Challenges in Predicting Recurrence After Resection of Node-Negative Non-Small Cell Lung Cancer

Lucas W. Thornblade, MD, MPH, **Michael S. Mulligan, MD**, **Katherine Odem-Davis, PhD**, **Billanna Hwang, MPH, DHSc**, **Rachel L. Waworuntu, BS**, **Erika M. Wolff, PhD**, **Larry Kessler, ScD**, **Douglas E. Wood, MD**, and **Farhood Farjah, MD, MPH**

Department of Surgery, University of Washington, Seattle; Clinical & Translational Research, Seattle Children's Hospital, Seattle; and Department of Health Services, University of Washington, Seattle, Washington

Abstract

Background.—One in 5 patients with completely resected early-stage non-small cell lung cancer will recur within 2 years. Risk stratification may facilitate a personalized approach to the use of adjuvant therapy and surveillance imaging. We developed a prediction model for recurrence based on five clinical variables (tumor size and grade, visceral pleural and lymphovascular invasion, and sublobar resection), and tested the hypothesis that the addition of several new molecular markers of poor long-term outcome (vascular endothelial growth factor C; microRNA precursors 486 and 30d) would enhance prediction.

Methods.—We performed a retrospective cohort study of patients with completely resected, node-negative non-small cell lung cancer from 2011 to 2014 (follow-up through 2016) using the Lung Cancer Biospecimen Resource Network. Cox regression was used to estimate the 2-year risk of recurrence. Our primary measure of model performance was the optimism-corrected C statistic.

Results.—Among 173 patients (mean tumor size, 3.6 cm; 12% sublobar resection, 32% poorly differentiated, 16% lymphovascular invasion, 26% visceral pleural invasion), the 2-year recurrence rate was 23% (95% confidence interval, 17% to 31%). A prediction model using five known risk factors for recurrence performed only slightly better than chance in predicting recurrence (optimism-corrected C statistic, 0.54; 95% confidence interval, 0.51 to 0.68). The addition of biomarkers did not improve the model's ability to predict recurrence (corrected C statistic, 0.55; 95% confidence interval, 0.52 to 0.71).

Conclusions.—We were unable to predict lung cancer recurrence using a risk-prediction model based on five well-known clinical risk factors and several biomarkers. Further research should consider novel predictors of recurrence to stratify patients with completely resected early-stage non-small cell lung cancer according to their risk of recurrence.

Address correspondence to Dr Farjah, Department of Surgery, University of Washington, 1959 NE Pacific St, Box 356310, Seattle, WA 98195; ffarjah@uw.edu.

The Supplemental Tables can be viewed in the online version of this article [https://doi.org/10.1016/j.athoracsur.2018.06.022] on http://www.annalsthoracicsurgery.org.

Patients with completely resected node-negative non-small cell lung cancer (NSCLC) account for approximately 17% of the lung cancer population (~40,000 patients/year) [1, 2]. This segment of the NSCLC population is expected to increase over time as a result of lung cancer screening and the rising incidence of incidentally detected lung nodules [3, 4]. Approximately 20% of patients with completely resected node-negative NSCLC will experience a recurrence, and most of those occur within the first 2 years after resection [5, 6].

One strategy for reducing the risk of recurrence is adjuvant therapy. However, guidelines generally recommend against radiotherapy or chemotherapy for patients with negative margins who have no nodal disease [7]. They do suggest consideration of adjuvant therapy for patients with pathologic findings associated with poor outcomes, such as poorly differentiated tumors, vascular invasion, wedge resections, visceral pleural involvement, and tumors exceeding 4 cm, but without further guidance [7]. Guidelines also recommend surveillance imaging for all resected patients [7]. Unfortunately, adherence rates with surveillance imaging are low for unclear reasons. The level of evidence supporting guideline recommendations on surveillance imaging is also low [8, 9]. Risk stratification may facilitate a personalized approach to the use of adjuvant therapy and surveillance imaging in patients with completely resected node-negative NSCLC.

A comprehensive literature review revealed five predominant clinical risk factors for lung cancer recurrence and several novel biomarkers associated with poor survival rates among patients with completely resected NSCLC [10–19]. The availability of multiple risk factors for recurrence suggests that the use prediction to facilitate risk stratification may be possible. We aimed to develop and internally validate a prediction model estimating the 2-year risk of recurrence in patients with completely resected, node-negative NSCLC and to test the hypothesis that the addition of biomarkers enhances prediction.

Patients and Methods

Study Design and Population

We performed a retrospective cohort study of NSCLC patients who underwent a margin-negative resection with no pathologic evidence of nodal disease between May 2011 and July 2014 (follow-up through October 2016). A total of 211 patients met criteria for inclusion in the study. After excluding patients with a history of prior lung cancer (n = 9), no follow-up data (n = 12), or missing covariates prespecified for model inclusion (n = 17), 173 were included (Fig 1). Data were obtained from the Lung Cancer Biospecimen Resource Network (LCBRN), a United States Department of Defense–funded biorepository collecting clinical data and tissue from consenting patients [20]. Study patients were recruited from the Medical University of South Carolina, the University of Virginia, and Washington University in St. Louis. Patients included active-duty, retired military personnel, and civilians. The University of Washington Institutional Review Board approved this investigation, which was performed in compliance with the Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK) checklist [21].

