Objective: Cancer staging is dynamic, reflecting accrual of knowledge and experience in treatment. The objectives of this study were to assess current esophageal cancer staging and to determine whether refinements of classification and stage grouping are necessary.

Methods: From 1983 through November 2000, 480 patients underwent esophagectomy without induction therapy. Depth of tumor invasion (T), regional lymph node status (N), distant status (M), number of metastatic regional lymph nodes, and histopathologic type and grade were subjected to survival-tree analysis, multivariable Cox and hazard function analysis, and residual misclassification risk analysis.

Results: Inhomogeneity of survival was found within and lack of distinction was found between current American Joint Committee on Cancer staging groups, supporting the need for refinement. T1 and N1 were redefined on the basis of survival differences. T1a is intramucosal cancer, T1b is submucosal cancer (P = .008), N1 is 1 or 2 metastatic regional lymph nodes, and N2 is 3 or more metastatic regional lymph nodes (P = .01). Current subclassification of M1 is not warranted (P = .9). Histopathologic type (P = .17) and grade (P = .3) minimally refined staging. Reassignment of staging groups constrained by American Joint Committee on Cancer definitions of stages 0 and IV produced less monotonic, distinctive, and homogeneous survival than free assignment of staging groups.

Conclusions: Current American Joint Committee on Cancer staging of esophageal cancer is inadequate. Refinement requires redefinition of T1, N1, and M1 classifications. Stage grouping within the constraints of American Joint Committee on Cancer definitions produces less accurate prognosis than free assignment based on survival data.
The AJCC Cancer Staging Handbook notes that “it is essential to recognize the kinetics of change of staging systems.” TNM classifications, additional factors, and stage groupings should be dynamic, reflecting increasing knowledge of cancer, its treatment, and its prognosis. Although TNM staging of esophageal cancer was updated in 2002, experience with these classifications and stage groupings indicates that review and refinement are needed. Therefore the purposes of this study were to determine whether the present American Joint Committee on Cancer (AJCC) staging of esophageal cancer adequately reflects modern prognosis and, if not, to refine staging of esophageal cancer by TNM reclassification, addition of other prognostic factors, or stage regrouping.

Patients and Methods
Patients
From January 1983 to December 2000, 480 patients with esophageal cancer who underwent resection without induction therapy were identified in the prospective esophageal surgery registry at The Cleveland Clinic Foundation. The Clinic’s Institutional Review Board has approved research based on data contained in this registry. Cross-sectional follow-up has been performed yearly. Follow-up was incomplete in 2 patients last contacted in 1998, 5 and 12 years after surgical intervention. Median follow-up of all patients was 1.5 years, with 25% followed for more than 3.5 years. Among 172 patients alive at last follow-up, median follow-up was 3.4 years, with 25% followed for more than 6.9 years. All-cause mortality (including operative) from the time of esophagectomy was used to calculate survival estimates.

Staging
Current staging of esophageal cancer, modified for clarity, is outlined in Table 1.

<table>
<thead>
<tr>
<th>TABLE 1. Modified current TNM staging of esophageal cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TNM classifications</strong></td>
</tr>
<tr>
<td><strong>T</strong></td>
</tr>
<tr>
<td>Tis High-grade dysplasia*</td>
</tr>
<tr>
<td>T1 Tumor invading lamina propria, muscularis mucosae, or submucosa†</td>
</tr>
<tr>
<td>T2 Tumor invading muscularis propria</td>
</tr>
<tr>
<td>T3 Tumor invading periesophageal tissue</td>
</tr>
<tr>
<td>T4 Tumor invading adjacent structures</td>
</tr>
<tr>
<td><strong>N</strong></td>
</tr>
<tr>
<td>N0 No regional lymph node metastases</td>
</tr>
<tr>
<td>N1 Regional lymph node metastases</td>
</tr>
<tr>
<td>M*</td>
</tr>
<tr>
<td>M0 No distant metastases</td>
</tr>
<tr>
<td>M1a Upper thoracic tumors metastatic to cervical nodes</td>
</tr>
<tr>
<td>M1b Other nonregional lymph node metastases or distant metastases</td>
</tr>
<tr>
<td><strong>Stage groupings</strong></td>
</tr>
<tr>
<td>Stage 0 Tis N0 M0</td>
</tr>
<tr>
<td>Stage I T1 N0 M0</td>
</tr>
<tr>
<td>Stage IIA T2 N0 M0</td>
</tr>
<tr>
<td>Stage IIB T1 N1 M0</td>
</tr>
<tr>
<td>Stage III T3 N1 M0</td>
</tr>
<tr>
<td>Stage IVA T1-4 N0-1 M1a</td>
</tr>
<tr>
<td>Stage IVB T1-4 N0-1 M1b</td>
</tr>
</tbody>
</table>

*Termed carcinoma in situ in current AJCC and UICC classifications.
†Muscularis mucosae not included in current AJCC or UICC classifications.
‡Complexity of distant metastases has been simplified.

