Classifying Clinical Severity to Help Solve Problems of Stage Migration in Nonconcurrent Comparisons of Lung Cancer Therapy

David G. Pfister, Carolyn K. Wells, Charles K. Chan, and Alvan R. Feinstein

ABSTRACT

To compare the effects of stage migration in the "traditional" 3-stage TNM (tumor, node, metastasis) system with those in a new "expanded" 5-stage system, which has two additional stages for the poor prognostic groups, we used both systems to classify a cohort of 178 patients with primary lung cancer. To check for migrations, the stages in both systems were first assigned using only "old" technological information and were then reassigned using all the available "new" as well as old technological data. Although the 5-stage system had more migrations than the 3-stage system, survival rates were relatively unaffected for patients in the two new stages with poor prognosis. In both TNM staging patterns, the effects of stage migration on survival statistics were most impressive in the prognostically better (TNM I and II) stages. A solution to the migration problem is offered by the "clinical severity" (CS) staging system. Like the expanded TNM system, the CS system has 5 stages and a sharp prognostic gradient among stages. The CS system, however, had fewer technology-induced stage migrations than either TNM system, and the migrations had no substantial impact on stage-specific survival results. The excellent prognostic discrimination and secular stability of the CS system make it superior to the TNM system for comparing treatment results from different eras, especially for patients with stage I and II disease.

INTRODUCTION

Patients with cancer are usually "staged" prognostically according to morphological evidence of the tumor's anatomic dissemination expressed in stages of various TNM systems (1). Although the stages demarcate useful prognostic gradients (2) that facilitate comparisons of concurrent patient groups, the prognostic value of the TNM system is impaired by a problem called stage migration (3) when results are compared for patients from different eras. Although misleading, the nonconcurrent comparisons are often used when practicing clinicians recall past experience in making current choices of treatment and when health policy makers and funding agencies evaluate the impact of changes in therapy over time.

Stage migration can occur whenever the diagnostic data used to assign disease stages are different for the compared patient groups. The situation most commonly arises when advances in diagnostic technology can be applied only to relatively recent groups of patients. Because the sensitive new methods can identify "silent" or "subclinical" metastatic lesions that would previously have been undetected, the new technological data allow recent patients with silent metastases to "migrate" or "shift" from localized stages with generally good prognoses (such as TNM I) into advanced stages with generally worse prognoses (such as TNM III).

For example, consider a patient with a small primary lung cancer and asymptomatic liver metastases. According to history, physical examination, and conventional roentgenograms (the main diagnostic evidence available in earlier eras), the patient would have been categorized as having stage I or localized disease. If a liver-spleen scan is performed, however, the demonstration of the clinically unsuspected liver metastases would shift the patient to a different stage, such as III, for distantly metastatic disease.

Although the overall survival rate for a cohort would be unaffected, this type of stage migration would elevate the survival rates in each of the constituent stages (3). Survival becomes better in the lower ("good") stages, which would contain fewer patients with metastatic disease. Survival is also better in the higher ("bad") stages, because the metastases in the transferred patients are silent rather than overt. Although first described and quantified in a cohort of lung cancer patients, stage migration has been shown in other research to pertain to many other cancers classified with morphological staging systems (4-8).

PROBLEMS UNDER STUDY

The impetus for the current study occurred when the American Joint Committee on Cancer and the Union Internationale Contre Cancer proposed that the old TNM system for lung cancer be expanded from 3 to 5 prognostic stages. As shown in Table 1, stages I and II of this new system (9) are similar to their counterparts in the conventional system. The new stages IIIA, IIIB, and IV, however, represent subdivisions of stage III from the old system. Although a finer prognostic gradient is demarcated, this new 5-stage system should theoretically also increase problems of stage migration, since more morphological stages offer more opportunities for migration. The first purpose of the current research, therefore, was to determine what effect the increased number of TNM stages would have on stage migration rates and survival statistics.

A second, more prominent purpose of the research was to consider methods of reducing the problems of stage migration, while maintaining a strong prognostic gradient in the staging system. Intuitively, a staging system based on clinical attributes of the patients would seem particularly desirable, because the clinical manifestations of disease can generally be identified and classified without intensive technological data and should, therefore, be relatively resistant to stage migration.