Clinical Variables

The LCBRN ascertained demographics and lung cancer risk factors (eg, smoking, environmental exposures, etc.) using a standardized intake form. Information on staging and treatment characteristics (eg, extent of resection) was abstracted from the medical record. We abstracted information on tumor size and grade, histologic type, presence of lymphovascular invasion (LVI) and visceral pleural invasion (VPI), pathologic stage, and the number of nodal stations sampled from redacted pathology reports.

Lung Cancer Recurrence

The LCBRN performed regular follow-up on study patients and reviewed medical records for up to 5 years, until death occurred, or until the time of specimen release (October 24, 2016), whichever occurred first. A pathologic diagnosis was the preferred method of determining recurrence, but abstractors also coded recurrence if the patient's providers had documented such a diagnosis. The LCBRN does not collect surveillance imaging, onset of new symptoms, or the manner in which recurrence was first detected and ultimately diagnosed; however, recurrences were distinguished from second primary lung cancers based on consensus between the treating thoracic surgeon, pathologist, and oncologist. Patients with second primaries remained at risk for recurrence through the course of the study. We were unable to account for the competing risk of death from noncancer causes because 13 deaths occurred after the follow-up period for recurrence. In these cases, we cannot determine whether the patient had a recurrence in the interval between the last follow-up period and date of death.

Variable Selection and Model Development

Two models were selected as candidates for predicting lung cancer recurrence. Model A included five clinical variables—VPI, LVI, tumor size, tumor grade, and sub-lobar resection (segmentectomy or wedge resection)—and was fit by Cox proportional hazards regression. A comprehensive literature review revealed these five variables to be the most commonly associated with recurrence [5, 10–18]. Model B consisted of all variables in model A with the intent of evaluating the addition of up to five biomarkers—vascular endothelial growth factor-C (VEGF-C) and microRNA (miRNA) precursors miR-1, miR-486, miR-499, and miR-30D. Two biomarkers were excluded because of very low rates of detection (miR-1 and miR-499 were expressed in 8% and 1% of patients, respectively). VEGF-C is a growth factor that stimulates lymphangiogenesis [22] and has been associated with lung cancer recurrence among patients with nodal disease [23]. A prior study showed variability in VEGF-C expression, suggesting it may be a predictor of recurrence [24].

The four miRNAs we initially intended to study were previously shown to be associated with poor long-term survival among treated lung cancer patients [25], and recurrence is the most common cause of death among patients with lung cancer [26]. For these reasons, we hypothesized that these biomarkers might improve our ability to predict lung cancer recurrence.

Variables with significantly skewed distributions, including tumor size and biomarkers, were log-transformed to an additive scale. Missing risk-factor data occurred in only 8.9% of

patients. There were no significant differences in demographic, clinical variables, levels of biomarkers, follow-up time, or rates of recurrence between patients with and without missing covariates (Supplemental Table A).

Biomarker Measurement

Tissue from resected tumor specimens was snap frozen, and RNA was isolated using the RNeasy Plus Universal Kit (Qiagen, Hilden, Germany) by LCBRN staff [20]. Samples were purchased from the LCBRN, and expression levels of VEGF-C and miRNAs were measured by quantitative polymerase chain reaction using commercially available primers [25]. Values were expressed as the fold-change relative to the housekeeping gene glyceraldehyde-3-phosphate dehydrogenase using the $2^{-\Delta\Delta CT}$ method [27]. Each biomarker was measured in triplicate, and the mean value for each patient was used for modeling. The mean within-sample coefficient of variation was calculated [28] and is reported in Supplemental Table B along with biomarker summary statistics. All coefficients of variation were less than 5%, indicating little intraassay variability in measurement. Investigators performing sample analysis were blinded to clinical outcomes data.

Model Performance and Comparisons

Receiver operating characteristic (ROC) curves (estimating sensitivity and specificity) were determined by nearest neighbor estimator smoothing [29]. These curves accounted for censoring in predicted and observed probability of recurrence at 2 years. The area under the ROC curve (AUC) was used to evaluate each model's discriminatory ability at 2 years [29, 30]. A goodness-of-fit test was used to evaluate each model's calibration, with small p values (eg, $p < 0.05$) indicating poor calibration [31]. We used bootstrapping (1,000 replicates) to internally cross-validate discrimination, reporting optimism-corrected C statistics, and nested bootstrapping (2,000 outer replicates and 1,000 inner replicates) to calculate corresponding 95% nonparametric confidence intervals (CIs) [32, 33]. Optimism correction accounts for bias arising from the use of the same data set for estimating coefficients and for evaluating the model's performance. Calibration plots were created with predicted probability of 2-year survival against corresponding observed survival probability. A likelihood ratio test was used to evaluate improvement in prediction performance of model B compared with model A [34]. Proportional hazards were assessed by the Schoenfeld residuals test for trend. Supplemental Table C demonstrates a post hoc univariate analysis of predictors of lung cancer recurrence. All analyses were performed using Stata 14.2 IC software (StataCorp, College Station, TX) and R 3.3.3 software (2017-03-06, The R Foundation for Statistical Computing, Vienna, Austria) [35].