<table>
<thead>
<tr>
<th>TABLE 2. Prevalence of pathologic tumor classification (TNM), additional tumor features, and stage groupings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Characteristic</strong></td>
</tr>
<tr>
<td><strong>TNM classification</strong></td>
</tr>
<tr>
<td>pT</td>
</tr>
<tr>
<td>Tis</td>
</tr>
<tr>
<td>1a</td>
</tr>
<tr>
<td>1b</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>pN</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>Metastatic nodes</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>≥4</td>
</tr>
<tr>
<td>pM</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td><strong>Tumor features</strong></td>
</tr>
<tr>
<td>Histopathologic type</td>
</tr>
<tr>
<td>Adenocarcinoma</td>
</tr>
<tr>
<td>Squamous cell</td>
</tr>
<tr>
<td>Adenosquamous cell</td>
</tr>
<tr>
<td>Unclassified</td>
</tr>
<tr>
<td><strong>Histopathologic grade</strong></td>
</tr>
<tr>
<td>Well</td>
</tr>
<tr>
<td>Moderately well</td>
</tr>
<tr>
<td>Moderate</td>
</tr>
<tr>
<td>Moderately poor</td>
</tr>
<tr>
<td>Poor</td>
</tr>
<tr>
<td>Unclassified</td>
</tr>
<tr>
<td><strong>AJCC stage groupings</strong></td>
</tr>
<tr>
<td>O</td>
</tr>
<tr>
<td>I</td>
</tr>
<tr>
<td>IIA</td>
</tr>
<tr>
<td>IIB</td>
</tr>
<tr>
<td>III</td>
</tr>
<tr>
<td>IVA</td>
</tr>
<tr>
<td>IVB</td>
</tr>
</tbody>
</table>

Clinic Foundation.
**TNM classification.** All staging was determined from examination of resected tissue at esophagectomy (pathologic stage). pT was determined microscopically and categorized according to maximum depth of invasion into each distinct esophageal layer. Subclassification of pT was based on esophageal layer: epithelium (HGD), lamina propria, muscularis mucosa, submucosa, muscularis propria, periesophageal tissue, and adjacent structures. pN was determined by microscopic examination of all resected regional lymph nodes. Subclassification of pN was based on number of regional lymph nodes containing metastatic cancer, although in the analysis both this and the percentage of resected nodes containing metastases were examined. The median number of nodes examined in resection specimens was 11 (interquartile range, 5-18). pM was determined by preoperative clinical staging and intraoperative assessment of distant sites within the surgical fields. Subclassification of pM was based on metastases to specific nonregional lymph nodes (pM1a, lower thoracic esophagus to celiac lymph nodes or upper thoracic to supraclavicular lymph nodes) and other nonregional lymph nodes or distant sites (pM1b). Patients' TNM classifications are listed in Table 2.

**Additional factors.** Histopathologic type was assessed and classified as adenocarcinoma, squamous cell cancer, and adenosquamous cell cancer (Table 2). Histopathologic grade was assessed and classified as well differentiated (plus HGD), well to moderately well differentiated, moderately differentiated, moderately poor to poorly differentiated, and poorly differentiated (undifferentiated) (Table 2).

**Stage groupings.** Current stage groupings are outlined in Table 1. JCC guidelines constrain assignment of staging groups between stage 0, HGD (in situ cancer), and stage IV, distant metastatic disease (Table 2).

**Methods**

**Adequacy of current staging.** Monotonicity is strictly ordered decreasing survival with increasing classification or grouping. Distinctiveness is separation between monotonic classifications or groupings based on survival. Homogeneity is absence of distinct subclassifications or subgroups within classifications or groupings based on survival.

Monotonicity, distinctiveness, and homogeneity of survival in the context of current TNM classifications and stage groupings were examined. Monotonicity was assessed by evaluating (1) stratified survival curves for decreasing survival between classifications and groups and (2) coefficients of Cox proportional hazard and nonproportional hazard function multivariable analyses. Distinctiveness is separation between monotonic classifications or groupings based on survival.