In the original description of the stage-migration phenomenon, the alternative approach proposed by the authors was a 4-stage clinical prognostic system based on the asymptomatic, primary, systemic, or metastatic symptom patterns present at
first treatment (3). In the current report, we have used an expanded CS system, which includes symptom patterns but also incorporates information on severity of illness and comorbidity of coexisting disease (10). In two previous cohorts, the 5-stage CS system produced prognostic gradients at least as strong as those found with the new 5-stage TNM system (10). The CS system would, therefore, seem to be a preferable alternative to the TNM system for making nonconcurrent comparisons of lung cancer patients.

In this report, we verify that both the 5-stage TNM and the CS systems demarcate important prognostic gradients. We also show the increased stage migration and associated statistical consequences produced by the new 5-stage TNM system. Finally, we demonstrate how application of the CS system can help solve these difficulties.

METHODS

The study population included patients with microscopic evidence of lung cancer, having a zero time at Yale-New Haven Hospital from July 1, 1981, through December 31, 1982. Zero time was defined as the date of first antineoplastic therapy for lung cancer or the date of the decision not to give therapy. Antineoplastic therapy was defined as surgery, radiotherapy, or chemotherpay. Potential study subjects were identified from the Yale-New Haven Hospital discharge diagnosis logbook. All histological types of lung cancer were included.

Among 201 patients whose zero time occurred during the selected 18-month interval, 23 were excluded from analysis: 7, because the medical record could not be found; 9, because of no tissue or cytological confirmatory evidence; 5, in whom zero time therapy occurred elsewhere; and 2, who had been followed for less than 6 months after zero time. Thus, the cohort under analysis here contains 178 patients.

The data were extracted from patients' medical records, according to methods previously described (11). Although the prerequisite microscopic evidence of cancer could have been obtained before or after zero time (e.g., at diagnostic lymph node biopsy, thoracotomy, necropsy), the zero time staging of morphological and clinical information was done exclusively with data that had been acquired before zero time. The coding of the morphological and clinical information extracted from the medical records and the assignment of TNM and CS stages were supervised and reviewed (as needed) by a second observer unaware of the patient's post-zero time outcome.

The TNM stages for both the 3- and 5-stage systems were assigned with two different sets of data using established criteria (2, 9). First, an "old" stage was assigned, using old morphological data, which included evidence available before the modern "revelation" in imaging and other diagnostic technology. The old evidence came from physical examination, bronchoscopy and other endoscopic procedures, common biopsy procedures (liver, lymph node biopsy), and conventional roentgenograms (plain films, ordinary tomography, films made with injected or ingested contrast media). The same patients were then assigned a "new" stage, which depended on all the available data, including new morphological information provided by radionuclide scans, computerized tomography, ultrasound, and relatively new biopsy procedures, such as mediastinoscopic biopsy before thoracotomy, or routine bone marrow biopsy at diagnosis in patients with small cell carcinoma.

Within each 3-stage or 5-stage TNM system, a stage migration was defined as occurring when the new-data stage differed from the old-data stage. The results of old technology were always included as old-stage data, even though the performance on the old test may have been prompted by a new-technology result. For example, the results of a positive liver biopsy were regarded as old-stage data, although the biopsy might not have been performed without the suggestion of metastases provided by a liver-scan scan. This approach, which was used in previous research (3), avoids invidious retrospective decisions about the reasons why certain tests were ordered and provides a more reproducible, albeit slightly lower, estimate of the number of stage migrations.

Equivocal scans occur commonly in the diagnostic evaluation of cancer patients (3, 12), and their classification is a source of potential ambiguity in the assignment of TNM stages. For the current analysis, equivocal scans were approached in two different ways to provide an estimated range for the migration effect. In the first approach, all equivocal scans were regarded as positive, because various oncologic colleagues had stated that lung cancer patients with equivocal scan results are commonly regarded as having metastatic disease. In the second approach, which would provide a conservative "lower limit" estimate of the migration effect, equivocal scans were regarded as negative unless they were clearly regarded as positive when the patient's treatment was selected. (For example, an equivocal scan result that led to cancellation of thoracotomy in someone with otherwise resectable disease would have been regarded as positive.) With the second approach, most equivocal scans were classified as negative.

The contents and criteria for the clinical severity stages are summarized in Table 2. Stage assignment depends on three main variables: the pattern or type of symptoms, classified hierarchically as none, pulmonic or systemic only, regional or mediastinal, or distant; the severity of symptoms, with progressive levels: none, major weight loss, severe weight loss or severe dyspnea, and severe tumor effects; and prognostic comorbidity classified as absent or present. Categories of the three component variables are grouped to create 5 stages, labeled A (with excellent prognosis) through E (worst prognosis). Patients are assigned to a CS stage according to the most severe clinical manifestation present at zero time. More complete, formal criteria for the CS system have been described elsewhere (10).

