Local recurrence after surgery for non–small cell lung cancer: A recursive partitioning analysis of multi-institutional data

Chris R. Kelsey, MD,[a] Kristin A. Higgins, MD,[b] Bercedis L. Peterson, PhD,[c] Junzo P. Chino, MD,[a] Lawrence B. Marks, MD,[d] Thomas A. D'Amico, MD,[e] and John M. Varlotto, MD[f]

Objective: To define subgroups at high risk of local recurrence (LR) after surgery for non–small cell lung cancer using a recursive partitioning analysis (RPA).

Methods: This Institutional Review Board–approved study included patients who underwent upfront surgery for I-IIIA non–small cell lung cancer at Duke Cancer Institute (primary set) or at other participating institutions (validation set). The 2 data sets were analyzed separately and identically. Disease recurrence at the surgical margin, ipsilateral hilum, and/or mediastinum was considered an LR. Recursive partitioning was used to build regression trees for the prediction of local recurrence-free survival (LRFS) from standard clinical and pathological factors. LRFS distributions were estimated with the Kaplan-Meier method.

Results: The 1411 patients in the primary set had a 5-year LRFS rate of 77% (95% confidence interval [CI], 0.74-0.81), and the 889 patients in the validation set had a 5-year LRFS rate of 76% (95% CI, 0.72-0.80). The RPA of the primary data set identified 3 terminal nodes based on stage and histology. These nodes and their 5-year LRFS rates were as follows: (1) stage I/adenocarcinoma, 87% (95% CI, 0.83-0.90); (2) stage I/squamous or large cell, 72% (95% CI, 0.65-0.79); and (3) stage II-IIIA, 62% (95% CI, 0.55-0.69). The validation RPA identified 3 terminal nodes based on lymphovascular invasion (LVI) and stage: (1) no LVI/stage IA, 82% (95% CI, 0.76-0.88); (2) no LVI/stage IB-IIIA, 73% (95% CI, 0.69-0.80); and (3) LVI, 58% (95% CI, 0.47-0.69).

Conclusions: The risk of LR was similar in the primary and validation patient data sets. There was discordance between the 2 data sets regarding the clinical factors that best segregate patients into risk groups. (J Thorac Cardiovasc Surg 2013;146:768-73)

Supplemental material is available online.

An accurate understanding of risk (eg, risk of recurrence) is essential in the field of oncology. Estimates of risk guide the development of clinical trials exploring alternative treatment strategies but are also used when considering treatment programs for individual patients. In specialties that deal with local modalities, more precise risks are particularly helpful (eg, risk of *local* recurrence). This is complicated statistically because of the issue of competing risks.[1] Nonetheless, possessing a reasonable appreciation of such is critical in surgical and radiation oncology practice.

Lung cancer remains the leading cause of cancer mortality in the United States.[2] Disease recurrence resulting in death is common despite complete surgical resection, even in patients with early-stage disease. Recurrences are generally subdivided into those developing at local sites (surgical margin and regional draining lymph nodes) and those developing at distant sites. Although adjuvant chemotherapy can potentially decrease the risk of both local and distant recurrence, the risk of local recurrence can also be reduced with the use of postoperative radiation therapy.

Understanding the risk of local recurrence in resectable non–small cell lung cancer (NSCLC) has 4 primary challenges. First, although overall recurrence rates are generally reported in randomized lung cancer trials, patterns of failure are often not described. This can lead to a general unawareness of recurrence patterns, which has relevance to the choice of adjuvant treatment modalities. Second, various definitions of local recurrence have been used in the literature.[3] This creates difficulty when one compares rates of local recurrence among different studies.[3,4] Third, although multiple risk factors for local recurrence have been observed using multivariate modeling, there are inconsistencies between studies. Finally, estimating the aggregate risk in

From the Department of Radiation Oncology,[a] the Cancer Institute Biostatistics,[c] and the Division of Thoracic Surgery,[e] Department of Surgery, Duke Cancer Institute, Durham, NC; the Department of Radiation Oncology,[b] Emory University, Atlanta, Ga; the Department of Radiation Oncology,[d] University of North Carolina, Chapel Hill, NC; and the Department of Radiation Oncology,[f] Penn State Hershey Cancer Institute, Hershey, Pa.

