Fluorine-18 fluorodeoxyglucose positron emission tomographic maximal standardized uptake value predicts survival independent of clinical but not pathologic TNM staging of resected non–small cell lung cancer

Robert J. Downey, MD,[a] Timothy Akhurst, MBBS, FRACP,[b] Mithat Gonen, PhD,[c] Bernard Park,[a] and Valerie Rusch, MD[a]

Robert J. Downey

Objectives: Positron emission tomographic maximal standardized uptake value has been shown to predict survival after resection of non–small cell lung cancer. The relative prognostic benefit of maximal standardized uptake value with respect to other clinical/pathologic variables has not been defined.

Methods: We reviewed patients who had positron emission tomographic imaging and an R0 resection for non–small cell lung cancer between January 1, 2000, and December 31, 2004, without induction or adjuvant therapy. The associations between overall survival, histology, pathologic TNM stage, pathologic tumor diameter, and standardized uptake value were tested.

Results: Four hundred eighty-seven patients met the study criteria. Median follow-up was 25.8 months. By using the median values for tumor size (2.5 cm) and standardized uptake value (5.3), standardized uptake value was an independent predictor of survival ($P = .03$), adjusting for tumor size ($P = .02$) and histology ($P < .01$). The optimal standardized uptake value for stratification was identified as 4.4, and this value was identified as an independent predictor of survival ($P = .03$) after adjusting for clinical TNM stage. Standardized uptake value was not an independent predictor of survival ($P = .09$), adjusting for pathologic TNM stage (stage IA vs IB vs stage II–IV, $P < .01$).

Conclusions: Standardized uptake value does not add to the prognostic significance of pathologic TNM stage. Standardized uptake value was an independent prognostic factor from clinical TNM stage.

From the Thoracic Surgery Service, Department of Surgery,[a] the Division of Nuclear Medicine, Department of Radiology,[b] and the Department of Epidemiology and Biostatistics,[c] Memorial Sloan-Kettering Cancer Center, New York, NY.

Read at the Eighty-sixth Annual Meeting of The American Association for Thoracic Surgery, Philadelphia, Pa, April 29-May 3, 2006.

Received for publication June 29, 2006; revisions received Dec 6, 2006; accepted for publication Jan 8, 2007.

Address for reprints: Robert J. Downey, MD, Thoracic Service, Department of Surgery, Memorial Sloan-Kettering Cancer Center, 1275 York Ave, New York, NY 10021. (E-mail: downeyr@mskcc.org).

J Thorac Cardiovasc Surg 2007;133:1419-27

0022-5223/$32.00

doi:10.1016/j.jtcvs.2007.01.041

Abbreviations and Acronyms

CT = computed tomography
^{18}F-FDG = fluorine-18 fluorodeoxyglucose
NSCLC = non–small cell lung cancer
pTNM = pathologic TNM
SUV = standardized uptake value
SUV_{MAX} = maximal standardized uptake value

GTS

The standardized uptake value (SUV) for fluorine-18 fluorodeoxyglucose (^{18}F-FDG) as measured by using positron emission tomography (PET) has been shown to correlate with several measures of tumor behavior, such as lesion doubling time[1] and Ki-67 staining,[2] suggesting that SUV might be a predictor of patient prognosis. Previously, to determine whether ^{18}F-FDG uptake in a malignancy correlated with prognosis, we performed a retrospective review of patients with histologically proved non–small cell lung cancer (NSCLC) or carcinoid cancer (pathologic T1-4N0-2M0) who had undergone R0 resections after PET imaging and without either neoadjuvant or adjuvant therapy.[3] We found that stratification of patients by the median SUV_{MAX} (which was 9) predicted survival; the 2-year survival for patients with an SUV_{MAX} of greater than 9 was 68%, and that for patients with an SUV_{MAX} of less than 9 was 96% ($P < .01$, log–rank test). However, an insufficient number of patients was available to allow an analysis of the relationship of SUV_{MAX} to pathologic TMN staging. Other reports have attempted to determine whether SUV_{MAX} is an independent predictor from pathologic TNM (pTNM) staging of survival, but none have contained sufficient patients to allow a definitive answer.[4-15] Results in the published studies have been mixed, and possible reasons for the conflicting results have been recently extensively analyzed in a review by Pillot and coauthors.[16] To analyze whether PET SUV was a predictor of survival independent of pTNM staging, we reviewed our experience with a much larger cohort of patients treated since our original report.