**TABLE 3. Depth of tumor invasion and survival in pN0 M0 cancers**

<table>
<thead>
<tr>
<th>Classification</th>
<th>Coefficient ± SE</th>
<th>P value</th>
<th>Hazard ratio</th>
<th>95% CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current classification*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>0.75 ± 0.46</td>
<td>.10</td>
<td>2.1</td>
<td>0.87-5.2</td>
</tr>
<tr>
<td>T2</td>
<td>1.50 ± 0.48</td>
<td>.002</td>
<td>4.5</td>
<td>1.74-11.5</td>
</tr>
<tr>
<td>T3</td>
<td>2.1 ± 0.44</td>
<td>&lt;.0001</td>
<td>8.5</td>
<td>3.6-20</td>
</tr>
<tr>
<td>T4</td>
<td>2.4 ± 0.71</td>
<td>.0006</td>
<td>11.3</td>
<td>2.8-46</td>
</tr>
<tr>
<td>Subclassifications of pT*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1a (intramucosal)</td>
<td>0.40 ± 0.51</td>
<td>.4</td>
<td>1.49</td>
<td>0.55-4.0</td>
</tr>
<tr>
<td>T1b (submucosal)</td>
<td>1.25 ± 0.50</td>
<td>.01</td>
<td>3.5</td>
<td>1.30-9.3</td>
</tr>
<tr>
<td>T2</td>
<td>1.50 ± 0.48</td>
<td>.002</td>
<td>4.5</td>
<td>1.74-11.5</td>
</tr>
<tr>
<td>T3</td>
<td>2.1 ± 0.44</td>
<td>&lt;.0001</td>
<td>8.5</td>
<td>3.6-20</td>
</tr>
<tr>
<td>T4</td>
<td>2.4 ± 0.71</td>
<td>.0006</td>
<td>11.4</td>
<td>2.8-46</td>
</tr>
<tr>
<td>Residual risk†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1b (submucosal)</td>
<td>1.00 ± 0.38</td>
<td>.008</td>
<td>2.7</td>
<td>1.30-5.7</td>
</tr>
<tr>
<td>T2</td>
<td>1.25 ± 0.35</td>
<td>.0004</td>
<td>3.5</td>
<td>1.76-7.0</td>
</tr>
<tr>
<td>T3</td>
<td>1.90 ± 0.29</td>
<td>&lt;.0001</td>
<td>6.7</td>
<td>3.8-11.7</td>
</tr>
<tr>
<td>T4</td>
<td>2.2 ± 0.63</td>
<td>.0005</td>
<td>8.9</td>
<td>2.6-31</td>
</tr>
</tbody>
</table>

CL, Confidence limits.

*Compared with Tis tumors.
†Compared with Tis plus T1 mucosal tumors.

Figure 1. Survival after esophagectomy according to T classification in pN0 M0 disease. A, Current classification. Number of patients in each classification and number traced at 5 years, respectively, were as follows: Tis (44 and 21), T1 (98 and 26), T2 (31 and 8), T3 (57 and 12), and T4 (4 and 0). B, Proposed subclassification of T1 into T1a (intramucosal, n = 66, 18 traced at 5 years) and T1b (submucosal, n = 32, 8 traced at 5 years). Shown for reference are Tis (HGD) and T2. Vertical bars represent 68% confidence limits.
distinctiveness was assessed by calculating confidence limits of these comparisons of P values in multivariable analysis. Homogeneity was assessed by analysis of residual risk, forcing the current classifications and groupings into multivariable models and identifying individual misclassifications and misgroupings.

Refining staging. Survival-tree analysis (recursive partitioning) analysis and multivariable analysis were used to determine subclassifications of TNM that identified monotonic, distinctive, and homogeneous survival (appendix). Subclassifications of TNM that did not contribute to monotonic, distinctive, and homogeneous survival were discarded.

For multivariable analysis, both Cox proportional hazard and hazard function nonproportional methods were used. All subclassifications of TNM and their interactions were examined.

Histopathologic type and grade were added to both the recursive partitioning analysis and the multivariable analysis to identify additional refinement of staging.

Stage groupings were first assigned based on AJCC guidelines that fix stage 0 and stage IV (constrained). The unassigned anatomic classifications were then assigned to stage I, II, or III based on recursive partitioning analysis and multivariable analysis. These groupings were verified by both methods of multivariable analysis and by analysis of residual risk to identify and correct residual inhomogeneity and to maximize goodness of fit (likelihood).

Stage groupings were then freed from AJCC constraints (unconstrained), and the analyses were repeated to allow TNM classifications to be freely assigned to 1 of 5 stage groupings (although 4 groupings were found to be adequate).

Results
Adequacy of Current Staging
In the absence of regional lymph node metastases (N0) and distant metastases (M0), current T classifications...
were monotonic and distinctive (Figure 1, A, and Table 3). Current classification of T1 tumors was inhomogeneous. Homogeneous classifications were obtained by dividing T1 into T1a (intramucosal) and T1b (submucosal; Figure 1, B, and Table 3). Within T1, there was a monotonically decreasing and distinctive survival with increasing depth of tumor invasion from the mucosa to the submucosa.

In the absence of distant metastases (M1), the current N classification was monotonic and distinctive (Figure 2, A, and Table 4). However, N1 was inhomogeneous. Survival decreased as the number of regional lymph node metastases increased (Figure 2, B, and Table 4; \( P < .0001 \)). In contrast to N0, stratification of N1 tumors by T classification was monotonic but not distinctive (Figure 2, C, and Table 4).

The current M classifications were monotonic and distinctive; however, within M1, survival was homogeneous (Figure 3 and Table 4).

Current stage groupings were monotonic (Figure 4 and Table 5). They were distinctive for stages I, IIA, and III and possibly for 0 versus I, but not for IIB versus III versus IV (Figure 4 and Table 5).

**Refining Staging**

*Subclassification of TNM.* Analysis of residual risk suggested subclassifying T1 into T1a (intramucosal) and T1b (submucosal; Figure 1, B, and Table 3) and subclassifying N1 by the number of metastatic regional lymph nodes (Figure 2, B, and Table 4). For purposes of distinct and homogeneous subclassifications, regional lymph node metastases were best classified as N1 (\( \geq 2 \) metastatic regional lymph nodes) and N2 (\( \geq 3 \) metastatic regional lymph nodes; Figure 2, B, and Table 4). Subclassifying M1 into M1a and M1b did not provide distinctive prognostic information because survival was homogeneous (Figure 3 and Table 4).