Disclosures: Authors have nothing to disclose with regard to commercial support.
Presented at the 55th Annual Meeting of the American Society for Therapeutic Radiology and Oncology, October 28-31, 2013, Boston, Mass.
Received for publication Feb 20, 2013; revisions received April 16, 2013; accepted for publication May 10, 2013; available ahead of print July 15, 2013.
Address for reprints: Chris R. Kelsey, MD, Department of Radiation Oncology, Duke Cancer Institute, DUMC 3085, Durham, NC 27710 (E-mail: kelse003@mc.duke.edu).
0022-5223/$36.00

http://dx.doi.org/10.1016/j.jtcvs.2013.05.041

Abbreviations and Acronyms

CI = confidence interval
HR = hazard ratio
LR = local recurrence
LRFS = local recurrence-free survival
LVI = lymphovascular invasion
NSCLC = non–small cell lung cancer
RPA = recursive partitioning analysis

an individual patient, or population, based on the presence or absence of numerous potential factors is a challenge.

With these issues in mind, using 2 independent databases of lung cancer patients undergoing surgery for NSCLC, we estimated the distribution of local recurrence-free survival (LRFS) (using a practical definition of local recurrence), performed multivariate analyses to assess risk factors for local recurrence, and finally performed recursive partitioning to better understand the risk of local recurrence in defined cohorts of lung cancer patients.

METHODS

The retrospective collection of data for this study was approved by the Institutional Review Board of each of the individual participating institutions. The primary data set included all patients who underwent upfront surgery for I-IIIA NSCLC at Duke University (Durham, NC) between 1995 and 2008. The validation data set included patients from the following institutions: University of North Carolina (Chapel Hill, NC), Penn State Hershey Cancer Institute (Hershey, Pa), Beth Israel Deaconess Medical Center (Boston, Mass), and the Veterans Administration hospitals in Boston, Mass, and Denver, Colo, including patients who underwent surgery between 1996 and 2008. For both data sets, patients who received preoperative chemotherapy and/or radiation therapy, presented with synchronous primary tumors, or died in the immediate postoperative period (30 days) were excluded. Patients with superior sulcus tumors or chest wall invasion were excluded because their patterns of local recurrence are different than patients with disease confined to the lung parenchyma and regional lymph nodes. Because the primary objective of the study was to evaluate the risk of local recurrence, we also excluded patients who had positive surgical margins or who received adjuvant postoperative radiation therapy. Because the effect of chemotherapy on local recurrence is not clear, patients receiving adjuvant chemotherapy were included.

Disease recurrence at the surgical resection margin, ipsilateral hilum, and/or mediastinum was defined as a *local recurrence*. This definition was used because it encompasses the anatomic sites that are included in a typical postoperative radiation field. All other sites of recurrence, including the supraclavicular fossa and contralateral hilum, were considered distant recurrences. Patterns of recurrence were assessed by follow-up imaging studies supplemented with invasive procedures, such as bronchoscopy, as clinically indicated.

Statistical Analyses

The primary and validation data sets were analyzed separately and identically. Recursive partitioning was used to fit regression trees for the prediction of LRFS, defined as the time from surgery to local recurrence, with distant recurrences ignored and deaths censored. The candidate predictors for both trees were sex, age, surgery type (wedge/segmentectomy vs lobectomy/pneumonectomy), histology (squamous/adenosquamous cell carcinoma, large-cell carcinoma, adenocarcinoma, and NSCLC not otherwise specified), lymphovascular space invasion (yes/no), pleural invasion (yes/no), number of hilar lymph nodes sampled, stage using the American Joint Committee on Cancer, 7th edition, classification (IA, IB, IIA, IIB, or IIIA, treated as a continuous variable), and adjuvant chemotherapy (yes/no). Other predictors were not included in the analysis, either because they had many missing values in the secondary data set (eg, surgical approach [open vs thoracoscopic] and grade) or because they were measures of the same construct as pathologic stage and were highly correlated with pathologic stage (eg, T-stage, N-stage, size of the primary tumor, and number of hilar nodes involved). We used stage as the candidate predictor instead of one of its correlates because stage had by far the largest univariate association with LRFS. Extent of mediastinal lymph node sampling was not included because the available data in the 2 databases were different (number of stations sampled vs number of lymph nodes sampled).

Recursive partitioning is a statistical method that groups patients into distinct cohorts based on maximizing the value of log-rank tests for the clinical end point of interest (in this case, LRFS). The first 2 cohorts are defined by assessing all possible dichotomizations of all predictor variables, whether categorical or continuous, to find the one dichotomization that produces the largest log-rank test statistic. The method then repeats this assessment within each of these 2 cohorts so that 1 of these 2 cohorts is further split into 2 smaller subgroups. The method proceeds in this manner until a complex stopping rule is met. For each cohort, 5-year LRFS rates are estimated.

