Esophageal Cancer: Associations With (pN+) Lymph Node Metastases

Thomas W. Rice, MD, Hemant Ishwaran, PhD,† Wayne L. Hofstetter, MD,‡ Paul H. Schipper, MD,§
Kenneth A. Kesler, MD,¶ Simon Law, MD,‖ Toni Lerut, MD, PhD,** Chadrick E. Denlinger, MD,††
Jarmo A. Salo, MD,‡‡ Walter J. Scott, MD,§§ Thomas J. Watson, MD,¶¶ Mark S. Allen, MD,‖‖
Long-Qi Chen, MD,*** Valerie W. Rusch, MD,††† Robert J. Cerfolio, MD,‡‡‡ James D. Luketich, MD,§§§
Andre Duranceau, MD,¶¶¶ Gail E. Darling, MD,‖‖‖ Manuel Pera, MD, PhD,****
Carolyn Apperson-Hansen, MStat,* and Eugene H. Blackstone, MD**

Objectives: To identify the associations of lymph node metastases (pN+), number of positive nodes, and pN subclassification with cancer, treatment, patient, geographic, and institutional variables, and to recommend extent of lymphadenectomy needed to accurately detect pN+ for esophageal cancer.

Summary Background Data: Limited data and traditional analytic techniques have precluded identifying intricate associations of pN+ with other cancer, treatment, and patient characteristics.

Methods: Data on 5806 esophagectomy patients from the Worldwide Esophageal Cancer Collaboration were analyzed by Random Forest machine learning techniques.

Results: pN+, number of positive nodes, and pN subclassification were associated with increasing depth of cancer invasion (pT), increasing cancer length, decreasing cancer differentiation (G), and more regional lymph nodes resected. Lymphadenectomy necessary to accurately detect pN+ is 60 for shorter, well-differentiated cancers (<2.5 cm) and 20 for longer, poorly differentiated ones.

Conclusions: In esophageal cancer, pN+, increasing number of positive nodes, and increasing pN classification are associated with deeper invading, longer, and poorly differentiated cancers. Consequently, if the goal of

lymphadenectomy is to accurately define pN+ status of such cancers, few nodes need to be removed. Conversely, superficial, shorter, and well-differentiated cancers require a more extensive lymphadenectomy to accurately define pN+ status.

Regional lymph node metastases (pN+) in esophageal cancer patients negatively affect outcome. However, the extent of lymphadenectomy necessary to maximize survival may be different from the extent necessary to accurately detect pN+. Understanding the associations of pN+ with extent of lymphadenectomy and other cancer, treatment, and patient characteristics may aid in treatment decisions and prognostication; however, their intricate associations have been difficult to identify because of limited data and traditional analytic techniques.[1]

To overcome prior data limitations, the Worldwide Esophageal Cancer Collaboration (WECC) uniquely provides, in a single database, cancer, treatment, patient, geographic, and institutional variables for a large cohort of esophagectomy patients.[2] To overcome prior analytic limitations, Random Forest analysis,[3] a modern machine learning technique,[4,5] permits exploration of such nonlinear, complex interrelationships.[3,6] Thus, the purpose of this study was to use this worldwide dataset and machine learning technique to identify WECC variables associated with pN+, number of positive nodes, and pN subclassification to develop recommendations for the number of regional lymph nodes that must be resected to accurately detect pN+.

PATIENTS AND METHODS

Patients

A total of 5806 patients in WECC, a worldwide consortium of institutions that have contributed deidentified patient data on esophagectomy for cancer, underwent esophagectomy alone (no preoperative or postoperative adjuvant therapy).[2] All datasets were approved for research by each site's Institutional Review Board, and data use agreements were executed when required. Worldwide Esophageal Cancer Collaboration variables included patient demographics (age, sex, and race), region of the world (East, West), institution, cancer characteristics (location in esophagus, histopathologic cell type, histopathologic grade, cancer length, pT, pN, pM, and number of regional lymph nodes containing metastatic cancer (herein termed positive nodes), and esophagectomy variables (number of regional lymph nodes resected, residual cancer, and year of surgery), 31 variables in all (Table 1).

TABLE 1. Patient and Esophageal Cancer Characteristics

Characteristic	n*	No. (%) or Mean ± SD
Age (y)	5664	63 ± 11
Male	5672	4392 (77)
Race	4063	
White		2834 (70)
Asian		1171 (28)
Other		58 (2.0)
East (part of world)	5673	1171 (21)
Location of cancer	4602	
Upper third		209 (5.0)
Middle third		1209 (26)
Lower third		3184 (69)
Cancer length (cm)	3150	3.5 ± 2.6
pT	5673	
is		370 (7.0)
1		1393 (25)
2		932 (16)
3		2787 (49)
4		191 (3.0)
pN	5639	
0		3199 (57)
+		2440 (43)
Number of regional lymph nodes	5238	
N0		3199 (61)
N1		952 (18)
N2		634 (12)
N3		453 (9.0)
Number of regional lymph nodes resected	4250	
0		438 (10)
1–5		1019 (24)
6–10		625 (15)
11–15		492 (12)
16–25		774 (18)
≥26		902 (21)
pM	5673	
0		5258 (93)
1		415 (7.0)
G (histologic grade)	4202	
G1 (well differentiated)		1422 (34)
G2 (moderately differentiated)		1349 (32)
G3 (poorly differentiated)		1425 (34)
G4 (undifferentiated)		6 (<1)
Histopathologic cell type	5673	
Squamous cell carcinoma		2154 (38)
Adenocarcinoma		3519 (62)
Resection margins	5194	
R0		4605 (89)
R1		464 (9.0)
R2		125 (2.0)
Year of esophagectomy	5806	
1970s		78 (1.3)
1980s		126 (22)
1990s		2303 (40)
2000s		2160 (37)

*Patients with data available.
SD indicates standard deviation.