**Additional factors.** Histopathologic type \(( P = .17 \) and grade \(( P = .3 \) provided little additional prognostic information for M0 tumors by multivariable analysis. Recursive partitioning analysis suggested that adenocarcinoma and higher-grade cancer resulted in poorer survival and refined classification of M1 tumors.

**Redefinition of stage groupings.** Current stage groupings were inhomogeneous. Analysis of residual risk identified patients in 4 TNM classifications whose survival was not well represented by this grouping \(( P < .05 \), Table 6). These were (1) submucosal cancers without regional lymph node metastases (T1b N0 M0), which are currently grouped with intramucosal cancers in stage I; (2) T2 N0 M0 cancers, which are currently grouped in stage IIA; (3) T3 N0 M0 cancers, which are currently grouped in stage IIB; and (4) tumors with 3 or more regional lymph node metastases, which are currently grouped in both stages IIB and III.

Based on the above refinements of TNM classification (Table 7 and Appendix Table 1), optimal stage groupings were constructed (Figure 5 and appendix). Unconstrained stage grouping produced 4 distinct groups with monotonically decreasing survival that were reasonably homogeneous within groups. Compared with current groupings, this system improved the monotonicity, distinctiveness, and homogeneity of other classifications.

**Discussion**

TNM staging is a powerful tool that plays a major role in treatment decisions, prognostication, and research. However, the process of staging cancer is dynamic. As knowledge of the disease and its treatment and experience with staging increase, there is an inherent need for revision of
GTS

TABLE 4. Regional lymph node status (N), metastatic disease (M), and survival

<table>
<thead>
<tr>
<th>Classification</th>
<th>Coefficient ± SE</th>
<th>P value</th>
<th>Hazard ratio</th>
<th>95% CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>N in pM0 disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pN1 vs pN0</td>
<td>1.44 ± 0.135</td>
<td>&lt;.0001</td>
<td>4.2</td>
<td>3.2-5.5</td>
</tr>
<tr>
<td>N and T* in pN1 M0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More metastatic nodes†</td>
<td>0.30 ± 0.120</td>
<td>.01</td>
<td>1.35</td>
<td>1.07-1.71</td>
</tr>
<tr>
<td>pT1</td>
<td>−0.88 ± 0.57</td>
<td>.12</td>
<td>0.41</td>
<td>0.136-1.26</td>
</tr>
<tr>
<td>pT2</td>
<td>−0.58 ± 0.49</td>
<td>.24</td>
<td>0.56</td>
<td>0.22-1.47</td>
</tr>
<tr>
<td>pT3</td>
<td>−0.36 ± 0.42</td>
<td>.39</td>
<td>0.70</td>
<td>0.30-1.59</td>
</tr>
<tr>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pM1 vs pM0</td>
<td>1.12 ± 0.178</td>
<td>&lt;.0001</td>
<td>3.1</td>
<td>2.2-4.3</td>
</tr>
<tr>
<td>M in pM1 disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pM1b vs pM1a</td>
<td>0.33 ± 0.34</td>
<td>.9</td>
<td>1.39</td>
<td>0.7-2.7</td>
</tr>
</tbody>
</table>

CL, Confidence limits.
*Compared with pT4 tumors.
†Ln (number of regional lymph nodes containing metastatic cancer) natural logarithmic transformation; coefficient and hazard ratio are for a log unit increase.

TABLE 5. Current stage groupings and survival

<table>
<thead>
<tr>
<th>Stage grouping*</th>
<th>Coefficient ± SE</th>
<th>P value</th>
<th>Hazard ratio</th>
<th>95% CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.76 ± 0.46</td>
<td>.1</td>
<td>2.1</td>
<td>0.87-5.3</td>
</tr>
<tr>
<td>IIA</td>
<td>1.94 ± 0.43</td>
<td>&lt;.0001</td>
<td>7.0</td>
<td>3.0-16.1</td>
</tr>
<tr>
<td>IIB</td>
<td>2.4 ± 0.46</td>
<td>&lt;.0001</td>
<td>10.8</td>
<td>4.4-27</td>
</tr>
<tr>
<td>III</td>
<td>2.8 ± 0.42</td>
<td>&lt;.0001</td>
<td>16.4</td>
<td>7.2-37</td>
</tr>
<tr>
<td>IVA</td>
<td>3.1 ± 0.46</td>
<td>&lt;.0001</td>
<td>23</td>
<td>9.1-56</td>
</tr>
<tr>
<td>IVB</td>
<td>3.5 ± 0.49</td>
<td>&lt;.0001</td>
<td>33</td>
<td>12.6-86</td>
</tr>
</tbody>
</table>

CL, Confidence limit.
*Compared with stage 0.

staging. The first practical TNM staging of esophageal cancer was introduced in 1987 and revised most recently in 2002. However, further revisions are and will be required to improve staging of esophageal cancer.

