers. The database was almost exclusively a surgical database with 2,958 (95.4%) of patients undergoing surgery. Of the 3,101 patients, a total of 785 (25.3%) were excluded due to inadequate staging data. The remaining 2,316 (74.7%) patients were combined to provide “best” staging in accordance with IMIG, AJCC, and UICC guidelines, and this information was used to generate survival curves. A relatively large number 729 (31.5%) of patients underwent palliative surgery only and 90 (3.9%) had no data or did not undergo surgery at all. A relatively paltry number of patients by international database standards (1,056; 34.1% of all patients submitted) had simultaneously both clinical and pathological staging information available. The inaccuracy of the prior clinical IMIG staging system particularly in early stage disease was demonstrated by the high rate of surgical up-staging (80%,
65%, and 22% in clinical stages I, II, and III, respectively. Clinical over-staging was a smaller but still significant issue with approximately 3%, 12%, and 12% of patients ultimately down-staged by surgery in clinical stages II, III, and IV, respectively.

The analysis and recommendations for changes to the T-, N-, and M-stage criteria were published separately (11-14). For T-stage, there was separation between clinical T1-4 disease but no difference was noted between T1a and T1b (12). Pathologically, however, there was no clear separation to any of the T1-3 survival curves, leaving a significant difference only between the T3 and T4 groups. This finding appears anatomically logical as there is no real rationale to assign more significance to any the individual T1-3 descriptor (pleura, superficial diaphragm, lung or pericardium, endothoracic fascia, mediastinal fat, and limited resectable chest wall/intercostal muscle). It also is remarkably reminiscent of the simplicity of T descriptors in Butchart’s Staging System. Interestingly, the most promising anatomic distinction that may prove useful in future staging system revisions is tumor thickness, which was available in 472 M0 patients, and/or tumor volume. When analyzing the mean sum of lower, middle, and upper pleural thickness measurements, survival by classification into quartiles decreased from the lowest (median survival = 23.4 months with <16.0 mm) to the highest quartile of pleural thickness (median survival = 13.2 months with >50.0 mm; P = 0.005 by log-rank test) (12). Increasing thickness sums with two cut-points (13 and 60 mm total pleural thickness) also was significantly associated with cT categories (P < 0.0001), node positivity (P < 0.0001), and overall stage (P < 0.0001) by a chi-square test of association (12). Even with a single cut-point (5.1 mm) significant survival differences were also observed (24.2 months < 5.1 mm versus 17.7 months > 5.1 mm; P < 0.0014) (12). The pattern of pleural spread also correlated somewhat with prognosis with minimal pleural involvement more favorable when compared to nodular and rind-like patterns (23.4 versus 18.2 and 14.5 months, respectively; P < 0.004) (12). The use of pleural and volumetric measurements in mesothelioma as potential criteria for T-stage over the current histological invasion approach is supported by a number of other studies and investigators as well (15,16).

The N-stage mesothelioma descriptors were analyzed using the IMIG-IASLC database which included 1,322 cases with clinical staging descriptions and 851 cases with pathological staging information (11,13). The findings revealed survival differences between patients with node-negative (pN0) versus node-positive (pN1 or pN2) disease but no difference between pN1 and pN2 (11,13). Furthermore, analysis of a small subset of 181 patients with data available showed no difference in survival according to the number of metastatic nodes (11). This resulted in perhaps the most significant change in mesothelioma staging: that is, collapse of the N1 and N2 groups into a single N1 designation. This certainly matches with clinical experiences of many thoracic surgeons. The corresponding change of N3 lymph nodes to N2 follows logically, but still has little, if any, data to support this continued separate designation and only future large datasets will elucidate the significance of these nodal stations.