Endpoints

The primary endpoint was pN+ disease. Secondary endpoints were number of positive nodes and pN classification: N0, no positive nodes; N1, 1 to 2 positive nodes; N2, 3 to 6 positive nodes; and N3, 7 or more positive nodes, based on the seventh edition of the American Joint Committee on Cancer (AJCC) staging manual.[7]

Data Analysis

Overview

Random Forest technology was chosen as the analytic strategy because of the known complex interactions among esophageal cancer characteristics identified in the seventh edition of the AJCC cancer staging endeavor.[4,7] The method is related to classification and decision-tree analyses, wherein the variable most related to an outcome of interest is first optimally split to improve prediction, then followed by more and more splits of it, and other variables, to create a tree (recursive partitioning, classification and regression trees). An individual tree ''grown'' by this method is inherently unstable, and this can be demonstrated by creating trees from bootstrap samples of the original dataset that split much differently (the bootstrap dataset is formed by random sampling with replacement until a dataset of equal size is generated; there will be some duplicated patients, and an average of 37% will not be sampled). To overcome this instability, a forest of trees is grown from such bootstrap samples, permitting an ensemble average to be formulated across the individual trees.[3] Because the method is completely nonparametric, with no restrictive underlying model assumptions, complex interactions among variables can be robustly accounted for. Validity of the forest is evaluated by assessing outcome among patients who were not selected by the bootstrap process.

Because some values were missing for some of the 31 variables, Random Forest imputation was employed to maximize use of the available data.[8]

Rather than *P* values, 2 metrics of prediction accuracy are generated based on the patients not selected (called the ''out-of-bag'' samples). The first ranks the importance of each variable in predicting the outcome of interest (variable importance).[3] The second quantifies the average number of branches before a variable is split (called ''minimal depth''): the closer to the trunk of the tree a variable is split, the more important that variable is to prediction accuracy.[9]

In summary, predictors of outcome using Random Forest technology are identified in 2 steps: building the forest based on cancer and other characteristics and the outcome of interest, and using the resulting forest to discover the importance of variables to the prediction of the outcome and their interrelationships with respect to outcome.

Details of how this method was used for this study are given in SDC Appendix E1, http://links.lww.com/SLA/A941 and briefly summarized as follows.

Predictors of lymph node metastases (pN+)

Predictors of N+ were identified using the randomForestSRC R package.[10] All 31 variables described in Table 1 plus surgical site were used to generate 1000 random bootstrap classification trees. The average 37% of patients not included in building a given tree were used to estimate the cross-validated probability of a patient being N+.

Predictors of Number of Positive Nodes

Predictors of number of positive nodes were identified using Random Forest nonparametric regression analysis.[10] A forest of 1000 trees was grown.

Predictors of Lymph Node (pN) Classification

A Random Forest strategy similar to that for identifying predictors of N+ was used for the ordinal outcome pN classification (N0, N1, N2, and N3) according to criteria in the seventh edition of the AJCC Cancer Staging Manual.[7]

Lymphadenectomy Needed to Accurately Detect Lymph Node Metastases (pN+)

To ascertain extent of lymph node resection needed to accurately detect pN+ for each combination of pT and length of cancer (dichotomized as <2.5 or ≥2.5 cm), the number of lymph nodes resected was replaced by a fixed cutoff value, and the predicted value of being pN+ was calculated using the previously constructed pN+ classification forest and the out-of-bag patients. The cutoff value was varied from 0 to the maximum number of resected nodes observed within the pT and cancer length categories. The average over all such predicted values yielded the adjusted predicted probability of pN+ for the given cutoff. The point at which these curves flattened was interpreted as the lymphadenectomy needed to accurately detect pN+.

RESULTS

Predictors of Lymph Node Metastases (pN+)

More advanced cancer characteristics—longer cancer length, higher pT, and higher G—were the strongest predictors of pN+ (Figs. 1, 2, and SDC Figure E1, http://links.lww.com/SLA/A941). Presence of pM1 (SDC Figure E2, http://links.lww.com/SLA/A941) and squamous cell cancer were predictive, but less so. The certainty of pN0 was improved with increasing number of negative lymph nodes resected, because the likelihood of resecting a positive node increased with increasing number of nodes resected (Fig. 3 and SDC Figure E3, http://links.lww.com/SLA/A941).

The complex interplay of cancer characteristics, number of resected nodes, and pN+ is illuminated in Figure 4. The relationships were more striking for adenocarcinoma than for squamous cell carcinoma. Identifying pN+ required fewer resected lymph nodes for deeply invasive, poorly differentiated, or long cancers than for superficial invasion, well differentiated, or short cancers. pT3, G3, or cancers longer than 4 cm were highly likely to be pN+ at all levels of nodes resected. Note that when no nodes were resected, the likelihood of pN+ became an average that increased as depth of invasion, grade, and length increased.