Balch and colleagues6,7 introduced a strategy to refine and improve staging of melanoma. We have used their combination of algorithmic and regression analyses to assess monotonicity and distinctiveness of survival. Analysis of residual risk has been added to improve homogeneity of classification and stage groupings. This combination analysis should be considered in future refinements of cancer staging. Further methodological refinement to minimize misclassification in series with larger number of patients would require the addition of multidimensional algorithmic modeling.8

TNM Classifications

For patients without metastases to regional lymph nodes (N0) or distant sites (M0), depth of tumor invasion (T) is remarkably monotonic and distinctive for survival. However, within T1, survival is not homogeneous. This requires reclassifying T1 tumors into T1a (intramucosal) and T1b (submucosal) to produce distinct classes. Both clinical staging by endoscopic ultrasonography and pathologic staging allow differentiation of T1a from T1b cancers. This is a practical classification that improves staging.9 It is noteworthy that neither the AJCC nor the International Union Against Cancer (UICC) mentions the muscularis mucosa specifically in the definition of T1 tumors.1,5 This omission must be addressed in future definitions of T1 classification.

For patients without distant metastases (M0), the current N classification is monotonic and distinctive. However, for patients with regional lymph node metastases (N1), stratification on the basis of depth of tumor invasion is not distinctive. Within N1 classification, survival is not homogeneous. Increasing lymph node burden, as reflected by the number of regional nodes that contain metastases, is associated with poorer survival.10,11 For purposes of staging, distinctive survival is produced by classifying regional lymph node metastases into N1 (≤2 nodes) and N2 (≥3 nodes). Survival differences are so small that only one factor is necessary to account for them. The number of regional nodal metastases is the one factor identified by the analyses; the percentage of resected lymph nodes containing metastases adds no further information. This finding might account for apparent discrepancies between our proposed refinement of staging and that of Korst and colleagues,12 which was based on univariable analyses.

The current classification of distant metastatic disease (M1) is monotonic and distinctive. Survival of patients with M1 disease is poor and homogeneous for all T and N classifications. Distinctive survival is not produced by the subclassification into M1a and M1b, and thus this subclassification should be abandoned.13 Although Steup and colleagues14 disagree with this finding, their data support the homogeneity of survival within these subclassifications.
Stage Groupings

Stage groupings assemble the multiple variables of TNM classifications into a small number of homogeneous clusters with monotonic and distinctive survival. Current stage groupings follow a pattern of increasing T, followed by increasing N and then increasing M. However, esophageal cancers do not grow and metastasize according to this orderly pattern. Some patients have small tumors with minimal invasion of the esophageal wall and a single regional nodal metastasis that decreases survival, but not to the extent implied by current stage groupings. In contradistinction, a patient with the same superficial cancer with extensive regional lymph node metastases has a prognosis similar to that of a patient with distant metastases. This requires regrouping not by pattern but by distinctive and homogeneous survival.

Assigning stage 0 to only HGD and stage IV to only distant metastases, as in the current AJCC guidelines, limits the definition of distinct and homogeneous groups. Patients with cancer confined to the epithelium (Tis N0 M0) have a survival that is not distinct from those with invasion beyond the basement membrane to involve either the lamina propria or muscularis mucosae (T1a N0 M0). This grouping is clinically practical because distinguishing HGD from intramucosal cancer might be difficult even for the most experienced pathologist. Patients with cancers that invade adjacent structures and metastasize to 1 or 2 regional lymph nodes (T4 N1 M0) and patients with multiple regional lymph node metastases (T1-3 N2 M0) do as poorly as those with distant metastases (M1). Freeing of the upper and lower constraints on grouping thus permits improved staging.

The goal of stage grouping is to condense TNM classification into a small and convenient number of groups to allow tabulation and analysis. The designation of stage 0 (HGD) and stage IV (M1) forces use of 5 stage groupings. There has been a trend toward making stage groupings more complex by introducing subgroups, such as A and B. These designations and trends defeat the purpose of stage grouping. Free assignment of stage groupings for esophageal cancer requires only 4 stages to represent survival adequately.

Additional Factors

TNM descriptors are unique to each cancer and, in addition to anatomic extent, may include such nonanatomic factors as histopathologic type (thyroid, histopathologic grade (sarcomas), serum biologic markers (testis), and patient demography (thyroid and gestational trophoblastic tumors). Additional histopathologic type or grade did not improve staging of esophageal cancers. Further staging classifications of esophageal cancer await identification of nonanatomic factors that predict survival and add additional information not provided by TNM classifications.