Although less affected than morphological data, clinical assessments can nevertheless be altered by new technology, so that migrations can occur within the CS stages. For example, a lung cancer patient with chest pain and no evidence of bone disease on physical examination or conventional roentgenograms would be classified as having primary symptoms only (CS stage B). If rib metastases are shown on a bone scan, however, prompting local radiotherapy to the involved region, the new technological data would make the patient's stage migrate into the regional symptom group (CS stage C).

In contrast to the unilateral direction of migrations from good to bad stages in the TNM system, the CS migrations can go in both directions. A reverse migration, from a worse to better prognostic stage, occurs when a new technology result (such as a negative CT scan of the head) removes the metastatic symptom classification for a suspicious, potentially metastatic manifestation (such as a new, persistent headache).

To determine the number and impact of such migrations on the CS system, the CS stages for each patient were also assigned both with and without new technological data. Because the attribution of symptoms, severity of illness, and comorbidity all depend on clinical decisions made with the available pre-zero time information, equivocal scans were classified only one way, i.e., as positive or negative according to the associated clinical decisions in diagnosis or therapy.

Each patient's zero time TNM and CS stage were correlated with the patient's subsequent outcome irrespective of therapy or histological type. Mortality information for each patient was obtained from the Yale-New Haven Tumor Registry. All deaths were counted, regardless

<table>
<thead>
<tr>
<th>Table 1 TNM staging systems for lung cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional 3-stage</td>
</tr>
<tr>
<td>I  Localized</td>
</tr>
<tr>
<td>II Regional</td>
</tr>
<tr>
<td>III Locally extensive, distant</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Distant</td>
</tr>
</tbody>
</table>

* Patients with T1N1M0 disease are included in stage I of the 3-stage system and in stage II of the 5-stage system.
of cause. Follow-up for 3 years or more was available for all but 3 patients in the cohort. The survival times for these 3 patients, up to their date of loss to follow-up, were included in our analysis. Six-month survival status was chosen as the primary outcome measure because the numerators and denominators can indicate the number and survival proportion of patients in each stage, while clearly identifying both the patients who migrated between stages and the statistical consequences of the migrations. Where appropriate, median survivals were also determined.

RESULTS

The study population of 178 patients had a median age of 64 years, had more men (58%) than women, and was predominantly white (93%). The cell types of the various lung cancers were apportioned as: squamous cell, 28%; adenocarcinoma (includes bronchoalveolar), 31%; large cell undifferentiated, 23%; small cell (includes oat cell), 16%; and others, 2%. In more than 50% of patients, disease had spread beyond the ipsilateral regional (hilar) nodes at zero time. This distribution of histological types and morphological extensiveness was similar to what has been reported from other centers (13–17), thus indicating no disparities in composition of the cohort that would cause quantitatively important distinctions in the morphological or clinical analyses.

New diagnostic imaging and biopsy procedures were commonly used in the cohort. A total of 352 new-data tests, not counting repeats of the same test in a given patient, were performed in the 178 patients. Proportions of tested patients were particularly high (61 and 46%, respectively) for radionuclide scans of the liver-spleen and bone. The number of equivocal readings was substantial (37 overall), relative to the number of positive readings (104 overall).

The subsequent analysis has two main parts. The first part shows the mechanism of stage migration and its survival effects when the number of TNM morphological stages is expanded from 3 to 5. The second part compares the migration effects and survival results for the CS and 5-stage TNM systems. Although patients with small cell and non-small cell histological types are often separated in lung cancer research, all cell types are included in the analyses here. The results were similar when the 16% of patients with small cell carcinoma are excluded.

Stage Migration in TNM Systems. Tables 3 and 4 demonstrate the mechanism and consequences of stage migration in the 3- and 5-stage TNM systems. As in previous research (3), equivocal results were classified as positive in this part of the analysis. Within the traditional 3-stage TNM system (Table 3), classifications with new data would make 16 and 10 patients, respectively, migrate from stages I and II into stage III, for a total of 26 migrations (15%). (When most equivocal scans were classified as negative, the rate of migration was 8%).
The migratory shifts in Table 3 increase the survival rates in stage I from 83 to 89%, in stage II from 61 to 62%, and in stage III from 39 to 44%. The survival rates improved in stages I and II because patients with relatively poorer prognoses (6-month survival rates of 63 and 60%, respectively) moved into stage III. The survival rate increased in stage III because the "immigrant" patients were prognostically more favorable than the group they joined. None of these changes within stages would affect the total survival rate, which remains at 57% (102 of 178 patients) in all 3 columns of the table.