Predictors of Number of Positive Nodes

The same cancer characteristics that predicted presence of pN+ also predicted higher number of positive nodes (Fig. 5).
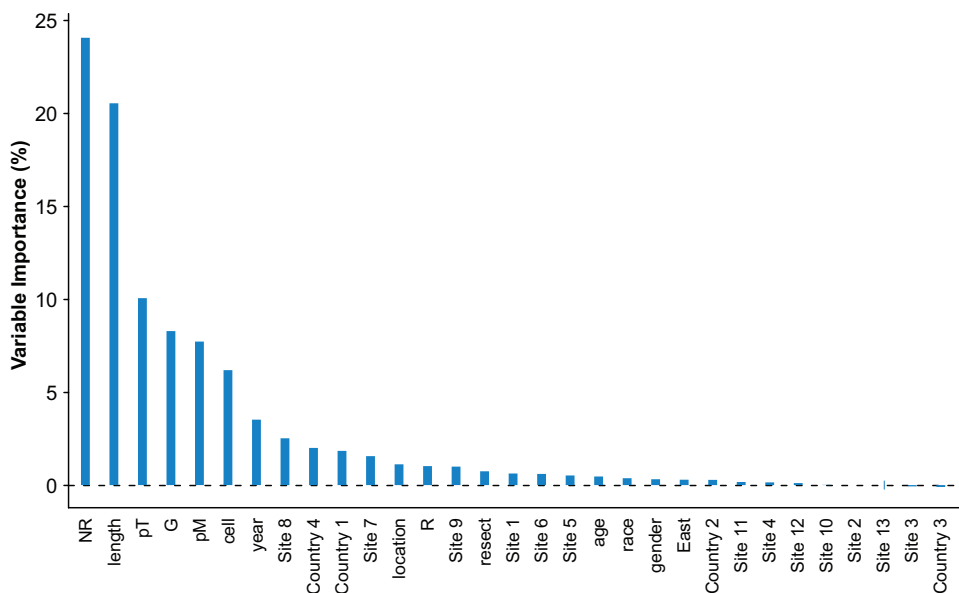
However, because number of positive nodes cannot exceed number of resected nodes, the relationship shifted to the right until about 11 to 20 nodes were resected. pT1 cancers that were node positive were predicted to have few positive nodes, but pT3/pT4 cancers that were node positive had a large number of positive nodes. Cancers longer than 4 cm were likely to have more than 5 positive nodes. Grades 1 and 2 had a similar number of positive nodes, but fewer than G3.

Predictors of Lymph Node (pN) Classification

The association of cancer characteristics with higher classification of pN+ was more striking in adenocarcinoma than in squamous cell carcinoma (Fig. 6). Superficial cancers were likely to be pN0, and if they were not, they were likely to be pN1 rather than pN2 or, rarely, pN3. In contrast, deeply invasive cancers were more likely to be pN2 or pN3 than pN1, and if they were poorly differentiated, they were unlikely to be pN0.

Lymphadenectomy Needed to Accurately Detect Lymph Node Metastases (pN+)

Extent of lymphadenectomy needed to accurately detect pN+ decreased with increasing cancer length (Fig. 7). Thus, for short cancers (<2.5 cm), the curves flatten at approximately 60 resected nodes. For cancers ≥2.5 cm, the curves clearly flatten at 20 nodes resected.

DISCUSSION

Principal Findings

The strongest associations with pN+ cancers reflect cancer growth, biology, and histology. Cancer growth, represented by 2 dimensions—depth of invasion and length—is strongly related to pN+. Adenocarcinoma invading beyond the muscularis propria (pT3) and cancers longer than 4 cm had a 60% to 80% prevalence of pN+. Cancer biology reflected by differentiation was strongly associated with pN+: G3 cancers had a 30% to 80% prevalence of pN+, depending on pT. However, we discovered a difference in prevalence of pN+ associated with histopathologic cell type; for the "same cancer," pN+ was more likely in adenocarcinoma than squamous cell cancer. The process of metastases resulting in spread to regional lymph nodes (pN+) and distant sites (pM1) not surprisingly links these 2 anatomic cancer
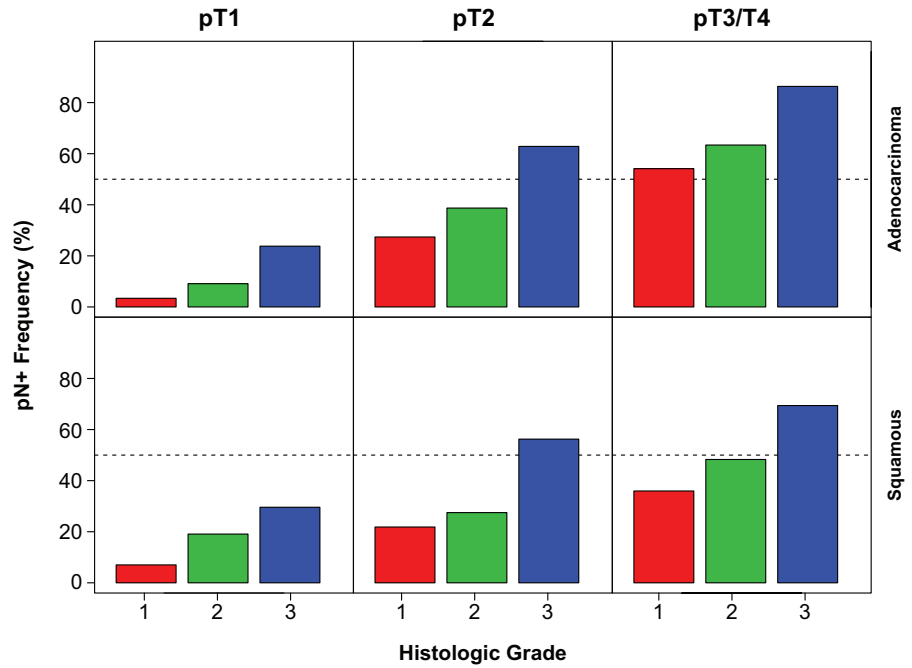