Strengths and Weaknesses

Although a single-institution study might be limited in numbers and ability to generalize inferences, its strength is

### Table 6. Residual risk analysis of stage groupings

<table>
<thead>
<tr>
<th>Stage grouping</th>
<th>Coefficient ± SE</th>
<th>P value</th>
<th>Hazard ratio</th>
<th>95% CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>0.45 ± 0.50</td>
<td>.4</td>
<td>1.57</td>
<td>0.6-4.2</td>
</tr>
<tr>
<td>Stage IIA</td>
<td>2.2 ± 0.43</td>
<td>&lt;.0001</td>
<td>8.6</td>
<td>3.7-20</td>
</tr>
<tr>
<td>Stage IIB</td>
<td>2.1 ± 0.47</td>
<td>&lt;.0001</td>
<td>8.6</td>
<td>3.4-21</td>
</tr>
<tr>
<td>Stage III</td>
<td>2.5 ± 0.43</td>
<td>&lt;.0001</td>
<td>12.4</td>
<td>5.3-29</td>
</tr>
<tr>
<td>Stage IVA</td>
<td>3.2 ± 0.46</td>
<td>&lt;.0001</td>
<td>24</td>
<td>9.6-59</td>
</tr>
<tr>
<td>Stage IVB</td>
<td>3.6 ± 0.49</td>
<td>&lt;.0001</td>
<td>35</td>
<td>13.4-31</td>
</tr>
<tr>
<td>Additional groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pT1b N0 M0 or pT1a N1 M0†</td>
<td>0.76 ± 0.37</td>
<td>.04</td>
<td>2.1</td>
<td>1.03-4.5</td>
</tr>
<tr>
<td>pT2 N0 M0‡</td>
<td>−0.66 ± 0.29</td>
<td>.03</td>
<td>0.52</td>
<td>0.29-0.92</td>
</tr>
<tr>
<td>pT4 N1 M0 or pTany N2 M0‖</td>
<td>0.54 ± 0.160</td>
<td>.0007</td>
<td>1.71</td>
<td>1.25-2.3</td>
</tr>
</tbody>
</table>

CL, Confidence limits.
*Compared with stage 0.
†Suggests Tis (stage 0) and pT1a N0 M0 (residual after accounting for pT1a N1 M0 and pT1b N0 M0) for a single-stage grouping.
‡Suggests these patients are at lower risk than remaining stage IIA.
§Suggests stage IVA (M1a) and stage IVB (M1b) should be coalesced into a single group.
‖Suggests these patients are at higher risk than remainder of stage III, and more similar to stage IV.

### Table 7. Refined stage groupings

<table>
<thead>
<tr>
<th>Stage</th>
<th>Constrained*</th>
<th>Unconstrained</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Tis (HGD)</td>
<td>—</td>
</tr>
<tr>
<td>I</td>
<td>T1a N0 M0</td>
<td>Tis (HGD), T1a N0 M0</td>
</tr>
<tr>
<td>II</td>
<td>T1b N0 M0, T1a N1 M0</td>
<td>T1b NO M0, T1a N1 M0</td>
</tr>
<tr>
<td>III</td>
<td>T3 N0 M0, T1b/T2 N1 M0</td>
<td>T3 N0 M0, T1b/T2 N1 M0</td>
</tr>
<tr>
<td>IV</td>
<td>T4 N1 M0, Tany N2 M0</td>
<td>T4 N1 M0, Tany N2 M0</td>
</tr>
</tbody>
</table>

*By AJCC definition of stage 0 and stage IV.
uniformity of patient selection, clinical staging, treatment decisions, surgical technique, postoperative care, and pathologic staging. This study has been purposely limited to surgical therapy, avoiding influences of multimodality therapy, which includes induction toxicity, biologic processes (e.g., hysteresis effect of downstaging), and disease progression during therapy.\(^\text{16,17}\) This staging system is valid and useful for both clinical and pathologic staging. A limitation of this strategy is that few patients with locally advanced T4 tumors are available for analysis, limiting the ability to recognize distinctiveness and homogeneity in locally advanced disease. We have not included completeness of resection in T classification because this is a surgical determinant and not a tumor characteristic.\(^\text{18}\)

Survival of patients who have undergone esophagectomy for cancer does not reflect the natural history of untreated disease. This is in part why staging is dynamic: it will change as therapy evolves. Furthermore, staging might differ for patients treated nonsurgically or by combination therapy. In all these instances, survival is influenced by many factors, of which disease stage is but one. Patient demographics, clinical status, clinical staging, surgical approach, and postoperative care are at play, and a combination of factors will best predict survival and direct decision making and treatment.\(^\text{9,13,14,16,17}\)

The utility of this refinement of esophageal cancer staging depends on modern clinical staging tools and complete pathologic assessment. Currently, endoscopic esophageal ultrasonography with fine-needle aspiration of lymph nodes and the combination of computed tomography and fluorodeoxyglucose positron emission tomography provide the most accurate clinical staging. Pathologic staging requires resection of the primary tumor, assessment of adjacent organs, adequate lymphadenectomy or lymph node sampling, and assessment of accessible distant sites.

We thank Diane Baisden for data management and Tess Knerik for editorial assistance.

References


Appendix: Refining Staging

Subclassification of TNM

Recursive partitioning analysis for constant hazard, log rank, Gordon-Olshen, adaptive normalization, and global normalization methods were used for child partitioning and pruning according to
monotonicity and distinctiveness of survival3 (Appendix Figure 1). Partitions were further coalesced to provide homogeneous survival. Within each partition there could be 1 or more classifications.