Recursive partitioning was done with R's rcart function (The R Foundation for Statistical Computing, Vienna, Austria). Overgrown trees were developed using 10 cross validations and then pruned down to a selected number of nodes. To select the number of nodes to retain, we plotted the mean of cross-validation errors from models of all sizes against their corresponding ''complexity parameters.'' We then noted the lowest-lying point in the plot (or, in practice, the leftmost of 2 or 3 similar low-lying points) and noted its complexity parameter value. Because each complexity parameter value is uniquely linked to a given number of nodes, the selected complexity parameter value determines the recommended number of nodes. For both trees, either 2 or 3 nodes were considered appropriate using this procedure.

The Kaplan-Meier product limit method was used to estimate 5-year LRFS rates. The proportional hazards model was used to determine the predictors of LRFS. By using the same 9 predictor variables as were used in the recursive partitioning analyses (RPAs), the model was fit by using a backwards elimination procedure with a significance level to stay in the model of 0.40.[5] Predictors that were retained in the model were not assessed for statistical significance, but their *P* values were used to show the strength of evidence against the null hypothesis.

RESULTS

The primary data set included 1411 patients, 199 of whom developed an LR, whereas the validation data set included 889 patients, 146 of whom developed an LR. The median follow-up among patients without LR was 26 months (range, 3 days to 175 months) and 33 months (range, 6 days to 175 months) in the two data sets, respectively. The primary and validation data sets had similar LRFS distributions (Figure E1). The data sets had 5-year LRFS rates of 77% (95% confidence interval [CI], 0.74-0.81) and 76% (95% CI, 0.72-0.80), respectively. Patient characteristics and surgical/pathological details are found in Table E1. The 2 cohorts were generally similar, although there were statistically significant differences in many of the factors given the many patients included in the analysis, despite identical inclusion and exclusion criteria.

GTS

TABLE 1. Hazard ratios from the multivariate analyses of local recurrence-free survival

	Primary data set (n = 1411)			Secondary data set (n = 889)		
Predictors*	**Hazard ratio**	**95% CI**	***P* value**	**Hazard ratio**	**95% CI**	***P* value**
Age (5-y increase)	0.91	0.84-0.99	.02	0.95	0.88-1.03	.29
Sex						
Female						
Male	1.13	0.85-1.52	.39	1.13	0.79-1.60	.51
Surgical procedure						
Lobectomy/pneumonectomy						
Wedge/segmentectomy	1.92	1.36-2.73	<.001	1.52	0.89-2.60	.12
No. of hilar nodes sampled (per 1 increase in node)	0.98	0.94-1.03	.43	0.98	0.95-1.01	.22
Histology						
Adenocarcinoma/NSCLC NOS						
Squamous/large-cell	1.75	1.31-2.34	<.001	1.43	1.02-2.00	.04
lymphovascular space invasion						
No/not stated						
Yes	1.32	0.95-1.83	.10	2.27	1.58-3.27	<.001
Visceral pleural invasion						
No/not stated						
Yes	1.78	1.30-2.43	<.001	1.24	0.77-1.98	.38
Stage (per 1 increase in unit)	1.40	1.22-1.61	<.001	1.32	1.13-1.54	<.001
Adjuvant chemotherapy						
No						
Yes	0.60	0.37-0.96	.03	1.66	1.05-2.61	.03

CI, Confidence interval; *NSCLC*, non–small cell lung cancer; *NOS*, not otherwise specified. *For dichotomous predictors, the first value listed is the reference value (ie, the numerator of the hazard ratio).

Predictors of LR

In the proportional hazards model, the size of the association between predictors of local recurrence depended on the data set (Table 1). In the primary data set, the factors most strongly associated with shorter LRFS included sublobar resection, compared with lobectomy/pneumonectomy (hazard ratio [HR], 1.92; 95% CI, 1.36-2.73); squamous or large-cell histology, compared with adenocarcinoma/NSCLC not otherwise specified (HR, 1.75; 95% CI, 1.31-2.34); visceral pleural invasion, compared with no pleural invasion (HR, 1.78; 95% CI, 1.30-2.43); and higher stage (HR for a 1-unit increase in stage, 1.40; 95% CI, 1.22-1.61). All 4 of these associations had *P* values < .001.

In the validation data set, the factors most strongly associated with a shorter LRFS included lymphovascular space invasion (LVI) compared with no invasion (HR, 2.29; 95% CI, 1.60-3.29) and higher stage (HR for a 1-unit increase in stage, 1.29; 95% CI, 1.11-1.50) (Table 1). Both of these associations had *P* values < .001.