FIGURE 1. Variable importance for each of the 31 variables from Random Forest analysis of lymph node metastases status.

FIGURE 2. Frequency of lymph node metastases (pN+) according to G, pT, and histopathologic cell type (red = G1, G1, green = G2, and blue = G3). Adeno indicates adenocarcinoma; squam, squamous cell cancer.

characteristics. Rarely was a pM1 patient pN0. No cancer location was associated with a higher prevalence of pN+.

Intuitively, the more one looks, the more one finds. Thus, the only noncancer characteristic associated with pN+ was number of regional lymph nodes resected—the extent of surgical lymphadenectomy. Proportionately, this was most important in short well-differentiated squamous cell cancers. Thus,

extent of lymphadenectomy necessary to accurately detect pN+ must be greater for short, less invasive, well-differentiated cancers than for longer, more invasive, and poorly differentiated ones.

We recommend that number of positive nodes and number of resected nodes be reported separately and not confounded by conversion to the continuous variable "lymph node ratio."[11] Its use
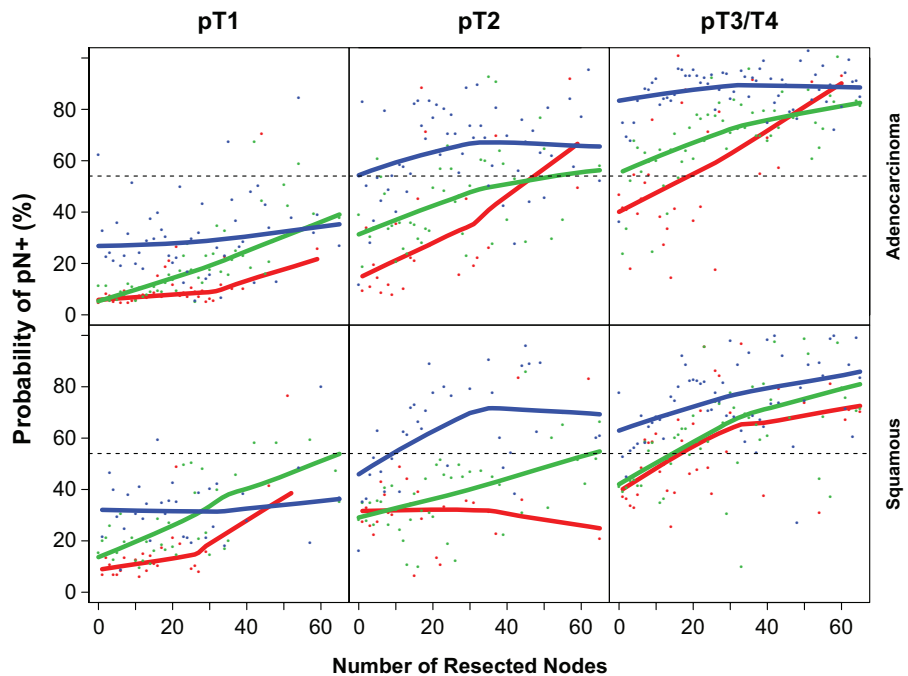