Redefinition of Stage Groupings
We initially allowed subgroupings (A, B, C) of groups 0 to IV that represented major lack of homogeneity within stage groupings. Within each stage grouping of constrained and unconstrained systems, inhomogeneity within groups was addressed by assignment to substages.

Constrained grouping produced lack of distinctiveness at the extremes of stage groupings (stages 0 and IV). Criteria set forth by the AJCC were best met by a redefined unconstrained staging system (Appendix Table 1 and Appendix Figure 2, A and B). Seven stage groupings were identified by the unconstrained analyses (Appendix Table 1). However, it was clear that the analysis had produced 3 pairs of stages and substages for stages II, III, and IV (Appendix Figure 2, A). Therefore, these pairs were coalesced to yield a total of 4 major stage groupings (Figure 5).

Appendix Figure 1. Recursive partitioning analysis of TNM classification using log-rank method. Within partitions are (1) number of patients and (2) median survival (years).

Appendix Figure 2. Survival after esophagectomy according to refined stage groupings. A. Groupings constrained by AJCC definitions of stages 0 and IV. Number of patients in each grouping and number traced at 5 years, respectively, were as follows: 0 (44 and 21), I (66 and 18), IIA (34 and 8), IIB (31 and 8), IIIA (72 and 14), IIIB (66 and 8), IIIC (125 and 2), and IV (42 and 0). B. Groupings free of AJCC constraints. Number of patients in each grouping and number traced at 5 years, respectively, were as follows: I (110 and 39), IIA (34 and 8), IIB (31 and 8), IIIA (72 and 14), IIIB (66 and 8), IVA (125 and 2), and IVB (42 and 0).

APPENDIX TABLE 1. Refined stage groupings

<table>
<thead>
<tr>
<th>Stage grouping</th>
<th>Constrained*</th>
<th>Unconstrained</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Tis (HGD)</td>
<td>—</td>
</tr>
<tr>
<td>I</td>
<td>T1a N0 M0</td>
<td>Tis (HDG), T1a N0 M0</td>
</tr>
<tr>
<td>IIA</td>
<td>T1b N0 M0, T1a N1 M0</td>
<td>T1b N0 M0, T1a N1 M0</td>
</tr>
<tr>
<td>IIB</td>
<td>T2 N0 M0</td>
<td>T2 N0 M0</td>
</tr>
<tr>
<td>IIIA</td>
<td>T3 N0 M0, T1b/T2 N1 M0</td>
<td>T3 N0 M0, T1b/T2 N1 M0</td>
</tr>
<tr>
<td>IIIB</td>
<td>T3 N1 M0, T4 N0 M0</td>
<td>T3 N1 M0, T4 N0 M0</td>
</tr>
<tr>
<td>IIIC</td>
<td>T4 N1 M0, Tany N2 M0</td>
<td>—</td>
</tr>
<tr>
<td>IVA</td>
<td>M1</td>
<td>T4 N1 M0, Tany N2 M0</td>
</tr>
<tr>
<td>IVB</td>
<td>M1</td>
<td>M1</td>
</tr>
</tbody>
</table>

*By AJCC definition of stage 0 and stage IV.

Appendix Table 1. Refined stage groupings

Discussion
Dr Carolyn E. Reed (Charleston, SC). This article confirms what most surgeons who deal with esophageal cancer know: the present staging system is inadequate. Dr Rice and his colleagues have used sophisticated and powerful statistical modeling to refine the esophageal staging system.

The reclassification of T1 into T1a and T1b makes sense from an anatomic-biologic standpoint and from the ease of identifying these subsets by means of endoscopic ultrasonography. The addition of the number of metastatic regional lymph nodes to improve survival stratification has been shown by others, although one could argue about the number used as a cutoff between N1 and N2.
I accept that patients with T4 any N and T1 through T3 N2 disease will do poorly with surgical intervention alone, mirroring patients with metastatic disease. However, keeping some subclassification of your unrestrained stage IV group might be useful for future studies involving induction therapy.

I have a couple of questions for you. You had only 26 patients with M1a disease. Because most of us are seeing large numbers of adenocarcinomas of the gastroesophageal junction, M1a becomes a significant subset. Does your institution offer these patients with stage IV disease induction therapy? Should we consider a single positive celiac lymph node either N1 and maybe stage III or equivalent to N2? I ask this because looking at N1 versus N2 might be important when we study the results of induction strategies.

I assume you would not offer surgical intervention alone to any of your new patients with stage IV disease. Would you use endoscopic ultrasonography and fine-needle aspiration results after induction therapy to continue to deny patients surgical intervention (ie, continued evidence of lymph nodal disease)?

I believe the time has come to redefine the staging system for esophageal cancer. Finally, I would make a plea that staging must incorporate endoscopic ultrasonography if it is to be considered adequate.

Dr Rice. I would like to thank Dr Reed for her kind comments. She is one of the pioneers in staging and has demonstrated its importance in the treatment of esophageal cancer.