Materials and Methods

We performed a retrospective review of patients who had PET imaging and an R0 resection for NSCLC without induction or adjuvant therapy at Memorial Sloan-Kettering Cancer Center during the period from January 1, 2000, through December 31, 2004. The primary goal of this study was to determine whether SUV_{MAX} of the primary site of disease was an independent predictor of survival from clinical and pTNM staging.

This review was performed after approval had been obtained from the Memorial Sloan-Kettering Cancer Center Institutional Review Board and in accord with an assurance filed with and approved by the Department of Health and Human Services.

Patient data were obtained from a prospectively maintained database in the Memorial Sloan-Kettering Cancer Center Thoracic Service, in which patient staging information is entered on a weekly basis under attending surgeon supervision, and survival data were updated at regular intervals by research study assistants. The primary end point was overall survival, which was calculated from the date of surgical intervention to the date of death or last contact with the patient.

Histologic characterization of the tumors was obtained from the operative pathologic report. Patients undergoing resections of what were believed by the treating physicians to be synchronous or metachronous (defined for the purpose of this study as within 5 years of prior lung cancer) primary NSCLCs were not included. Tumors were grouped into 3 categories: adenocarcinoma, squamous cell carcinoma, and "other." Patients who had no nodes sampled at the time of resection were treated as having pathologic N0 disease. Adenosquamous carcinomas were included in the "other" category. If recorded, SUV_{MAX} of the primary tumors was obtained from the radiology report, and if not recorded, it was calculated by a Nuclear Medicine physician (TA) from PET images. For statistical analyses, clinical and pTNM stages were compressed to 3 groups: stage IA, stage IB, and stage II to IV.

Associations between overall survival and tumor histology, clinical TNM stage, pTNM stage, and pathologic maximal tumor diameter were tested by using the log–rank test, and the associations between survival and SUV_{MAX} of the primary site of disease were evaluated with Cox regression. In determining the interrelation between SUV and variables that can be determined preoperatively, we used pathologically measured tumor size determined from pathology reports as a surrogate for imaging-based estimates of tumor size and pathologic histology as a surrogate for histology determined from a fine- or core-needle biopsy. Estimation of the optimal values for stratification was performed by using the maximal χ^2 method. Multivariate modeling to identify independent prognostic factors was performed by using Cox regression. Poor, intermediate, and good risk groups were identified by merging subgroups with similar subgroups through backward elimination following the Cox model. Thirty-three patients in the current cohort had been included in the previously reported analysis[3]; analyses were performed with these patients excluded, and the results did not differ significantly from those found in the material presented below (data not shown).

Results

Four hundred eighty-seven patients met the study criteria. Patient demographics are recorded in Tables 1A, 1B, and 1C. There was a predominance of male subjects (67%) and of adenocarcinoma (69%). The majority of patients (87.5%) underwent anatomic resections. Pathologic stage was IA in 249 patients, IB in 132 patients, and IIA through IV in 106 patients. With a median follow-up of 25.8 months, 17 (3%) patients have been lost to follow-up.

By using univariate analysis, tumor size (Figure 1) and SUV were determined to be significant predictors of survival ($P < .02$ and .03, respectively). Survival after resection stratified by histology (Figure 2) demonstrated significant differences between adenocarcinoma and squamous carcinoma ($P = .05$), adenocarcinoma and other histologies ($P < .01$), and squamous carcinoma and other histologies ($P = .05$).

TABLE 1A. Patient demographics

Sex	
Male subjects	230
Female subjects	257
Age (y), median (range)	69 (27–87)
Extent of resection	
Bilobectomy	18
Lobectomy	347
Pneumonectomy	19
Segmentectomy	41
Wedge	62
Histology	
Adenocarcinoma	337
Squamous	104
Other	46
Pathologic tumor size (mm), median (range)	25 (4–140)
SUV_{MAX} primary tumor	5.3 (0.6–36.3)
Follow-up (mo), median (range)	25.8 (1–66)
Status at last follow-up	
Alive with disease	38
Dead of disease	69
Dead of other causes	20
No evidence of disease	343
Lost to follow-up	17

SUV_{MAX}, Maximal standardized uptake value.