FIGURE 3. Out-of-bootstrap predicted probability of lymph node metastases (pN+) as a function of number of nodes resected, stratified by G, pT, and histopathologic cell type (red = G1, green = G2, and blue = G3). Individual dots represent predicted probabilities and solid lines are locally weighted scatterplot smoothing values of predicted probabilities. Adeno indicates adenocarcinoma; squam, squamous cell cancer.
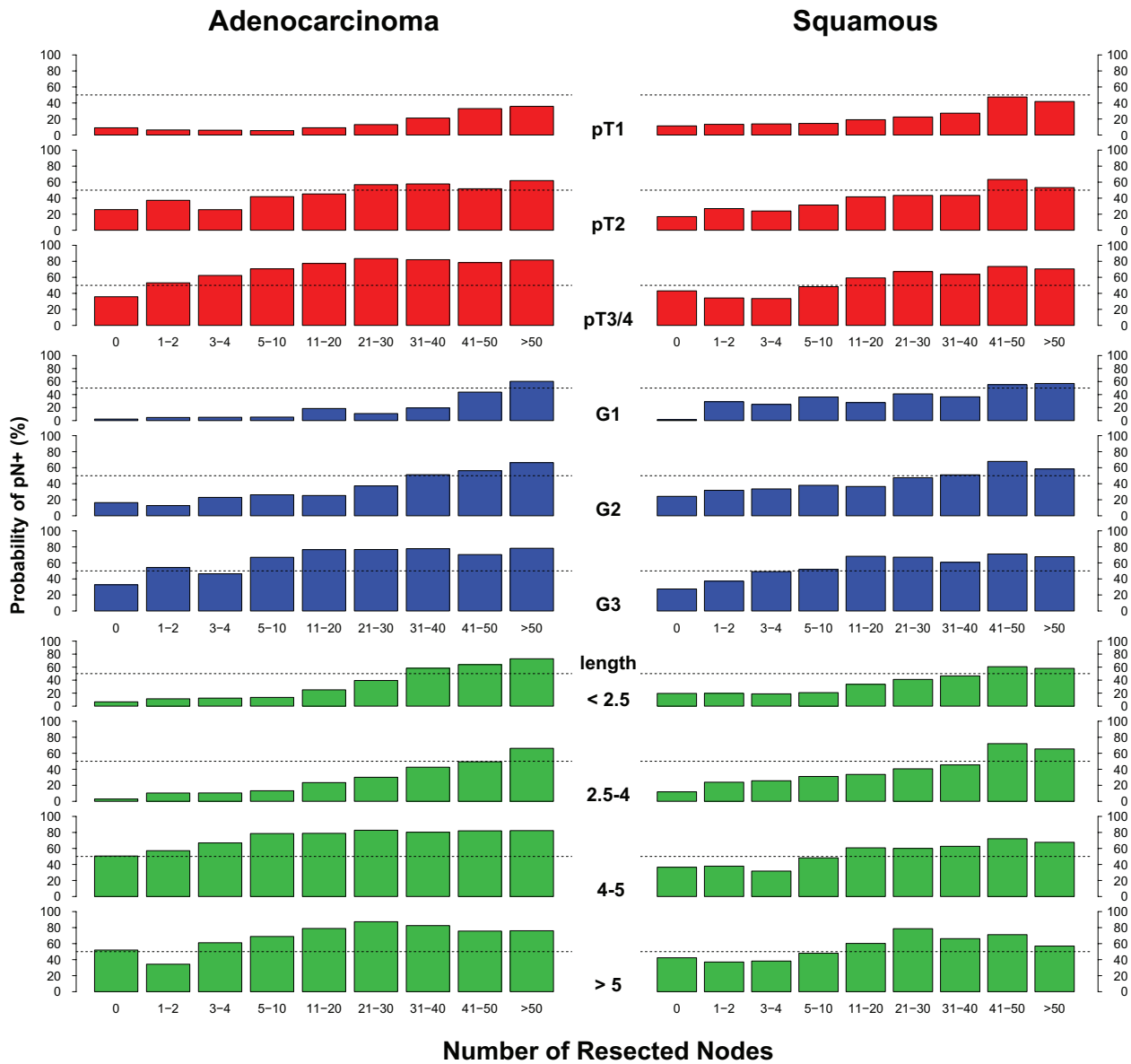
FIGURE 4. Probability of lymph node metastases (pN+) according to number of resected nodes and various esophageal cancer characteristics.

should be discouraged because it is a puzzling blend of cancer biology and surgical technique and offers no further information than revealed by its numerator and denominator.

The Literature

The associations of other cancer characteristics with pN+ are rarely investigated; most publications focus instead on associations with survival. Traditional statistical analysis has demonstrated the associations of T and N.[1] Staging of esophageal cancer in the seventh edition of the AJCC manual added 3 nonanatomic cancer characteristics—cancer location, histologic grade, and histopathologic cell type—providing potentially new associations to examine.[7] It was this effort and the statistical methods behind stage groupings[4] that led us to reexamine associations with pN+.