In parallel to and as an integral part of this staging revision effort for the 8th edition of the AJCC and UICC TNM staging system, an attempt was made by some IMIG surgeons led by Rice and colleagues to standardize terminology applied to various surgical procedures performed under the umbrella of “pleurectomy and decortication” or P/D, due to the increasing use of lung-sparing surgery (17). A web-based questionnaire was sent to 130 surgeons from 59 centers worldwide over a 3-week period from October 11 through October 29, 2010. Surgeons responding totaled 62 (47.7%) and were affiliated with 39 different medical centers in 14 countries on four continents (Europe 27/43.5%; North America 26/41.9%; Asia 8/12.9%; and Australia 1/1.6%). Most responding surgeons (59; 95.2%) felt that there was a need to refine surgical nomenclature to account for the procedural differences between P/D for palliation and P/D with the stated goal of attaining an R1 resection. The responses clearly demonstrated that surgeons across the globe used a wide array of terminology to describe an almost equally broad selection of surgical procedures. For instance, 42 of 58 respondents (72.4%) considered the term pleurectomy and decortication or “P/D” to imply an R1 resection; yet, 15 surgeons (25.9%) considered “P/D” to be subtotal removal of the parietal and visceral tumor for palliation. Of surgeons performing P/D with the goal of an R1 resection, 23 surgeons (40%) preferred the term “P/D,” whereas 22 (38%) preferred “total pleurectomy”. If diaphragm and/or pericardium was removed, most surgeons (64%) used the term: “radical P/D”. The results of this exercise were to prove that there were truly no uniform definitions of pleural procedures. In the final “consensus” statement, the IASLC Mesothelioma Domain “suggested” the use of the term “extended P/D” rather than “radical P/D” since it was felt that the latter implied a “completeness of resection with...
added therapeutic benefit” that the Domain felt was not justified. This action of the IASLC Mesothelioma Domain is based on no or certainly unclear data and principles, resulting in a “consensus” by dictum. Similarly applying this approach and conclusion to breast cancer, a mastectomy which includes the pectoralis muscle would be termed an “extended mastectomy” not a “radical mastectomy” because of an obvious lack of efficacy. Further confusion in staging terminology results from the use of the term “macroscopic complete resection” instead of the broadly-accepted TNM residual tumor designation of microscopic residual tumor, i.e., “R1” resection. Perhaps this reflects surgeons’ desire to focus on what was accomplished and not what was left behind; but clearly much of the confusion emanates from a lack of consistency within our own oncology and surgical communities.

**Staging deficiencies**

Clearly the IASLC staging projects for both lung cancer and mesothelioma have produced improvements in the understanding of staging and outcomes; however, significant deficiencies do exist, already resulting in challenges (18). First, the retrospective nature of the data is a huge limitation. Retrospective data is an excellent starting point, but now we have to move on to prospective data. Not to be left out of the data-driven staging revolution, the International Thymic Malignancies Interest Group (ITMIG) led by Frank Detterbeck similarly partnered with IASLC (and CRAB) to create a Thymic Malignancies Domain of the Staging and Prognostic Factors Committee (SPFC-TD) in 2010 to guide future staging revisions (18). This project included a relatively brief period of retrospective data collection (amassing 10,808 cases from 105 centers in 16 countries) but then quickly launched into detailed prospective data collection with an accompanying tissue bank and a prospective imaging repository. Unlike the mesothelioma and lung cancer staging efforts, ITMIG sought to include all interested centers internationally and did not limit the data collection to specific elite large “experienced” centers. In an orphan disease, this was necessary in order to accumulate adequate numbers of cases but undoubtedly also will lead to a more robust staging system that will be applicable worldwide.

Secondly, clinical staging is a significant deficiency. All of the staging projects have been dominated by surgical data which are available only following surgical resection and cannot be universally applied prior to initiation of therapy. Functional imaging (e.g., PET), magnetic resonance imaging, as well as volumetric and other novel imaging measurements need to be actively evaluated and incorporated in the process to improve clinical staging and decrease the differences between clinical and pathological staging (18,19). Thirdly, biases have not been eliminated. This is most evidence in N staging. Both mesothelioma and thymic tumors have much less clear nodal progression than lung cancer and any designation beyond nodal positivity versus negativity has not support by data in either of these diseases. Therefore, the inclusion of N2 designations in both these disease is purely based on speculation and serves no function. Perhaps in an effort to match the detail in the lung cancer staging system, mesothelioma staging, in particular, has remained more granular than is justifiable. Perhaps the limited T stages and N stages of Butchart’s Staging System currently are closer to the mark. Others also have proposed much simpler mesothelioma staging systems as having more clinical relevance (20-22). Finally, TNM staging remains strictly an anatomic system—in 2018, this rapidly is becoming a liability. To date, the AJCC and UICC have resisted efforts to include genetic and mutational information in staging but at some point, this “elephant in the room” will have to be constructively addressed to keep TNM staging relevant. With extensive analysis of tissue and “liquid” biopsies, future therapy may be more uniformly systemic in nature even for early cancers and dependent on “genetic” parameters rather than purely anatomic findings.

**Acknowledgements**

None.

**Footnote**

*Conflicts of Interest:* The author has no conflicts of interest to declare.

**References**

4. Goldstraw P. The 7th Edition of the TNM Classification
18. Abdel-Rahman O. Challenging a dogma; AJCC 8th staging system is not sufficient to predict outcomes of patients with malignant pleural mesothelioma. Lung Cancer 2017;113:128-33.

Cite this article as: Cameron RB. Staging in the era of international databases: documented improvements with remaining challenges. J Thorac Dis 2018;10(2):682-687. doi: 10.21037/jtd.2018.01.23