TABLE 1B. Clinical TNM stage

Clinical stage	
T	
T1	326
T2	153
T3	5
T4	3
N	
N0	404
N1	54
N2	26
N3	2
M	
M0	476
M1	4
Clinical TNM	
IA	284
IB	105
IIA	20
IIB	34
IIIA	26
IIIB	6
IV	4
Not available	8

TABLE 1C. Pathologic TNM stage

Postresection stage	
T	
T1	283
T2	173
T3	13
T4	18
N	
N0	395
N1	55
N2	27
N3	0
None sampled	10
M	
M0	479
M1	8
Pathologic TNM	
IA	249
IB	131
IIA	21
IIB	35
IIIA	27
IIIB	16
IV	8

The median SUV for adenocarcinoma was significantly different from that for the squamous carcinomas ($P < .01$) and that for the other histologies ($P < .01$). The median SUV values for squamous carcinoma and for other histologies were not different ($P = .69$, Figure E1).

We first performed an analysis of significant prognostic variables using the median values for SUV and for pathologic tumor size to avoid inadvertent bias. In an analysis using the median values for pathologically measured tumor size (2.5 cm) and tumor SUV_{MAX} (5.3), SUV_{MAX} was an independent predictor of survival ($P = .03$) after adjusting for tumor size ($P = .02$) and histology ($P < .01$). After demonstrating that the median value for SUV was a significant independent predictor of survival, the optimal cut-off point for stratification was calculated as a tumor size of 3.3 cm ($P < .03$) and an SUV of 4.3 ($P < .01$, Figure 3). This optimal value for SUV was an independent predictor of survival ($P = .03$) after adjusting for clinical TNM stage (as stage IA vs IB vs II-IV, $P < .01$). There was an interaction between SUV and clinical TNM stage in that SUV acted most strongly in the stratification of clinical TNM stage IB (Figure E2). Survival stratified by a combination of the optimal values for SUV and size demonstrated that patients with a size of greater than 2.5 cm and an SUV of greater than 4.3 had a significantly worse survival when compared with all other patients (Figure 4).

The SUV in patients with pTNM stage IA disease was significantly lower than that in patients with stage IB disease ($P < .01$) and patients with stage II to IV disease ($P <$

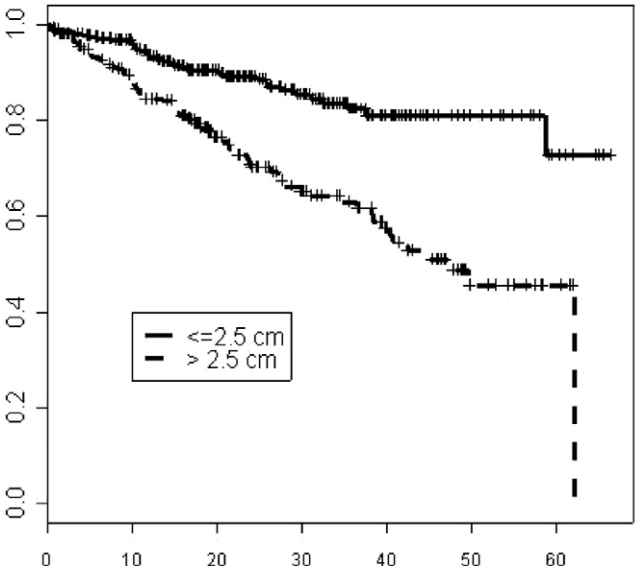

Figure 1. Survival after resection stratified by optimal value for maximal tumor diameter.

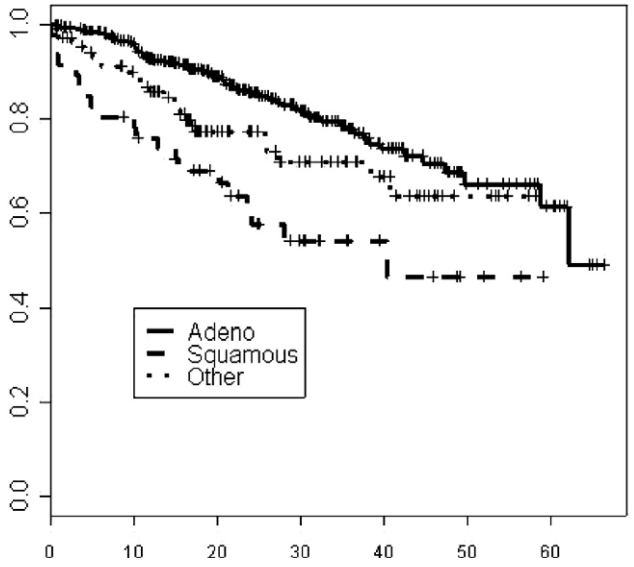

Figure 2. Survival after resection stratified by histology (*P* < .01 for adenocarcinoma *[Adeno]* vs other, *P* = .05 for adenocarcinoma vs squamous, *P* = .05 for squamous vs other).