Recursive Partitioning Analysis

The primary RPA identified 3 terminal nodes based on stage and histology: stage I/adenocarcinoma or NSCLC not otherwise specified, stage I/squamous or large cell, and stage II-IIIA (Figure 1, *A*). The 5-year LRFS rates (and 95% CIs) for each node were 87% (95% CI, 0.83-0.90), 72% (95% CI, 0.65-0.79), and 62% (95% CI, 0.55-0.69), respectively (Figure 1, *B*).

The validation RPA identified 3 terminal nodes based on LVI and stage: no LVI/stage IA, no LVI/stage IB-IIIA, and LVI (Figure 2, *A*). The 5-year LRFS rates (and 95% CIs) for each node were 82% (95% CI, 0.76-0.88), 73% (95% CI, 0.69-0.80), and 58% (95% CI, 0.47-0.69), respectively (Figure 2, *B*).

DISCUSSION

By using a definition of local recurrence that includes those anatomic sites that are included in a typical postoperative radiation field, the risk of local recurrence was the same in the primary and validation data sets. The 5-year LRFS rates were 77% and 76%, emphasizing that local recurrence remains an obstacle for cure, even in resected NSCLC. Given the negative impact of conventional postoperative radiation therapy in an unselected population of patients with early-stage NSCLC,[6] there is a need to identify and validate a population of patients at highest risk who would be optimal for further study of postoperative radiation therapy.

By using standard clinical and pathological factors, we examined whether recursive partitioning would stratify patients into risk groups with widely varying rates of local recurrence. Although recursive partitioning did segregate patients into distinct groups, the size of the risk differences between the RPA nodes was smaller than we had hoped, on the order of approximately a 10% to 20% absolute difference in risk. Such differences might be too small to

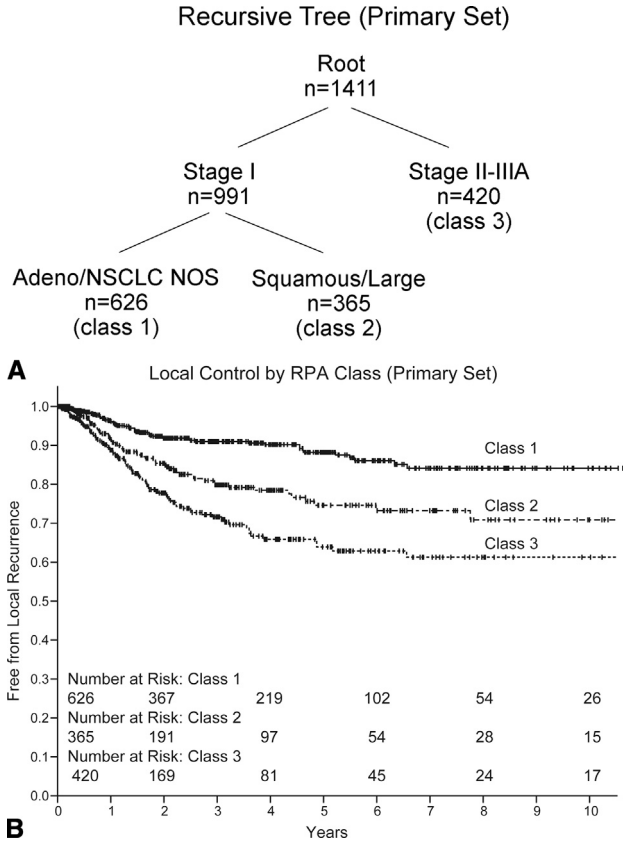

FIGURE 1. A, Results of the recursive partitioning analysis (*RPA*) of the primary data set. B, Five-year local recurrence-free survival rates are shown according to RPA node. *NSCLC*, Non–small cell lung cancer; *NOS*, not otherwise specified.