The overall prevalence of positive lymph nodes in esophageal cancer patients was similar for both histopathologic cell types in 1059 esophagectomy patients at a single institution.[12] Cancer length correlated with pT/ypT, but its association with positive lymph nodes was not evaluated.[13] In a study of 240 patients with esophageal squamous cell carcinoma, histologic grade was 1 of 5 variables univariably associated with positive lymph nodes; however, in multivariable analysis, it did not add additional information.[14] In a multiinstitutional study, a larger number of positive lymph nodes predicted a higher likelihood of distant metastatic cancer[15]: the probability of systemic disease exceeded 50% when 3 or more positive nodes were present, and approached 100% when 8 or more were present. Only 1 study has examined the relationship of pN+ and number of lymph nodes resected.[16] Sensitivity of classifying pN+ continued to improve up to 100 nodes resected; however, maximum increase of
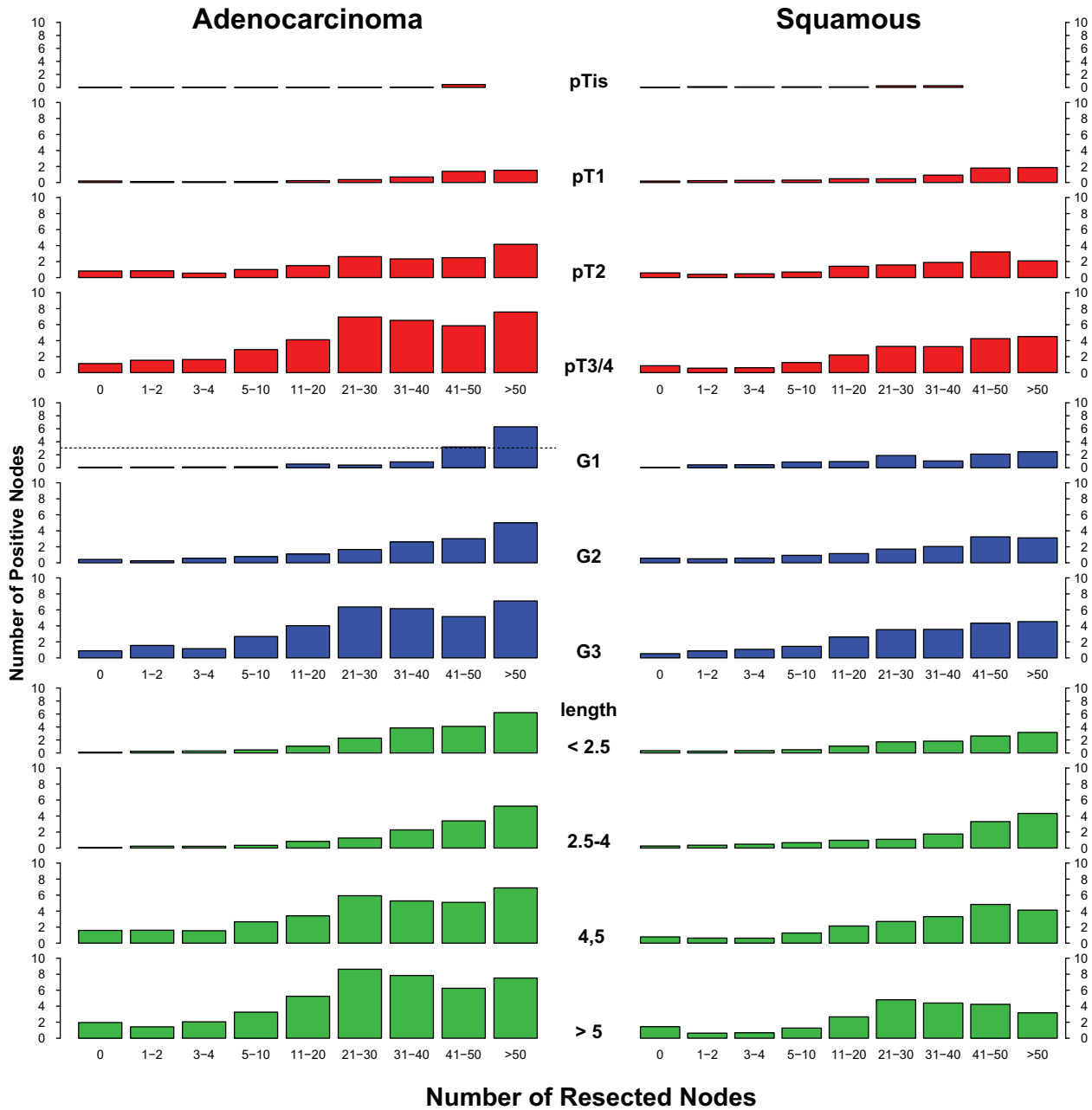
FIGURE 5. Predicted number of positive nodes according to number of resected nodes and various esophageal cancer characteristics.

sensitivity occurred from 0 to 6 nodes, and over 90% sensitivity was reached at 12. Because no further associations were examined, composition of this study group could have influenced the resulting lymphadenectomy recommendations.

Strengths and Limitations

Strengths and limitations of the WECC database have previously been outlined in detail,[2,5] the most notable being that it represents advanced cancers and to a much lesser extent superficial cancers with their detailed subclassifications. This study is based on pathologic staging data rather than clinical staging data, which is necessary to establish the relationships described. However, adding

length, cell type, and differentiation to depth of invasion[1] improves transferability to clinical staging.