.01). The SUV distribution between stage IB disease and stage II to IV disease did not differ ($P = .39$, Figure E3). Survival after resection stratified by pTNM stage (Figure E4) demonstrated significant differences between pTNM stages IA, IB, and II to IV ($P < .01$). Neither the median value for SUV nor the calculated optimal SUV were independent predictors of survival ($P = .09$ for both) after adjusting for pTNM stage (stages IA vs IB vs II-IV, $P < .01$; Figures E5–E7).

Given that SUV was an independent prognostic variable from clinical TNM stage, we attempted to determine whether there was a combination of prognostic variables available preoperatively, including SUV, that could approximate the prognostic information provided by the postoperative variables of pTNM staging and pathologic histology.

Combining histology and pTNM stage, we identified the following 3 statistically significant postoperative prognostic categories (Figure 5): good—adenocarcinoma, pTNM stage IA; poor—large cell/sarcomatoid, pTNM stages II to IV; intermediate, all other patients.

Combining SUV_{MAX}, pathologic tumor size, and histology, we identified the following statistically significant preoperative prognostic categories (Figure 6): good—adenocarcinoma, SUV of less than 4.4 and size of less than 2.5 cm; poor—large cell/sarcomatoid histology, SUV of greater than 4.4 and size of greater than 2.5; intermediate—all others.

After complete pathologic staging, 394 (81%) of 487 patients remained in the same prognostic group that they had been assigned to on the basis of clinical staging. The 19% (93/487) of patients who change prognostic groups after resection do so primarily by moving from a good to an intermediate prognosis (Table 2). The 3-year survival for patients in the preoperative and postoperative good category was 86% and 87%, respectively; 72% and 67%, respectively, for the intermediate group; and 45% and 45%, respectively, for the poor group.

Discussion

To determine whether PET SUV was a clinically relevant prognostic parameter, we examined a uniform population of patients who all had NSCLC of the lung resected with curative intent after PET imaging at one institution and who did not receive either induction or adjuvant therapy. Even though the number of patients available for analysis was more than twice the number (225) of surgically treated patients in the prior largest series,[17] to provide the highest quality statistical analyses, we compared SUV as a prognostic variable against 3 prognostic groupings of TNM stages. After patients had been stratified into groups as having pTNM stage IA disease (small node-negative tumors), stage IB disease (large node-negative tumors), and stage II to IV disease (82/105 or 78% of whom had lymph node involvement), further stratification by means of PET SUV did not further define prognosis.

However, further stratification by means of PET SUV after clinical TNM staging significantly improved the prognosis. This suggests that there might be a role for PET SUV in defining patient prognosis preoperatively to help in determin-

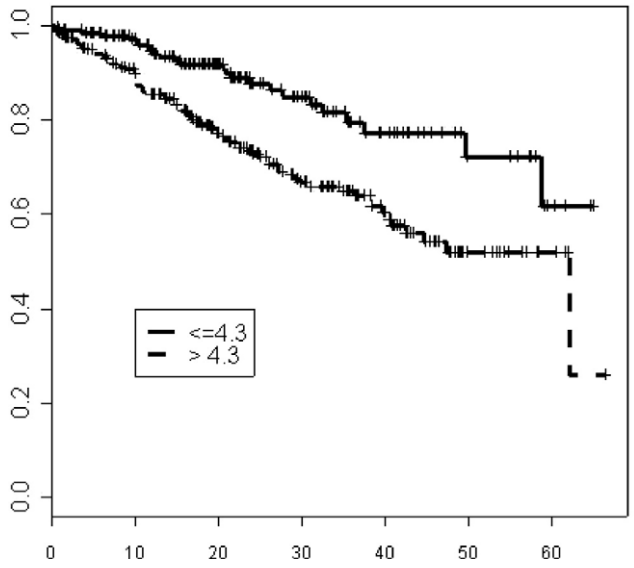

Figure 3. Survival after resection stratified by optimal value for maximal standardized uptake value of the primary tumor.