A similar analysis, when applied to the new 5-stage TNM system (Table 4), shows many more migrations. Eighteen patients with stage IIIA or IIIB disease, who previously could not migrate when classified as stage III in the 3-stage system, were now able to shift. With this change, the total number of shifts increased from 26 in the 3-stage system to 44 (25%) in the 5-stage system. (When most equivocal scans were classified as negative, the migration rate was 19% in the 5-stage system.)

The survival consequences of the shifts in Table 4 are similar in mechanism and direction to those previously described for the 3-stage TNM system. Specifically, the 6-month survival rates increased in TNM stage I from 83 to 89%, in II from 62 to 67%; in IIIa from 52 to 57%; in IIIb from 37 to 45%, and in IV from 33 to 40%.

To provide more complete quantitative information on the potential impact of these stage migrations on lung cancer survival statistics for the 5-stage TNM system, the changes in 6-month rates and median survival times are summarized in Table 5. The table shows the migration-induced changes when most equivocal scans are classified as negative (middle) and when all equivocal scans are classified as positive (far right). The survival changes for the 3-stage system were similar in direction and magnitude to those described in the 5-stage system but are not formally presented in an analogous table here.

Table 5 shows that the stage migrations led to improved 6-month survival rates and/or median survival times in all stages, but the changes in median survival were quantitatively most impressive in stages I and II. When all equivocal scans were classified as positive, the median survivals increased in stage I from 31.6 to 46.9 months and in stage II from 11.3 to 18.9 months. When most equivocal scans were classified as negative, the changes in median survivals were smaller but still distinctive. Regardless of how equivocal scans were classified, however, the change in median survival was minor in stages IIIa and IIIb and nonexistent in stage IV.

For the 3-stage rather than 5-stage TNM system, these same trends in median survivals were found (with data not shown here) in the better (stage I and II) and worst (stage III) prognostic groups.

**CS versus 5-Stage TNM Systems.** For the 6-month survival rates and median survival times of the CS system, Table 6 shows a distinctive strong prognostic gradient, similar to what was previously noted for the 5-stage TNM system. From best to worst CS stages, respectively, the survival rates extend from 91 to 22% and the median survival times from 30.6 to 1.9 months.

The stage-migration effect of new technology on the CS is demonstrated in Table 7. The number of migrations was much smaller [10/178 (6%)] than what occurred with the 5-stage TNM system. Stages A and E, the extreme prognostic groups, were unaffected by any migrations. Six of the 10 shifts traversed only one stage, and the remaining 4 traversed only 2. Although TNM migrations always went from better to worse prognostic stages, 7 of the 10 stage shifts in the CS system went in the opposite prognostic direction. Because these reverse migrants had a better prognosis than other patients remaining in the original stage, the countermigration tends to offset the migration effect of patients who shifted from better to worse prognostic groups. Consequently, as shown in Tables 7 and 8, the CS migrations had a relatively small effect on the stage-specific survival statistics. The distinction was particularly striking when the effects of migration were compared in the better prognostic stages, A and B of the CS system versus I and II of the TNM system.

**DISCUSSION**

These results offer further evidence that changing diagnostic technology and criteria will create problems of stage migration in cancer-staging systems. Because the migration-induced elevations in stage-specific survival rates can easily be mistaken for real improvements, the results may lead to inappropriate conclusions about the benefits of certain "modern" therapies. The issue of stage migration becomes especially important in view of contentions, in both the medical (18) and lay (19) press, that the improved results claimed for modern cancer therapy may arise from statistical artifacts.

The described limitations of the TNM and the relative superiority of the CS system are pertinent only when results of treatment are compared for patients from different eras. For concurrent comparison, the TNM system maintains its prognostic excellence, since stage migration will not occur unless the concurrent results are compared for nations or regions with major disparities in the availability of advanced-technology data. Nevertheless, even when TNM staging is unaccompanied by problems in stage migration, CS stages can be used to augment TNM stages, demonstrating important prognostic distinctions that cannot be discerned with morphology alone (10, 20). This additional prognostic information can help clinicians make better therapeutic decisions, particularly regarding the role of aggressive therapy in patients with notably poor prognoses.