Results

Patients in this cohort were a mean age of 66 years, 50% were women, 91% with a history of smoking (mean of 49 pack-years), and 25% had a family history of lung cancer. The mean primary tumor size was 3.6 cm, and the most common histologic type was adenocarcinoma (57%). Most patients underwent a lobectomy (86%) with a median sampling of 4 hilar and mediastinal lymph node stations (interquartile range [IQR], 3 to 5). After a median follow-up time of 30 months (IQR, 15 to 43.2 months), 48 patients had documented recurrence (2-

year recurrence rate, 23%; 95% CI, 17% to 31%). Among patients with recurrence, 14 were local (29%), 15 were regional (31%), and 19 were distant (40%). Median time to recurrence was 11 months (IQR, 7 to 27 months). Figure 2 demonstrates a Kaplan-Meier curve for recurrence-free survival among those with and without each clinical risk factor. Complete demographic, clinical, biomarker, and recurrence features are summarized in Table 1.

The prediction model based on prespecified clinical variables (model A) had an AUC of 0.65 (95% CI, 0.53 to 0.77). Model B (clinical risk factors plus biomarkers) had an AUC of 0.66 (95% CI, 0.58 to 0.79). Figures 3A and 3B display ROCs and corresponding calibration plots for both prediction models. Goodness-of-fit tests show no evidence of poor fitting models (all $p > 0.05$). The likelihood ratio test indicated that the addition of biomarkers to model A did not improve prediction ($p = 0.36$). Optimism-corrected C statistics were 0.54 (95% CI, 0.51 to 0.68) for model A and 0.55 (95% CI, 0.52 to 0.71) for model B. The performance characteristics of the models are summarized in Supplemental Table D. Schoenfeld residuals tests for trend did not provide evidence of nonproportional hazards in either model.

Comment

Despite having no clinical or pathologic evidence of residual disease, patients with completely resected node-negative NSCLC remain at risk for recurrence. We aimed to develop and validate a prediction model for recurrence as a first step toward designing a risk-stratified approach for the selective use of adjuvant therapy and surveillance imaging. We were unable to develop a model that reliably performed better than chance in predicting recurrence despite using five previously identified clinical risk factors for recurrence [5, 10–18]. The addition of novel biomarkers associated with poor survival [23, 25] did not improve the model's ability to predict recurrence.

The most likely explanation for our study's findings is the relatively small magnitude of associations between clinical risk factors and recurrence. Variables associated with less than threefold risk of an outcome are poor predictors of individual outcomes [36]. VPI, grade, and tumor size were previously reported to have hazard ratios of less than 3 [12, 16]. We hypothesized that the cumulative contribution of each risk factor would lead to reasonable performance, as demonstrated in other settings, such as predicting nodal disease [11]. None of the variables in our model had a hazard ratio greater than 3. Our inability to predict lung cancer recurrence highlights the challenges of using associations derived from population-based studies to predict individual outcomes. Another possibility is that prior reports of associations between clinical variables and outcomes were "false discoveries" due to uncorrected selection bias [37]. However, the consistency of reported associations across studies between lung cancer recurrence and tumor grade and size, VPI, LVI, and sublobar resection argues against this possibility. Furthermore, the relationships between risk factor and outcome in our sample matched those of prior reports (eg, LVI is associated with a higher risk of recurrence).

VEGF-C has been implicated in dissemination of cancer by the lymphatic system through promotion of lymphangiogenesis [22]. Others have shown VEGF-C to be associated with

recurrence among patients with nodal disease, but this association has not been shown among patients without nodal disease [23]. The rationale for investigating this association in node-negative patients is that surgeons are known to variably assess lymph nodes, resulting in potential misclassification of true nodal status [38]. In other words, elevated VEGF-C levels could be a surrogate for inadequate nodal staging in a patient who truly has occult nodal disease, and nodal disease is a very strong predictor of recurrence. This assumption, however, is not verifiable using this data set, which lacks individual surgeon identifiers.

A previous study showed miRNA biomarkers had high-magnitude associations with poor long-term survival among patients with stage I to IIIA NSCLC [25]. We hypothesized that these biomarkers might predict recurrence. However, the original biomarker study included a heterogeneous population of patients with and without nodal disease, and miRNA expression may be collinear with nodal disease. If true, miRNA would not be expected to predict recurrence among patients without nodal disease. Another possibility is that miRNA expression may be associated with poor long-term survival through a pathway other than cancer recurrence.

Our study had several limitations. First, our cohort was relatively small. To avoid over-parameterization, we constructed a parsimonious model. Another problem with a small cohort is the inability to divide it into development and validation sets. Instead we prespecified an optimism-corrected C statistic to validate model performance and safeguard against overly optimistic results. This approach revealed a model that predicted recurrence no better than chance (ie, a conservative result). Nonetheless, our study is, as far as we can determine, the largest cohort of patients from multiple centers in the United States with data on clinical variables, tissue samples for biomarker testing, and follow-up information on recurrence.