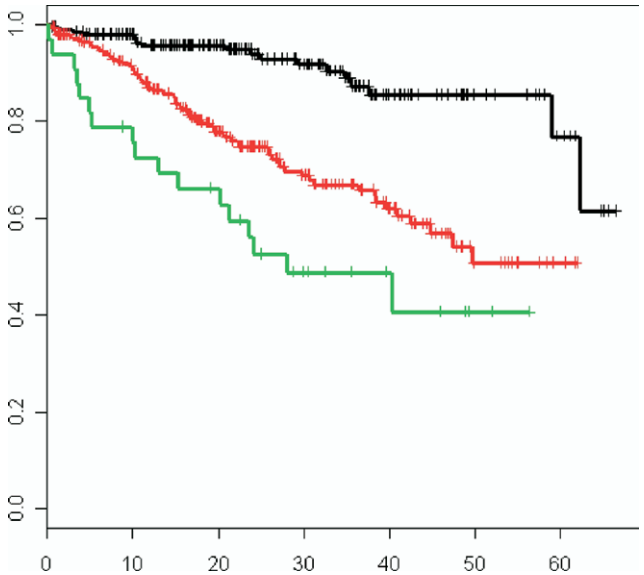

Figure 5. Survival stratified by a combination of histology and pathologic TNM stage.

ing which patients should be considered for induction therapy or for definitive chemoradiotherapy rather than surgical intervention. When we stratified patients on the basis of tumor maximum diameter and histology (both of which are variables that can be approximated by preoperative imaging and needle aspiration, respectively), the addition of PET SUV further improved the definition of prognosis. It is worth noting that because we used pathologic tumor size as a surrogate for radiographically determined size and final pathology as a surrogate for histology determined from a fine-needle aspiration biopsy, it is likely that when radiographic size and histology by fine-needle aspiration are compared with PET that the relative prognostic benefit of determining SUV will improve. Radio-

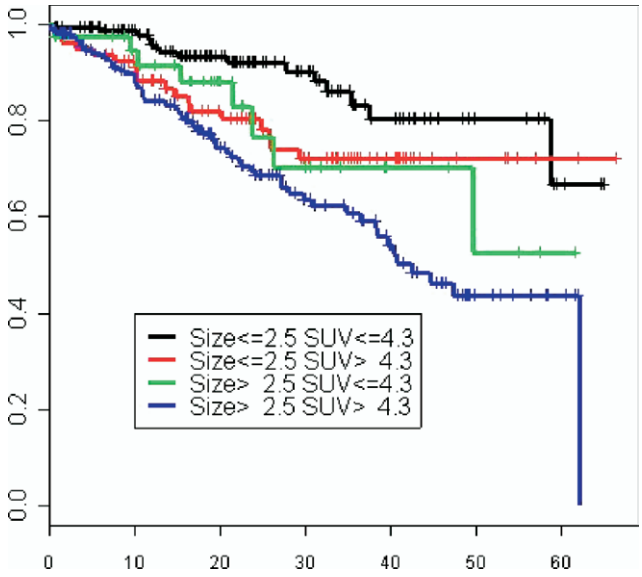

Figure 4. Survival stratified by a combination of the optimal values for standardized uptake value *(SUV)* and tumor size, demonstrating that patients with a tumor size of greater than 2.5 cm and an SUV of greater than 4.3 had significantly worse survival when compared with all other patients ($P < .01$).

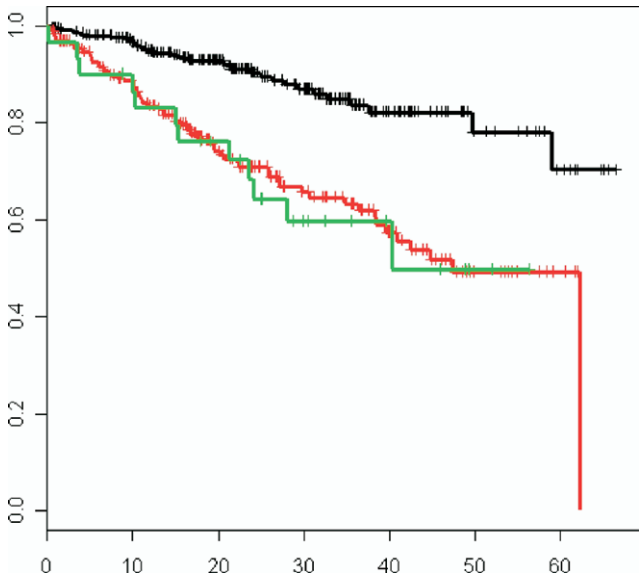

Figure 6. Survival stratified by a combination of histology, standardized uptake value, and tumor size.