An inherent limitation of a dataset such as WECC is lack of detailed information about the location of lymph nodes and variability of not only the surgical lymphadenectomy, but also the pathologic processing of the specimen. This could lead to both over- and underestimation of cancer characteristics. Extreme values may also be influential; inherent in Random Forest methodology is resampling, which mitigates against sampling extreme values, as does averaging over many trees (in this case, over 1000) and multivariable modeling. Institutional and region-of-the-world clustering effects were generally small, but were taken into account by
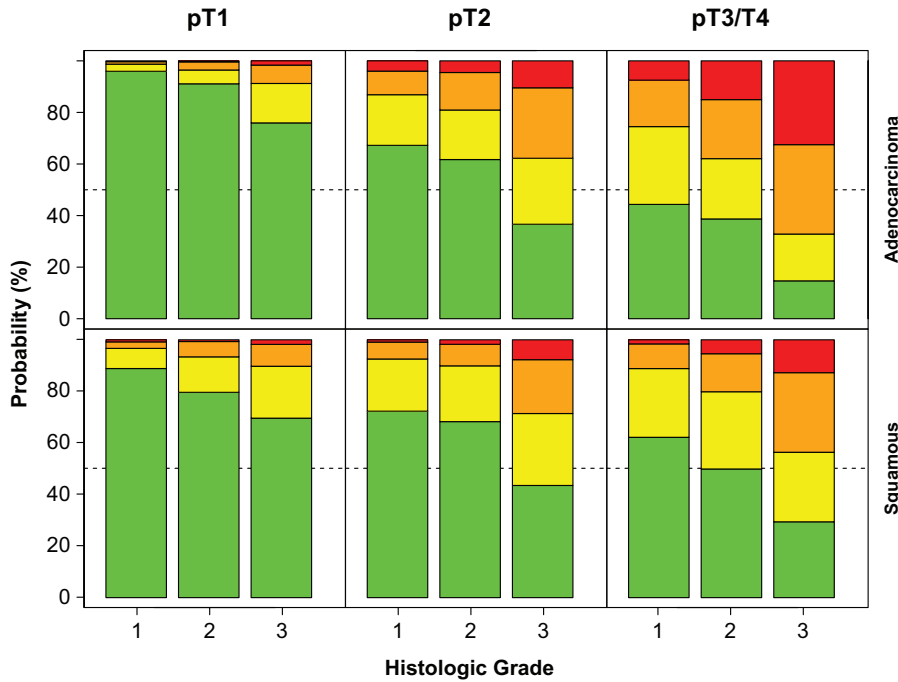
FIGURE 6. Out-of-bootstrap predicted probability of pN0, pN1 (1–2 positive nodes), pN2 (3–6 positive nodes), and pN3 (7 or more positive nodes) cancers, stratified by G, pT, and histopathologic cell type (red = pN0, green = pN1, blue = pN2, and turquoise = pN3). Adeno indicates adenocarcinoma; squam, squamous cell cancer.

incorporating sites and region of the world into risk-adjusted estimates. Nonlinearities in the data and complex interactions among cancer variables that have been identified for esophageal cancer, a challenge for traditional statistical methods, are inherently accounted for by Random Forest methods. These are the primary reasons for choosing this methodology for WECC.

A strength of this study is that it is a large surgery-alone series that spans the spectrum of esophageal cancer characteristics,
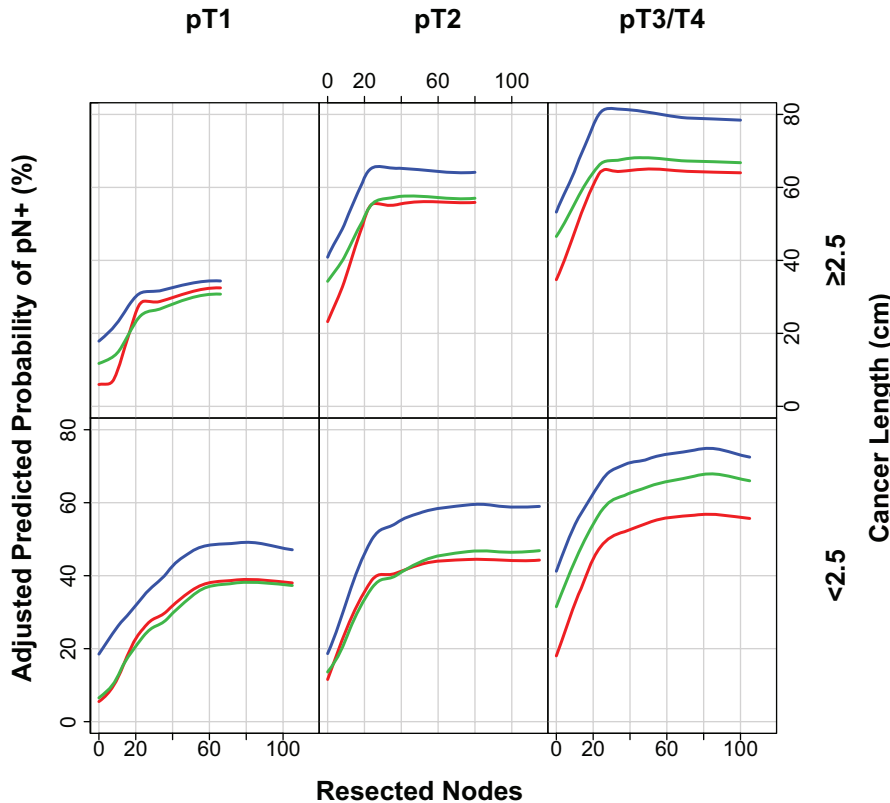

FIGURE 7. Adjusted predicted probability of lymph node metastases (pN+) for various cutoff values of number of resected nodes, according to pT and cancer length. Where curves plateau is interpreted as the lymphadenectomy necessary to accurately detect pN+.

and this will likely never be repeated in the age of neoadjuvant therapy.

Therapeutic Implications

Treatment That Assumes Lymph Node Metastases Status

Accurate N classification is extremely important for therapeutic decisions and prognostication in esophageal cancer. Currently, the only accurate method is lymphadenectomy (pN). Although technology (cN) continues to improve, it is still less accurate than histologic examination of excised lymph nodes. Failure to accurately classify N has implications not only for therapy, but also for clinical trials that use cN for allocation of patients to different trial arms. This may lead to erroneous conclusions.