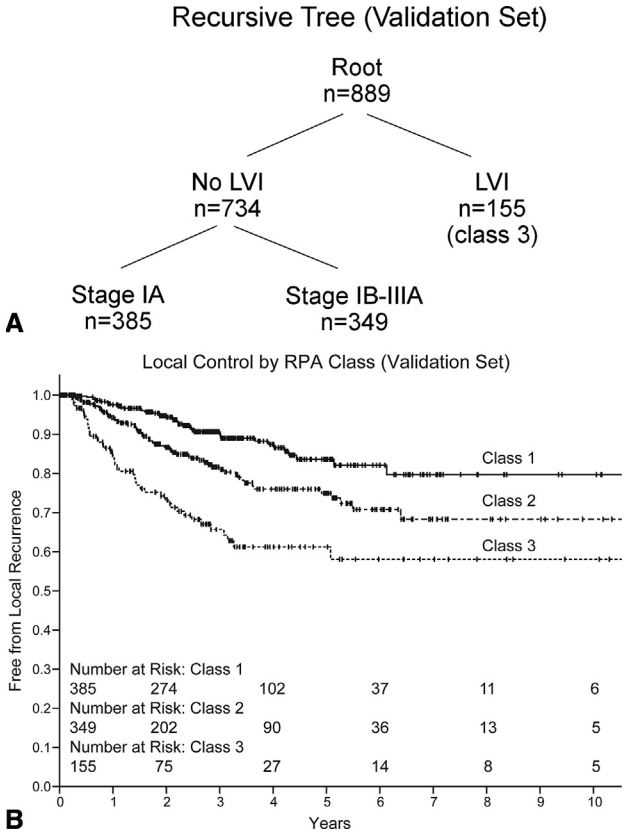

FIGURE 2. A, Results of the recursive partitioning analysis (*RPA*) of the validation data set. B, Five-year local recurrence-free survival rates are shown according to RPA node. *LVI*, Lymphovascular invasion.

optimally identify patients who can benefit most from postoperative radiation therapy. Another disappointment was that the resulting trees differed between the 2 data sets. This may be due, in part, to the large correlation among the risk factors, because multicollinearity is known to produce unstable results.

Although a defined population of patients with resected NSCLC was identified prospectively to be included in the study, there were some differences between the 2 groups of patients (Table E1), perhaps reflecting geographic variations in disease presentation, tumor characteristics, and treatment approaches. Furthermore, despite a consistent definition of local recurrence, the multivariate models of local recurrence differed between the data sets. This was also reflected in the disparate results obtained with recursive partitioning, although both data sets implicated pathologic stage as being a major predictive factor.

Prior studies have identified several potential risk factors for local recurrence. Many relate to the extent and quality of surgery. These include more limited surgical procedures (wedge or segmentectomy),[4,7-13] positive surgical margins,[10,14-16] and lack of or limited mediastinal sampling.[11,17,18] Other factors relate to the pathologic findings, including stage,[10,12,16] extent of regional lymph node involvement,[14,19] size of the primary tumor,[10,13,20] squamous histology,[12,21] visceral pleural invasion,[13,19] and lymphovascular space invasion.[4,12,20] Patient-related factors, such as smoking history[18] and comorbid disease,[22] may also be contributory. Randomized studies have shown that adjuvant radiation therapy decreases the risk of local recurrence but may negatively affect long-term survival, at least for an unselected population with early-stage disease.[16,23-26] The effect of chemotherapy on local recurrence is not clear.[22,27,28]

Although multiple surgical and pathologic factors, previously outlined, have been associated with a higher risk of local recurrence after surgery for NSCLC, these factors have not been consistent across studies (Table 2). Furthermore, most studies have not performed comprehensive multivariate modeling to identify factors independently associated with local recurrence. Thus, consistently defining cohorts at highest risk of local recurrence in NSCLC has proven elusive.

There are limitations with our analysis. Patients were treated over many years at multiple institutions and undoubtedly there has been an evolution in surgical

TABLE 2. Studies using multivariate modeling to assess risk factors for local recurrence

Risk factor	Kelsey[12] (n = 975)	Hung[18] (n = 756)	Varlotto[22] (n = 373)	Saynak[20] (n = 335)	Lee[10] (n = 211)	Lopez Guerra[13] (n = 1402)
Patient characteristics						
Age	−	−	−		−	−
Sex	−	−		−	−	−
Diabetes	−		+			
Smoking history		+	−			
PET-CT staging			−			
Surgical characteristics						
Open/VATS	−			−		
Sublobar resection	+		−		+	+
Positive margins	−		−		+	
Lymph node sampling	−	+	−			
Pathologic characteristics						
Right vs left	−	−				
Stage	+		−	−	−	−
Extent of regional lymph node involvement			−		−	
Tumor size	−		−	+	−	+
Histologic grade			−	−	−	−
Squamous histology	+	−	−	−	−	
Visceral pleural invasion	−	−	−	−		+
Lymphovascular space invasion	−		+	+		−
Adjuvant therapy						
Radiation therapy	−					
Chemotherapy	−		+			−

Some studies evaluated numerous additional factors (all positive factors listed in the table). +, Associated with local recurrence; −, not associated with local recurrence; *blank cells*, the factor was not assessed; *PET*, positron emission tomography; *CT*, computed tomography; *VATS*, video-assisted thoracoscopic surgery.

approaches over the time period. It is also likely there were differences among the institutions regarding surgical expertise and technique, pathological assessment and interpretation, including histologic classification,[29] lymph node identification and ascertainment of involvement,[30] and/or reporting of lymphovascular space invasion and visceral pleural invasion. Serial imaging to assess for local recurrence was not standardized between or within centers. Thus, some local recurrences were probably not identified. Furthermore, local recurrences were largely identified using imaging studies that have limitations with both sensitivity and specificity. Finally, only those factors that were available from each institution's database were studied. It is possible other clinical or pathological factors, such as comorbidities or smoking status, could be important. Nonetheless, this endeavor is one of the largest studies evaluating local recurrence in operable NSCLC and included data from multiple institutions.