Second, LCBRN did not measure information about surveillance imaging and visits. The lack of standardized surveillance would be expected to affect time to detection of recurrence and possibly the "stage" of recurrence (eg, local only, locoregional, distant), assuming that surveillance truly leads to early detection of recurrence. Accordingly, it is unlikely that variation in surveillance biased our measurement of any recurrence at 2 years, and our overall recurrence rates were similar to prior reports [5, 6].

Third, it is possible that misclassification of the outcome (eg, calling a second primary a recurrence) or of risk factors partly explains our inability to predict recurrence. Unfortunately, without adjudication of each case, which is not possible with the LCBRN, we cannot characterize the frequency of misclassification. Although the frequencies of VPI, LVI, and tumor grade in our sample are similar to prior reports (our cohort tended to have larger tumors and less frequent sublobar resection) [5, 6], frequency comparisons do not provide insights into the rate of misclassification. To the extent that misclassification of outcome occurred, we expect it to be uncommon because the risk of recurrence is substantially higher than the risk of second primary in the first 2 years after resection [6].

Despite the negative findings of our study, this line of investigation remains important given the limited evidence and guidance on the use of adjuvant therapy and surveillance imaging

in patients with completely resected node-negative NSCLC [11]. Stratification into high-risk and low-risk groups allows for a simple decision-making algorithm for clinicians. For example, high-risk patients might receive adjuvant therapy and more frequent surveillance imaging for the first 2 years, whereas low-risk patients would not receive additional therapy and undergo less intense surveillance imaging. This strategy could lead to more personalized and higher-value lung cancer care, and developing and validating a risk-prediction model is the first step toward achieving this broader goal. The discovery of novel predictors of recurrence would make this possible. Near the conclusion of our study, an investigation was published revealing an association between long-term outcomes and intratumor heterogeneity mediated through chromosome instability in resected lung specimens [39]. Specifically, elevated copy-number heterogeneity was associated with a five-fold increased risk of recurrence or death, making this molecular marker a candidate predictor for recurrence in future studies.

In conclusion, we were unable to predict lung cancer recurrence using five known clinical risk factors of recurrence. The addition of several biomarkers to the model did not improve our ability to predict lung cancer. Rapid discovery of novel predictors of recurrence motivates continued efforts to develop and validate a prediction model for recurrence for the purpose of risk stratification.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. Adv Exp Med Biol 2016;893:1–19. [PubMed: 26667336]

2. Little AG, Gay EG, Gaspar LE, Stewart AK. National survey of non-small cell lung cancer in the United States: epidemiology, pathology and patterns of care. Lung Cancer 2007;57:253–60. [PubMed: 17451842]

3. Gould MK, Tang T, Liu ILA, et al. Recent trends in the identification of incidental pulmonary nodules. Am J Respir Crit Care Med 2015;192:1208–14. [PubMed: 26214244]

4. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011;365:395–409. [PubMed: 21714641]

5. Kiankhooy A, Taylor MD, LaPar DJ, et al. Predictors of early recurrence for node-negative T1 to T2b non-small cell lung cancer. Ann Thorac Surg 2014;98:1175–83. [PubMed: 25134863]

6. Lou F, Huang J, Sima CS, Dycoco J, Rusch V, Bach PB. Patterns of recurrence and second primary lung cancer in earlystage lung cancer survivors followed with routine computed tomography surveillance. J Thorac Cardiovasc Surg 2013;145: 75–82. [PubMed: 23127371]

7. National Comprehensive Cancer Network. Non-small cell lung cancer (Version 5.2017). Available at https://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf. Accessed March 18, 2017.

8. Srikantharajah D, Ghuman A, Nagendran M, Maruthappu M. Is computed tomography follow-up of patients after lobectomy for non-small cell lung cancer of benefit in terms of survival? Interact Cardiovasc Thorac Surg 2012;15:893–8. [PubMed: 22859511]

9. Backhus LM, Farjah F, Zeliadt SB, et al. Predictors of imaging surveillance for surgically treated early-stage lung cancer. Ann Thorac Surg 2014;98:1944–52. [PubMed: 25282167]

10. Ko KH, Hsu HH, Huang TW, et al. Predictive value of 18F-FDG PET and CT morphologic features for recurrence in pathological stage IA non-small cell lung cancer. Medicine (Baltimore) 2015;94:e434. [PubMed: 25621697]

11. Zhang Y, Sun Y, Xiang J, Zhang Y, Hu H, Chen H. A clinicopathologic prediction model for postoperative recurrence in stage Ia non-small cell lung cancer. J Thorac Cardiovasc Surg 2014;148:1193–9. [PubMed: 24667022]

12. Kuo SW, Chen JS, Huang PM, Hsu HH, Lai HS, Lee JM. Prognostic significance of histologic differentiation, carcinoembryonic antigen value, and lymphovascular invasion in stage I non-small cell lung cancer. J Thorac Cardiovasc Surg 2014;148:1200–7.e3. [PubMed: 24993621]

13. Chen YY, Huang TW, Tsai WC, et al. Risk factors of postoperative recurrences in patients with clinical stage I NSCLC. World J Surg Oncol 2014;12:10. [PubMed: 24410748]

14. Neri S, Yoshida J, Ishii G, et al. Prognostic impact of microscopic vessel invasion and visceral pleural invasion in non-small cell lung cancer: a retrospective analysis of 2657 patients. Ann Surg 2014;0:1–6.