<table>
<thead>
<tr>
<th>TNM stage</th>
<th>6-mo survival rate (%)</th>
<th>Median survival (mo)</th>
<th>New-data classification</th>
<th>6-mo survival rate (%)</th>
<th>Median survival (mo)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>60</td>
<td>83</td>
<td>31.6</td>
<td>53</td>
<td>87</td>
</tr>
<tr>
<td>II</td>
<td>26</td>
<td>62</td>
<td>11.3</td>
<td>18</td>
<td>67</td>
</tr>
<tr>
<td>IIIa</td>
<td>25</td>
<td>52</td>
<td>6.7</td>
<td>25</td>
<td>56</td>
</tr>
<tr>
<td>IIIb</td>
<td>27</td>
<td>37</td>
<td>5.0</td>
<td>24</td>
<td>42</td>
</tr>
<tr>
<td>IV</td>
<td>40</td>
<td>33</td>
<td>4.0</td>
<td>58</td>
<td>34</td>
</tr>
</tbody>
</table>

*In the middle columns, most equivocal scans are classified as negative; in the far right columns, all equivocal scans are classified as positive.*
The statistical variables may be either missing or affected by the categories of staging systems and by resorting, instead, to This approach would use clinical information while maintaining the consistency of the human observers. and have also been validated by other observers (31, 32) during problems that have already been demonstrated for observer variability in the interpretation of hard data for histopathology and function evaluation, and rigid bronchoscopy. As diagnostic technology continues to advance and become more sensitive, the ability of any prognostic system to resist these changes depends on whether the component data come from methods that are intrinsically stable over time. If the available technological information cannot be equally applied for all the compared groups, the effects of migration will distort the comparison, although new technological information can be effectively applied if the comparisons are concurrent. Because a resolution of these problems will require further research, the best approach now may be to assign CS stages routinely so that nonconcurrent comparisons can be more reliable in the future. It is not surprising that the statistical artifacts caused by stage migration are more striking in the better prognostic stages of both the 3- and 5-stage TNM systems. Patients with metastatic lung cancer, whether silent or overt at zero time, usually have substantially worse prognoses than patients with limited, surgically resectable disease. Migration in the TNM system is, therefore, likely to produce greater improvements in the good prognostic stages that the patients leave than in the poor prognostic stages that they enter. This same rationale would explain why the migration effect is quantitatively modest for the expanded poor stages in the 5-stage TNM system. As subdivisions of the old stage III disease, the new groups IIIa, IIIb, and IV all have a relatively poor prognosis at baseline. Thus, for nonconcurrent comparisons of lung cancer patients in poor morphological stages, the effects of stage migration are less important than in the good stages. This distinction may not apply, however, to other cancers, in locations such as breast, where metastatic disease is not as rapidly fatal as in lung cancer and where stage migration may produce dramatic effects on survival statistics in the more advanced morphological stages. Data from other investigators, using different research methods, confirm our impression about a substantial “reservoir” of potential migrant cases. In two studies of autopsies (22, 23) in patients who had died within 30 days of a “curative” resection for lung cancer, the investigators found that about 25% had formerly undetected distant metastases. Subsequent radionuclide scans (11) showed definite abnormalities attributable to the spread of cancer in 6 (11%) of 55 patients whose lung cancer was still regarded as surgically resectable after clinical examination, chest roentgenogram, liver function tests, pulmonary function evaluation, and rigid bronchoscopy. As diagnostic technology continues to advance and become more sensitive, more patients with asymptomatic micrometastases will be identified, leading to more stage migration in the future. Fear of observer variability is a separate problem that inhibits the application of staging systems that use “soft” clinical rather than “hard” morphological data. This fear is particularly likely to arise if the investigators are unaware of the major unresolved problems that have already been demonstrated for observer variability in the interpretation of hard data for histopathology (24, 25), roentgenograms (26, 27), and radionuclide scans (28–30). Like all classifications that involve human perception and judgment, the CS system requires careful attention to its specifications and criteria. The clinical staging criteria for this system have been developed and standardized by ourselves (10) and have also been validated by other observers (31, 32) during the past 20 years. Although the question of reproducibility is important for all forms of cancer staging, morphological and nonmorphological, the main issues in the problem of stage migration arise from the technological sources of data, not from the consistency of the human observers.
ACKNOWLEDGMENTS

The authors thank Elizabeth Pesapane and Mary Lyman for preparing the manuscript, Roberta Lawrence of the Yale-New Haven Hospital Tumor Registry for supplying mortality data, and the staff of the Medical Records Department of Yale-New Haven Hospital for their technical assistance.

REFERENCES


Classifying Clinical Severity to Help Solve Problems of Stage Migration in Nonconcurrent Comparisons of Lung Cancer Therapy


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/50/15/4664

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.