This study, in concert with previous analyses evaluating risk factors for local recurrence, demonstrates that, although standard clinical and pathologic factors can be useful, they are unlikely to provide optimal discrimination between low- and high-risk cohorts. More refined methods to assess risk are necessary before an enriched population of patients can be identified for further study of postoperative radiation therapy. Whether biomarkers or genomic signatures will help in this regard is, at present, unknown, but should be explored. Furthermore, more customized and conformal radiation fields than have been historically used may be advantageous for patients with early-stage disease, as was shown in a randomized trial that showed a benefit for postoperative radiation therapy in stage I NSCLC.[26] We advocate good surgical and pathologic practices that may decrease the risk of local recurrence, including rigorous preoperative staging, complete resection with negative margins, adequate lymph node sampling/dissection and examination, and multidisciplinary assessment when appropriate.

CONCLUSIONS

By using 2 large independent data sets, we observed discordance regarding the parameters that best segregate patients into risk groups of local recurrence, although stage appears to be an important factor. Furthermore, although recursive partitioning is able to define subgroups at variable risk of local recurrence, the differences in the risk groups are modest. More precise methods of assessing risk are necessary to optimally define high-risk patients.

References

1. Gelman R, Gelber R, Henderson IC, Coleman CN, Harris JR. Improved methodology for analyzing local and distant recurrence. *J Clin Oncol*. 1990;8:548-55.
2. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. *CA Cancer J Clin*. 2010;60:277-300.