15. Al-Alao BS, Gately K, Nicholson S, McGovern E, Young VK, O'Byrne KJ. Prognostic impact of vascular and lymphovascular invasion in early lung cancer. Asian Cardiovasc Thorac Ann 2014;22:55–64. [PubMed: 24585645]

16. Koike T, Koike T, Yoshiya K, Tsuchida M, Toyabe S. Risk factor analysis of locoregional recurrence after sublobar resection in patients with clinical stage IA non-small cell lung cancer. J Thorac Cardiovasc Surg 2013;146:372–8. [PubMed: 23870323]

17. Hung JJ, Jeng WJ, Chou TY, et al. Prognostic value of the new International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society Lung Adenocarcinoma classification on death and recurrence in completely resected stage I lung adenocarcinoma. Ann Surg 2013;258:1079–86. [PubMed: 23532112]

18. Lopez Guerra JL, Gomez DR, Lin SH, 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. [PubMed: 23002278]

19. Tang Y, Qiao G, Xu E, Xuan Y, Liao M, Yin G. Biomarkers for early diagnosis, prognosis, prediction, and recurrence monitoring of non-small cell lung cancer. Onco Targets Ther 2017;10:4527–34. [PubMed: 28979144]

20. Lung Cancer Biospecimen Resource Network. Available at http://lungbio.sites.virginia.edu/. Accessed May 11, 2017.

21. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. REporting recommendations for tumor MARKer prognostic studies (REMARK). Breast Cancer Res Treat 2006;100:229–35. [PubMed: 16932852]

22. Alitalo K, Tammela T, Petrova TV. Lymphangiogenesis in development and human disease. Nature 2005;438:946–53. [PubMed: 16355212]

23. Chen G, Liu XY, Wang Z, Liu FY. Vascular endothelial growth factor C: the predicator of early recurrence in patients with N2 non-small-cell lung cancer. Eur J Cardiothorac Surg 2010;37:546–51. [PubMed: 19758816]

24. Farjah F, Madtes DK, Wood DE, et al. Vascular endothelial growth factor C complements the ability of positron emission tomography to predict nodal disease in lung cancer. J Thorac Cardiovasc Surg 2015;150: 796–803. [PubMed: 26320776]

25. Hu Z, Chen X, Zhao Y, et al. Serum microRNA signatures identified in a genome-wide serum microRNA expression profiling predict survival of non-small-cell lung cancer. J Clin Oncol 2010;28:1721–6. [PubMed: 20194856]

26. al-Kattan K, Sepsas E, Fountain SW, Townsend ER. Disease recurrence after resection for stage I lung cancer. Eur J Cardiothorac Surg 1997;12:380–4. [PubMed: 9332915]

27. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and. Methods 2001;25:402–8. [PubMed: 11846609]

28. Reed GF, Lynn F, Meade BD. Use of coefficient of variation in assessing variability of quantitative assays. Clin Diagn Lab Immunol 2002;9:1235–9. [PubMed: 12414755]

29. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics 2000;56:337–44. [PubMed: 10877287]

30. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the Cstatistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Stat Med 2011;30: 1105–17. [PubMed: 21484848]

31. May S, Hosmer DW. A simplified method of calculating an overall goodness-of-fit test for the Cox proportional hazards model. Lifetime Data Anal 1998;4:109–20. [PubMed: 9658770]

32. Harrell FE, Jr. Regression modeling strategies with applications to linear models, logistic regression, and survival analysis Springer Series in Statistics. New York: Springer-Verlag; 2001.

33. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. Stat Med 2000;19:1141–64. [PubMed: 10797513]

34. Pepe MS, Janes H, Li CI. Net risk reclassification P values: valid or misleading? J Natl Cancer Inst 2014;106: dju041. [PubMed: 24681599]

35. The R Foundation. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing Available at https://www.r-project.org/. Accessed May 22, 2017.

36. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. Am J Epidemiol 2004;159:882–90. [PubMed: 15105181]

37. Storey J The positive false discovery rate: a Bayesian interpretation and the q-value. Ann Stat 2003;31:2013–35.

38. Thornblade LW, Wood DE, Mulligan MS, et al. Variability in invasive mediastinal staging for lung cancer: a multicenter regional study. J Thorac Cardiovasc Surg 2018;155: 2658–71. [PubMed: 29534904]