Endoscopic therapies are recommended when it is reasonable to assume pN0. Absence of lymph node metastases is most likely to be true for well-differentiated short cancers with minimal invasion, and also more likely for squamous cell carcinoma than adenocarcinoma. Strictly adhering to these observations will rarely result in endoscopic mucosal therapy of a pN+ cancer.

Conversely, neoadjuvant therapy typically is recommended when it is reasonable to assume pN+. Realizing the inaccuracies of clinical staging, a deeply invasive long, poorly differentiated cancer reported to be cN0 is likely to be pN+. The high likelihood of a clinical staging error in this patient should prompt the use of neoadjuvant therapy. Improvements in clinical staging of cN are necessary, as illustrated by these examples at the extremes of early and advanced esophageal cancers.

Clinical Implications

The aims of lymphadenectomy are 2-fold: to achieve accurate staging and to provide a possible therapeutic (survival) benefit with respect to extent of lymphadenectomy. These goals are conflicting. This study found that extent of lymphadenectomy required to accurately detect pN+ is greatest for early-stage cancers; in our previous study, extent of lymphadenectomy required to maximize survival benefit is greatest for advanced-stage cancers.[17] Esophagectomy for deeply invasive, long, poorly differentiated (advanced-stage) cancers that are reported as pN0, but with fewer than 20 nodes resected, are unlikely to actually be pN0 because the extent of lymphadenectomy is inadequate for both staging and a possible survival benefit. For superficial, short, well-differentiated (early-stage) cancers, one must balance the risk of extensive lymphadenectomy to accurately detect pN+ (60 nodes) and the benefit of lesser lymphadenectomy (10 nodes for pT1) to maximize 5-year survival.[17] Although "the more one looks, the more one finds," this must be tempered by high-risk associations with pN+.

REFERENCES

1. Rice TW, Zuccaro G Jr, Adelstein DJ, et al. Esophageal carcinoma: depth of tumor invasion is predictive of regional lymph node status. *Ann Thorac Surg*. 1998;65:787–792.
2. Rice TW, Rusch VW, Apperson-Hansen C, et al. Worldwide Esophageal Cancer Collaboration. *Dis Esophagus*. 2009;22:1–8.
3. Breiman L. Random forests. *Mach Learn*. 2001;45:5–32.
4. Ishwaran H, Blackstone EH, Apperson-Hansen C, et al. A novel approach to cancer staging: application to esophageal cancer. *Biostatistics*. 2009;10:603–620.
5. Rice TW, Rusch VW, Ishwaran H, et al. Cancer of the esophagus and esophagogastric junction: data-driven staging for the 7th edition of the American Joint Committee on Cancer/International Union against Cancer Staging Manuals. *Cancer*. 2010;116:3763–3773.
6. Breiman L. Bagging predictors. *Mach Learn*. 1996;24:123–140.
7. Edge SB, Byrd DR, Compton CC, et al. *American Joint Committee on Cancer Staging Manual*. New York, NY: Springer-Verlag; 2010.
8. Ishwaran H, Kogalur UB, Blackstone EH, et al. Random survival forests. *Ann Appl Stat*. 2008;2:841–860.
9. Ishwaran H, Kogalur UB, Gorodeski EZ, et al. High-dimensional variable selection for survival data. *J Am Stat Assoc*. 2010;105:205–217.
10. Ishwaran H, Kogalur UB. Random Forests for Survival, Regression, and Classification (RF-SRC). R Package Version 1.6.0. 2015. Available at: http://cran.r-project.org/web/packages/randomForestSRC/index.html 2015.
11. Rice TW, Blackstone EH. Lymph node ratio: a confounded quotient. *Ann Thorac Surg*. 2013;96:744.
12. Siewert JR, Stein HJ, Feith M, et al. Histologic tumor type is an independent prognostic parameter in esophageal cancer: lessons from more than 1,000 consecutive resections at a single center in the Western world. *Ann Surg*. 2001;234:360–367.
13. Bollschweiler E, Baldus SE, Schroder W, et al. Staging of esophageal carcinoma: length of tumor and number of involved regional lymph nodes. Are these independent prognostic factors? *J Surg Oncol*. 2006;94:355–363.
14. Tajima Y, Nakanishi Y, Ochiai A, et al. Histopathologic findings predicting lymph node metastasis and prognosis of patients with superficial esophageal carcinoma: analysis of 240 surgically resected tumors. *Cancer*. 2000;88:1285–1293.
15. Peyre CG, Hagen JA, DeMeester SR, et al. Predicting systemic disease in patients with esophageal cancer after esophagectomy: a multinational study on the significance of the number of involved lymph nodes. *Ann Surg*. 2008;248:979–985.
16. Dutkowski P, Hommel G, Bottger T, et al. How many lymph nodes are needed for an accurate pN classification in esophageal cancer? Evidence for a new threshold value. *Hepatogastroenterology*. 2002;49:176–180.
17. Rizk NP, Ishwaran H, Rice TW, et al. Optimum lymphadenectomy for esophageal cancer. *Ann Surg*. 2010;251:46–50.