3. Kelsey CR, Marks LB, Wilson LD. Postoperative radiation therapy for lung cancer: where do we stand? *Oncology (Williston Park)*. 2008;22:301-10; discussion 10, 14-5, 19.
4. Varlotto JM, Recht A, Flickinger JC, Medford-Davis LN, Dyer AM, DeCamp MM. Varying recurrence rates and risk factors associated with different definitions of local recurrence in patients with surgically resected, stage I non-small cell lung cancer. *Cancer*. 2010;116:2390-400.
5. Harrell FE. *Regression Modeling Strategies With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York, NY: Springer; 2001.
6. PORT Meta-analysis Trialists Group. Postoperative radiotherapy in non-small-cell lung cancer: systematic review and meta-analysis of individual patient data from nine randomised controlled trials. *Lancet*. 1998;352:257-63.
7. Ginsberg RJ, Rubinstein LV, Lung Cancer Study Group. Randomized trial of lobectomy versus limited resection for T1 N0 non-small cell lung cancer. *Ann Thorac Surg*. 1995;60:615-22; discussion 22-3.
8. Harpole DH Jr, Herndon JE II, Young WG Jr, Wolfe WG, Sabiston DC Jr. Stage I nonsmall cell lung cancer: a multivariate analysis of treatment methods and patterns of recurrence. *Cancer*. 1995;76:787-96.
9. Warren WH, Faber LP. Segmentectomy versus lobectomy in patients with stage I pulmonary carcinoma: five-year survival and patterns of intrathoracic recurrence. *J Thorac Cardiovasc Surg*. 1994;107:1087-93; discussion 93-4.
10. Lee JH, Machtay M, Kaiser LR, Friedberg JS, Hahn SM, McKenna MG, et al. Non-small cell lung cancer: prognostic factors in patients treated with surgery and postoperative radiation therapy. *Radiology*. 1999;213:845-52.
11. Martini N, Bains MS, Burt ME, Zakowski MF, McCormack P, Rusch VW, et al. Incidence of local recurrence and second primary tumors in resected stage I lung cancer. *J Thorac Cardiovasc Surg*. 1995;109:120-9.
12. Kelsey CR, Marks LB, Hollis D, Hubbs JL, Ready NE, D'Amico TA, et al. Local recurrence after surgery for early stage lung cancer: an 11-year experience with 975 patients. *Cancer*. 2009;115:5218-27.
13. Lopez Guerra JL, Gomez DR, Lin SH, Levy LB, Zhuang Y, Komaki R, et al. Risk factors for local and regional recurrence in patients with resected N0-N1 non-small-cell lung cancer, with implications for patient selection for adjuvant radiation therapy. *Ann Oncol*. 2013;24:67-74.
14. Hsu HC, Wang CJ, Huang EY, Sun LM. Post-operative adjuvant thoracic radiotherapy for patients with completely resected non-small cell lung cancer with nodal involvement: outcome and prognostic factors. *Br J Radiol*. 2004;77:43-8.
15. Sawyer TE, Bonner JA, Gould PM, Foote RL, Deschamps C, Trastek VF, et al. Effectiveness of postoperative irradiation in stage IIIA non-small cell lung cancer according to regression tree analyses of recurrence risks. *Ann Thorac Surg*. 1997; 64:1402-7; discussion 1407-8.
16. Sawyer TE, Bonner JA, Gould PM, Foote RL, Deschamps C, Trastek VF, et al. The impact of surgical adjuvant thoracic radiation therapy for patients with non-small cell lung carcinoma with ipsilateral mediastinal lymph node involvement. *Cancer*. 1997;80:1399-408.
17. Lardinois D, Suter H, Hakki H, Rousson V, Betticher D, Ris HB. Morbidity, survival, and site of recurrence after mediastinal lymph-node dissection versus systematic sampling after complete resection for non-small cell lung cancer. *Ann Thorac Surg*. 2005;80:268-74; discussion 274-5.
18. Hung JJ, Jeng WJ, Hsu WH, Chou TY, Huang BS, Wu YC. Predictors of death, local recurrence, and distant metastasis in completely resected pathological stage-I non-small-cell lung cancer. *J Thorac Oncol*. 2012;7:1115-23.
19. Fujimoto T, Cassivi SD, Yang P, Barnes SA, Nichols FC, Deschamps C, et al. Completely resected N1 non-small cell lung cancer: factors affecting recurrence and long-term survival. *J Thorac Cardiovasc Surg*. 2006;132: 499-506.
20. Saynak M, Veeramachaneni NK, Hubbs JL, Nam J, Qaqish BF, Bailey JE, et al. Local failure after complete resection of N0-1 non-small cell lung cancer. *Lung Cancer (Amsterdam, Netherlands)*. 2011;71:156-65.
21. Sawyer TE, Bonner JA, Gould PM, Foote RL, Deschamps C, Lange CM, et al. Factors predicting patterns of recurrence after resection of N1 non-small cell lung carcinoma. *Ann Thorac Surg*. 1999;68:1171-6.
22. Varlotto JM, Recht A, Flickinger JC, Medford-Davis LN, Dyer AM, Decamp MM. Factors associated with local and distant recurrence and survival in patients with resected nonsmall cell lung cancer. *Cancer*. 2009;115: 1059-69.
23. The Lung Cancer Study Group. Effects of postoperative mediastinal radiation on completely resected stage II and stage III epidermoid cancer of the lung. *N Engl J Med*. 1986;315:1377-81.
24. Feng QF, Wang M, Wang LJ, Yang ZY, Zhang YG, Zhang DW, et al. A study of postoperative radiotherapy in patients with non-small-cell lung cancer: a randomized trial. *Int J Radiat Oncol Biol Phys*. 2000;47:925-9.
25. Mayer R, Smolle-Juettner FM, Szolar D, Stuecklschweiger GF, Quehenberger F, Friehs G, et al. Postoperative radiotherapy in radically resected non-small cell lung cancer. *Chest*. 1997;112:954-9.
26. Trodella L, Granone P, Valente S, Valentini V, Balducci M, Mantini G, et al. Adjuvant radiotherapy in non-small cell lung cancer with pathological stage I: definitive results of a phase III randomized trial. *Radiother Oncol*. 2002; 62:11-9.
27. Arriagada R, Dunant A, Pignon JP, Bergman B, Chabowski M, Grunenwald D, et al. Long-term results of the international adjuvant lung cancer trial evaluating adjuvant cisplatin-based chemotherapy in resected lung cancer. *J Clin Oncol*. 2010;28:35-42.
28. Douillard JY, Rosell R, De Lena M, Carpagnano F, Ramlau R, Gonzales-Larriba JL, et al. Adjuvant vinorelbine plus cisplatin versus observation in patients with completely resected stage IB-IIIA non-small-cell lung cancer (Adjuvant Navelbine International Trialist Association [ANITA]): a randomised controlled trial. *Lancet Oncol*. 2006;7:719-27.
29. Roggli VL, Vollmer RT, Greenberg SD, McGavran MH, Spjut HJ, Yesner R. Lung cancer heterogeneity: a blinded and randomized study of 100 consecutive cases. *Hum Pathol*. 1985;16:569-79.
30. Ramirez RA, Wang CG, Miller LE, Adair CA, Berry A, Yu X, et al. Incomplete intrapulmonary lymph node retrieval after routine pathologic examination of resected lung cancer. *J Clin Oncol*. 2012;30:2823-8.