39. Jamal-Hanjani M, Wilson GA, McGranahan N, et al. Tracking the evolution of non-small-cell lung cancer. N Engl J Med 2017;376:2109–21. [PubMed: 28445112]
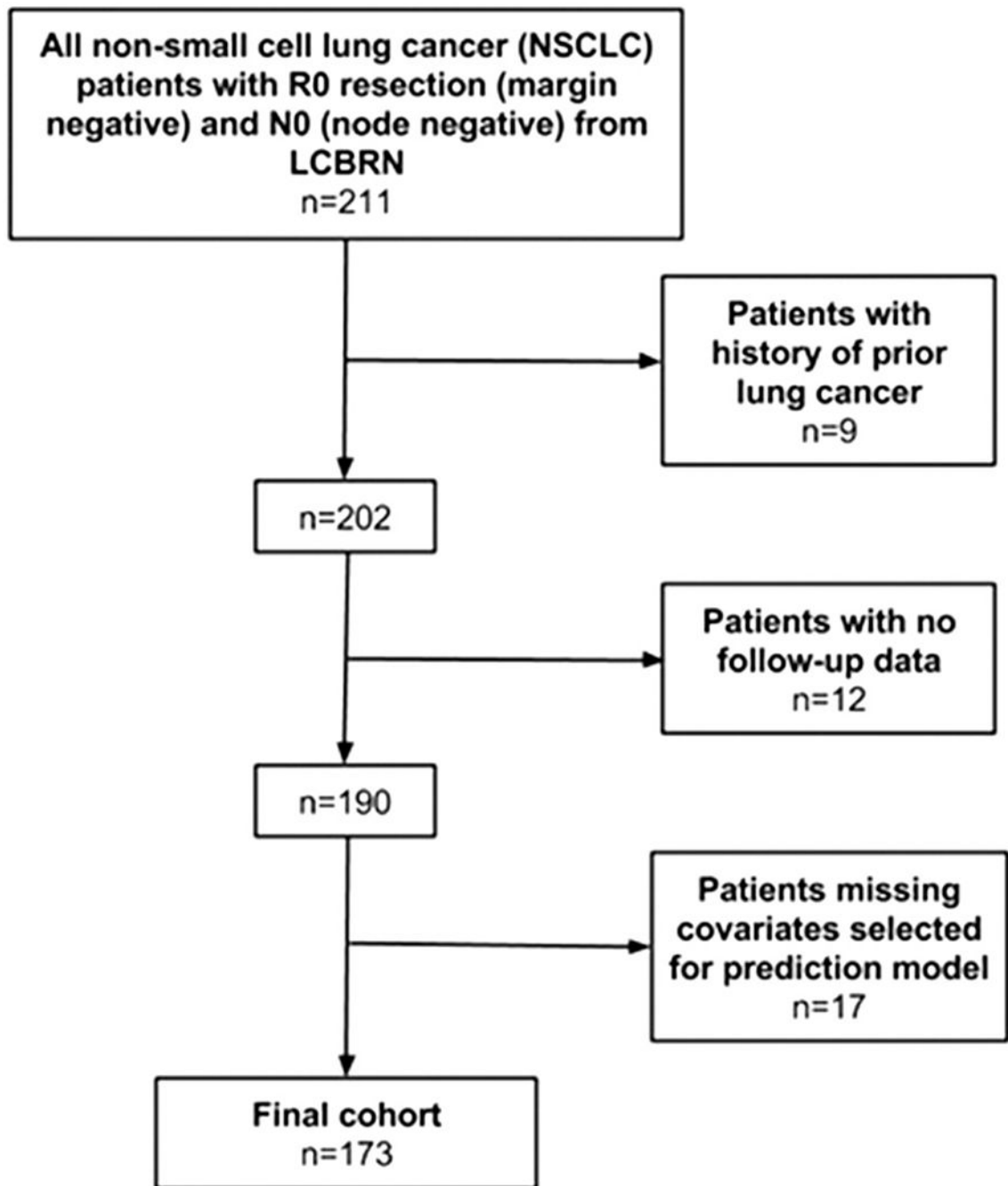
Fig 1.
Diagram of cohort selection. (LCBRN = Lung Cancer Biospecimen Resource Network.)
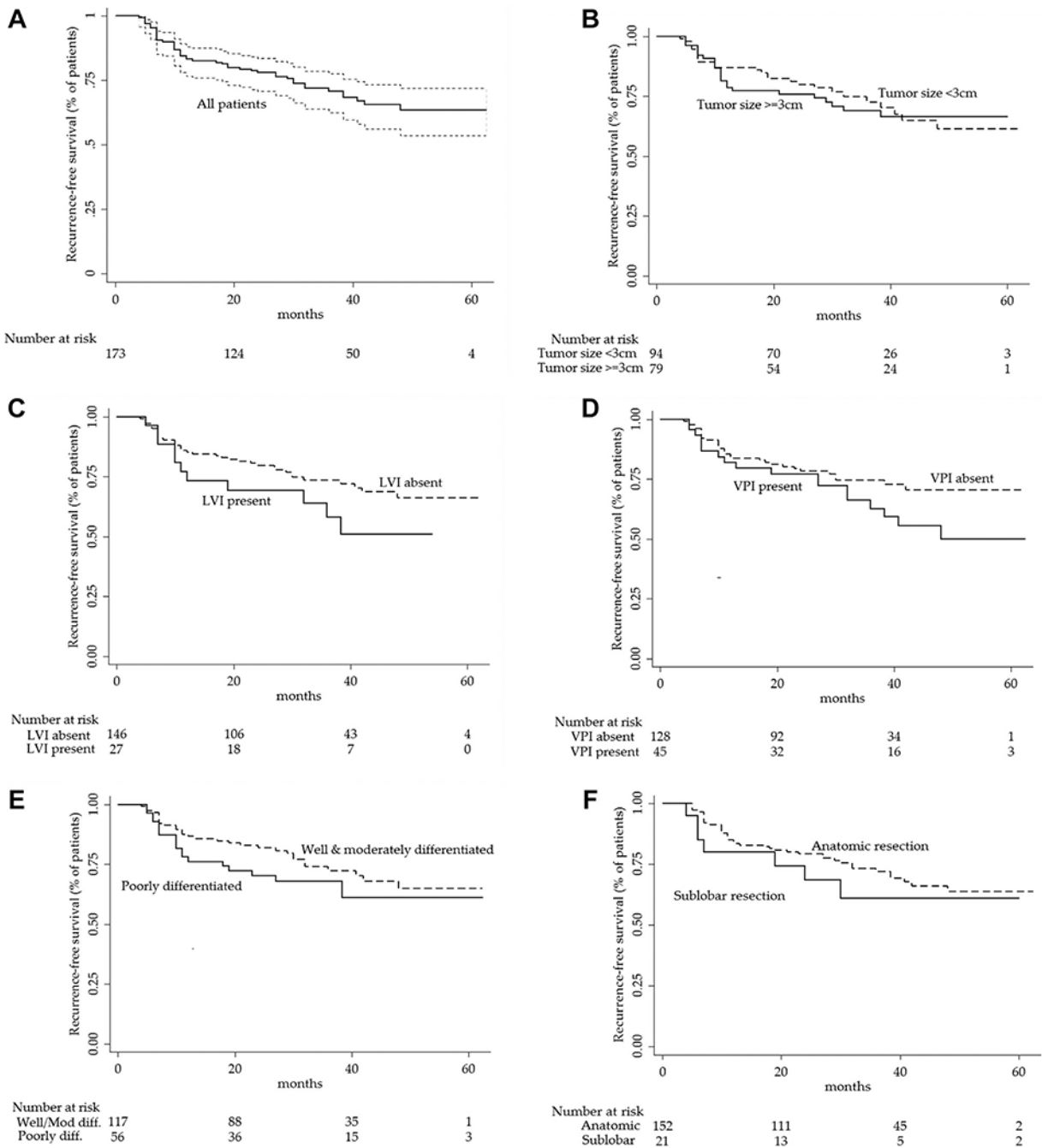
Fig 2.
(A) Kaplan-Meier survival estimates for recurrence-free survival among patients with completely resected, node-negative non-small cell lung cancer and for five pre-specified clinical risk-factors: (B) tumor size, (C) lymphovascular invasion (LVI), (D) visceral pleural invasion (VPI), (E) tumor grade, and (F) sublobar resection.
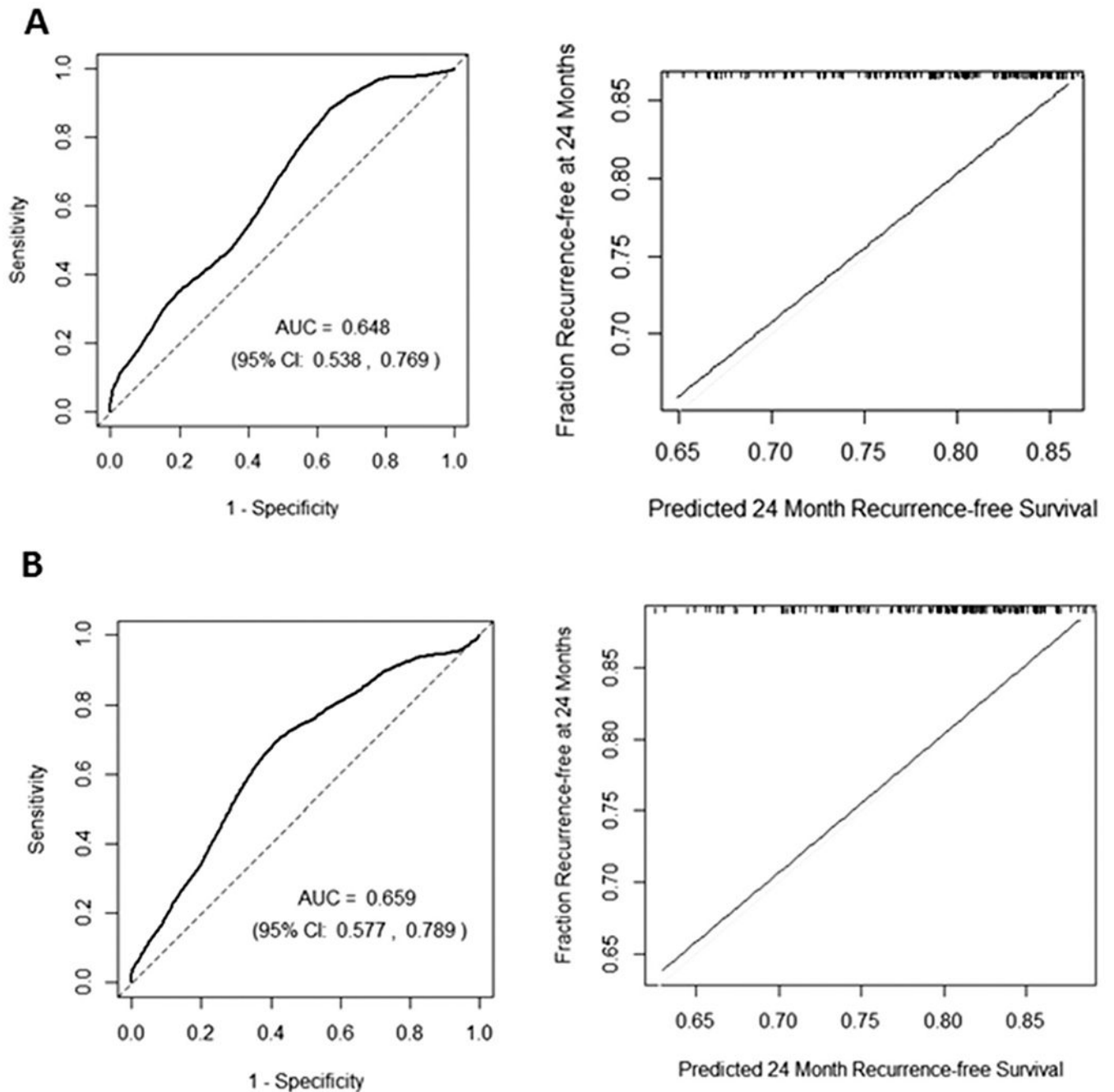
Fig 3.
Receiver operating characteristic curves (left) and calibration plots (right) for (A) model A (visceral pleural invasion, lymphovascular invasion, log-tumor size, tumor grade, and sublobar resection and (B) model B (model A plus vascular endothelial growth factor C, micro-RNA [miR]-486, and miR-30d). The time-dependent receiver operating characteristic curve was evaluated at 2 years using nearest neighbor estimator smoothing with a span equal to $0.25 \times n^{-0.2}$, with n equal to the number of patients. The calibration plot reflects "predicted" survival probabilities at 24 months (recurrence-free survival) and corresponding "observed" survival probabilities by Kaplan-Meier (fraction recurrence-free at 24 months).