GTS

TABLE 2. Preoperative versus postoperative risk group stratification

Preoperative (rows) vs postoperative (columns)	I	II	III
I	190	72	0
II	12	177	6
III	0	3	27

graphic estimates of the size of tumors can vary depending on whether the adjacent lung is inflated at the time of sectioning. It is possible that image-based estimates of tumor size might vary, and therefore the utility of combining image-based tumor size with SUV to define prognosis should be tested formally.

In a model of how a preoperative stratification might work, we construct good-, intermediate-, and poor-prognosis groups based on the postoperative prognostic variables of pTNM stage and pathologic histology. Statistical analysis suggested that combinations of potentially available preoperative variables (SUV, tumor size, and histology) allowed construction of good-, intermediate-, and poor-prognosis groups that closely correlated with the postoperative categories in that only 19% of patients moved from one to another group preoperatively to postoperatively. This compares very favorably with the current relationship between clinical and pathologic staging, in which approximately 50% of patients will shift stage. For example, Roberts and colleagues[18] found that T status as determined by means of computed tomographic (CT) imaging was concordant with pathologic T stage in only 56% of patients, with CT overstaging 20% and understaging 24%. Similarly, Cerfolio and coauthors[17] found that after PET/CT and CT imaging, 52% of patients clinically staged as N2 positive were actually N2 negative, and 14% of patients clinically staged as N2 negative were actually pathologically N2 positive.

Validation of the utility of PET SUV as a prognostic variable and the relative benefit of SUV in comparison with imaging measurements and tumor characteristics, such as gene expression, has been undertaken in an attempt to better guide patient therapy at the time of diagnosis rather than after resection.