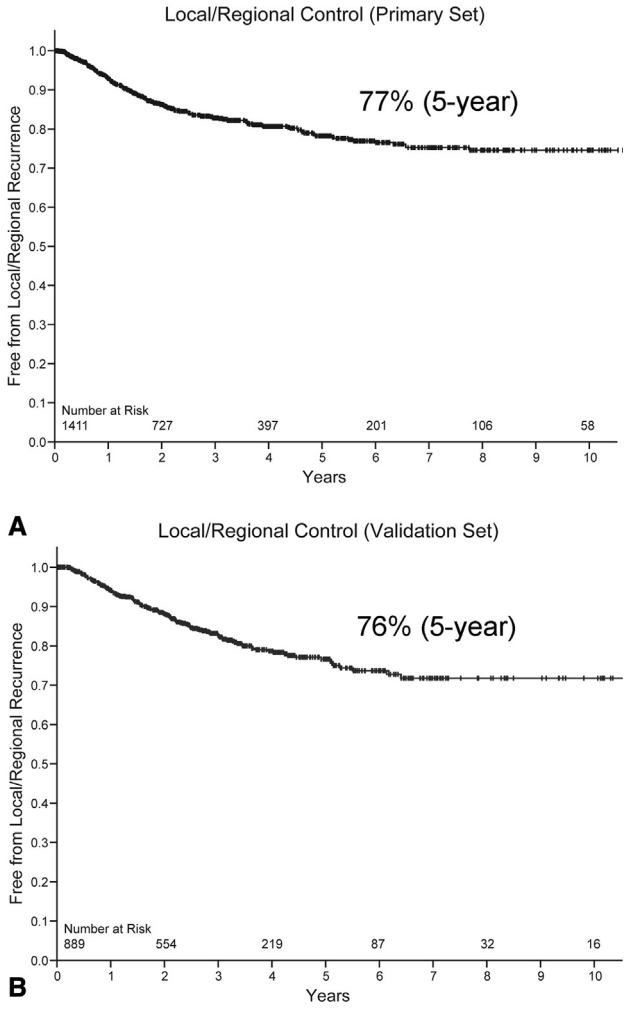

FIGURE E1. Kaplan-Meier curves of local recurrence-free survival in the primary data set (A) and validation data set (B).

TABLE E1. Patient and treatment characteristics

Characteristic	Primary data set (n = 1411)		Validation data set (n = 889)		*P* value
Factors used in multivariate and recursive partitioning analyses					
Age, y					
Median	68		67		.09
Range	21-93		35-96		
Unknown			2		
Sex					
Male	747	53	534	60	<.001
Female	664	47	353	40	
Unknown			2	0.2	
Surgical procedure					
Wedge/segmentectomy	285	20	103	12	<.001
Lobectomy*	1048	74	706	79	
Pneumonectomy	78	6	80	9	
Hilar lymph node sampling					
Yes	1150	82	770	87	.001
No	261	18	119	13	
Histology					
Adenocarcinoma	682	48	429	48	<.001
Squamous cell†	540	39	377	42	
Large cell	87	6	43	5	
NSCLC NOS	102	7	40	4	
Lymphovascular invasion					
Yes	297	21	155	17	.03
No/not stated	1113	79	734	83	
Visceral pleural invasion					
Yes	314	22	95	11	<.001
No/not stated	1097	78	794	89	
Pathologic stage					
IA	609	43	426	48	.006
IB	382	27	203	23	
IIA	231	16	153	17	
IIB	108	8	78	9	
IIIA	81	6	29	3	
Adjuvant chemotherapy					
Yes	141	10	106	12	.14
No	1270	90	783	88	
Factors not used in multivariate and recursive partitioning analyses					
Mediastinal lymph node sampling					
Yes	1234	87	694	78	<.001
No	177	13	195	22	
Histologic differentiation					
Well	130	9	136	15	<.001
Moderate	641	45	419	47	
Poor	463	33	286	32	
Unknown	177	13	48	6	
Size, cm‡					.91
Median	2.6		2.7		
Range	0.5-14		0.3-15		

Values are given as number (%) unless otherwise indicated. *NSCLC*, Non–small cell lung cancer; *NOS*, not otherwise specified. *Including bilobectomy and sleeve lobectomy. †Including adenosquamous (n = 15 in primary and n = 20 in validation cohorts). ‡There were 4 and 9 missing values on size in the primary and validation cohorts, respectively.