The tick marks at the top of this figure indicate points observed with corresponding predicted survival. (AUC = area under the receiver operating characteristic curve; CI = confidence interval.)

Table 1.

Characteristics of Patients With Completely Resected Node-Negative Non-Small Cell Lung Cancer

Variables	Data (N = 173)
Demographics	
Age, mean (SD), years	66 (10)
Female, No. (%)	86 (50)
Race, No. (%)	
White	158 (91)
Black	12 (7)
Asian	2 (1)
Native American	1 (1)
Environmental exposures, No. (%)	
Second-hand smoke	157 (91)
Asbestos	70 (40)
Coal mining	10 (6)
Uranium	3 (2)
Radon	6 (3)
Clinical variables	
Body mass index, mean (SD), kg/m^2	28.0 (6.1)
ECOG, No. (%)	
0	82 (48)
1	75 (44)
2	14 (8)
Family history of lung cancer, No. (%)	43 (25)
Smoking status, No. (%)	
Never	16 (9)
Former	110 (64)
Current	47 (27)
Pack years,[a] mean (SD)	49 (30)
SUV$_{max}$ of primary tumor, mean (SD)	9.7 (6.8)
Sublobar resection, No. (%)	21 (12)
Pathologic features	
Tumor size, mean (SD), cm	3.6 (2.3)
Tumor histology, No. (%)	
Squamous cell carcinoma	62 (36)
Adenocarcinoma	99 (57)
Other	12 (7)
Tumor grade, No. (%)	
Well differentiated	25 (14)
Moderately differentiated	92 (53)
Poorly differentiated	56 (32)
Lymphovascular invasion, No. (%)	27 (16)

Variables	Data (N = 173)
Visceral pleural invasion, No. (%)	45 (26)
Carcinoma in situ at bronchial margin, No. (%)	5 (3)
Lymph node stations sampled, median (IQR), No.	4 (3–5)
Mediastinal	3 (2–4)
Hilar	1 (1–2)
Biomarkeras,[b] median (minimum-maximum) VEGF-C	
VEGF-C	0.002 (0–0.127)
miR-1	0 (0–4.464)
miR-486	0.252 (0–40.651)
miR-499	0 (0–451,773)
miR-30d	0.438 (0–19.056)
Follow-up,[c] median (IQR), months	30 (15–43.2)

[a] Among current and former smokers.

[b] Biomarker values (mean of triplicate values) are reported as the fold-change expression compared with housekeeping genes and therefore are unitless.

[c] Time to recurrence, loss to follow-up, or 5 years of follow-up, whichever occurred first.

ECOG = European Cooperative Oncology Group; IQR = interquartile range; miR = microRNA precursor; SUV_{max} = maximum standardized uptake value; VEGF = vascular endothelial growth factor.