Your questions are really about the M1a group. This analysis considers patients who undergo surgical intervention alone. Therefore only 26 patients with M1a disease were available for analysis. It is our practice that if we identify people with lower thoracic esophageal cancers or cancers of the esophagogastric junction with positive celiac nodes, we offer them induction chemoradiation therapy. I think that if you have endoscopic ultrasonographic and fine-needle aspiration confirmation of persistent disease after such therapy, patients should not proceed to surgical intervention.

Dr Alan G. Casson (Halifax, Nova Scotia, Canada). I did not see this in your abstract, and I hope I did not miss it in your talk, but could you tell us how many of these were squamous tumors and how many were adenocarcinomas?

Dr Rice. The majority were adenocarcinomas, approximately 80% were adenocarcinomas, and 20% were squamous cell carcinomas. Although I could not cover it in the presentation, we included histologic type and grade of tumors, and by means of recursive partitioning analysis, those factors became important only in grouping for advanced-stage cancers, with adenocarcinomas doing a little bit worse than squamous cell carcinomas for the stage IV cancers, and poorly differentiated cancers doing worse than more differentiated cancers.

Dr Casson. Thanks for clarifying that because I think there is increasing biologic information that these are very different diseases, and I think to lump them together just because they happen to share the same organ might not necessarily be the correct thing to do in the future.

Dr Bryan Meyers (St Louis, Mo). Many of us have had the experience after a transhiatal esophagectomy in which the pathologist reports the T stage and then reports that there are no lymph nodes in the resected specimen. If you call them back and say, “You need to keep looking,” they will call back and say, “Perhaps there were 9 or 12 lymph nodes.” I certainly know that if you have zero lymph nodes in the specimen that N0 status means nothing. Has that ever occurred in your experience, and do you think those events occur frequently enough to be influential in your staging algorithm? By the same token, is there a certain number of lymph nodes that a specimen should have to allow the results of your strategy to be generalizable to that particular individual?

Dr Rice. The AJCC staging manual states that if there are fewer than 6 nodes in a specimen, that is an inadequate resection. That is the bare minimum. I think at transhiatal esophagectomy you have to be a lymph node sampler, and the reason to add a thoracotomy is to add lymphadenectomy for prognostication. Others believe this should be added for therapy, but we do it at least for prognostication. If you do not sample a lymph node, everybody has N0 disease or NX disease, and 5-year survival is going to be approximately 20% for all comers. Therefore I think lymph node sampling and lymphadenectomy whenever possible are crucial to have this staging system work.

Dr Scott J. Swanson (New York, NY). I have 2 brief questions. One, as we do more induction therapy for these patients, does that affect how you ultimately fit them into this new staging system?

Dr Rice. This staging system is for surgical intervention only. What you will find with induction therapy is that it just produces 2 groups, responders and nonresponders. I plan to write about that next. But basically there is stage I and stage II disease. Either you respond and have an intermediate survival, or you do not respond, have persistent disease, have nodal disease, and your survival is as bad or worse than if we had left you alone.

Dr Swanson. I have one other related question. Given what you see for 5-year survival, even in your stage II group, what would you say is the best therapy for that group going forward?

Dr Rice. Well, that is always a question: when do you add induction chemoradiation therapy? We have tended to use the cutoff point of 50% 5-year survival, and you usually get that with tumors confined to the esophageal wall without regional lymph node metastases. Therefore we have made a dividing line. If you have T2 N0 cancer or less, we will operate. If you have T3 N0 or any N1 cancer, you will be offered induction chemoradiation therapy because your survival with T3 N0 is around 20% to 25% at 5 years in our experience, and with N1 cancer it is about 7%.

Dr Robert J. Ginsberg (Toronto, Ontario, Canada). I think it was a beautiful article and a nice analysis of surgically resected esophageal cancer, but we have to remember what the TNM staging system is. Number one, it has to be relevant to doctors practicing in the middle of Africa, in the middle of China, in the middle of Indochina, as well as in the United States, where they have endoscopic ultrasonographic and fine-needle aspiration ability. Number 2, for treatment decision making, you must perform clinical staging. The staging system has to be relevant not only to pathologic staging but also to clinical staging. I think you would do us a great service if you went back and looked at the clinical stage of those patients and once again analyzed the data and ensured that this same distribution of survival, albeit probably lower for each group other than stage Ia (and even that might be lower), is relevant for clinical staging as well. There are 3 types of staging: clinical staging; postinduction staging, which is y staging; and then p staging, which includes all relevant information. And for every-
body’s information, NX implies that no nodes were examined at p staging only. At clinical staging, you have to bite the bullet and estimate, in this case, the number of nodes. In other words, how are you going to clinically stage the number of involved nodes in which NX is not an option?

Dr Rice. Thank you very much for your comments.

There is no doubt that when we compare survival on the basis of pathologic stage with that of clinical stage, it is always less. This is particularly true with N0 cancer, but we have not examined it for patients with N1 disease. I agree with you that it has to work for clinical, y, and p staging. There is also r staging (retreatment) and a (autopsy) staging. But your comments are appreciated.