References

1. Duhaylongsod FG, Lowe VJ, Patz EF Jr, Vaugh Al, Coleman RE, Wolfe WG. Lung tumor growth correlates with glucose metabolism measured by fluorine-18 fluorodeoxyglucose positron emission tomography. *Ann Thorac Surg*. 1995;60:1348-52.
2. Vesselle H, Schmidt RA, Pugsley JM, Li M, Kohlmyer SG, Vallires E, et al. Lung cancer proliferation correlates with [F-18] fluorodeoxyglucose uptake by positron emission tomography. *Clin Cancer Res*. 2000;6:3837-44.
3. Downey RJ, Akhurst T, Gonen M, Vincent A, Bains MS, Larson S, et al. Preoperative F-18 fluorodeoxyglucose-positron emission tomography maximal standardized uptake value predicts survival after lung cancer resection. *J Clin Oncol*. 2004;22:3255-60.
4. Ahuja V, Coleman RE, Herndon J, Patz EF. The prognostic significance of fluorodeoxyglucose positron emission tomography imaging for patients with nonsmall cell lung carcinoma. *Cancer*. 1998;83:918-24.
5. Cerfolio RJ, Bryant AS, Ohja B, Bartolucci AA. The maximum standardized uptake values on positron emission tomography of a non-small cell lung cancer predict stage, recurrence, and survival. *J Thorac Cardiovasc Surg*. 2005;130:151-9.
6. Dhital K, Saunders CAB, Seed PT, O'Doherty MJ, Dussek J. 18F]Fluorodeoxyglucose positron emission tomography and its prognostic value in lung cancer. *Eur J Cardiothorac Surg*. 2000;18:425-8.
7. Higashi K, Ueda Y, Arisaka Y, Sakuma T, Nambu Y, Oguchi M, et al. 18F-FDG uptake as a biologic prognostic factor for recurrence in patients with surgically resection non-small cell lung cancer. *J Nucl Med*. 2002;43:39-45.
8. Jeong H-J, Min J-J, Park JM, Chung J-K, Kim BT, Jeong JM, et al. Determination of the prognostic value of [18F]fluorodeoxyglucose uptake by using positron emission tomography in patients with non-small cell lung cancer. *Nucl Med Commun*. 2002;23:865-70.
9. Kieninger AN, Welsh R, Bendick PJ, Zelenock G, Chmielewski GW. Positron-emission tomography as a prognostic tool for early-stage lung cancer. *Am J Surg*. 2006;191:433-6.
10. Port JL, Andrade RS, Levin MA, Korst RJ, Lee PC, Becker DE, et al. Positron emission tomographic scanning in the diagnosis and staging of non-small cell lung cancer 2 cm in size or less. *J Thorac Cardiovasc Surg*. 2005;130:1611-5.
11. Sasaki R, Komaki R, Macapinlac H, Erasmus J, Allen P, Forster K, et al. [18F]fluorodeoxyglucose uptake by positron emission tomography predicts outcome of non-small-cell lung cancer. *J Clin Oncol*. 2005;23:1136-43.
12. Vansteenkiste JF, Stroobants SG, Dupont PJ, De Leyn PR, Verbeken EK, Deneffe GJ, et al. Prognostic importance of the standardized uptake value on 18F-fluoro-2-deoxy-glucose-positron emission tomography scan in non-small-cell lung cancer: An analysis of 125 cases. *J Clin Oncol*. 1999;17:3201-6.
13. Sugawara Y, Quint LE, Iannettoni MD, Orringer MB, Russo JE, Recker BE, et al. Does the FDG uptake of primary non-small cell lung cancer predict prognosis?: A work in progress. *Clin Positron Imaging*. 1999;2:111-8.
14. Vesselle H, Turcotte E, Wiens L, Schmidt R, Takasagui JE, Lalani T, et al. Relationship between non-small cell lung cancer fluorodeoxyglucose uptake at positron emission tomography and surgical stage with relevance to patient prognosis. *Clin Cancer Res*. 2004;10:4709-16.
15. Borst GR, Belderbos JSA, Boellaard R, Comans EFI, De Jaeger K, Lammertsma AA, et al. Standardised FDG uptake: a prognostic factor for inoperable non-small cell lung cancer. *Eur J Cancer*. 2005; 41:1533-41.
16. Pillot G, Siegel BA, Govindan R. Prognostic value of fluorodeoxyglucose positron emission tomography in non-small cell lung cancer. A review. *J Thorac Oncol*. 2006;1:152-9.
17. Cerfolio RJ, Bryant AS, Ojha B, Eloubeidi M. Improving the inaccuracies of clinical staging of patients with NSLCC: a prospective trial. *Ann Thorac Surg*. 2005;80:1207-13.
18. Roberts JR, Blum MG, Arildsen R, Drinkwater DC Jr, Christian KR, Powers TA, et al. Prospective comparison of radiologic, thoracoscopic, and pathologic staging in patients with early non-small cell lung cancer. *Ann Thorac Surg*. 1999;68:1154-8.

Discussion

Dr Carolyn R. Reed *(Charleston, SC)*. When PET came on the scene, all of us hoped that it would remarkably improve the accuracy of clinical staging. Although PET, especially integrated PET/CT, has increased staging accuracy, the multi-institutional studies have somewhat dampened our initial enthusiasm. As we have seen in other papers today, we must look at subsets of patients to refine the utility of PET.

Recently, there has been a new focus on the intensity of FDG uptake in an individual tumor as a surrogate marker of the biologic behavior of that tumor. This is particularly attractive because assessment of such things as cell proliferation markers and gene expression profiling require tissue samples. Dr Downey's paper gives us insight into the true usefulness of PET and also, I believe, its limitations.

Dr Downey, I think your study corroborates several known things: the size cut-off point between T1 and T2 tumors should be lower than 3 cm; small adenocarcinomas have better prognosis, perhaps because histologies such as BAC and other GGOs are included; and most importantly, postoperative pathologic stage is still the key.

I do have one problem with the preoperative use of SUV_{MAX} because although at your institution an SUV_{MAX} of less than 4.3 might be predictive of good prognosis, that might not be the case at other institutions. SUV is a continuous variable, and a binary cut-off point might not be appropriate. There are many factors, as we all know, that affect the SUV determination. Therefore my first question is this: How do we use your finding of an SUV_{MAX} of 4.3 as a prognostic cut-off point? Should we establish our own institutional cut-off point?

Dr Downey. This is a problem that has been addressed before, and it is a very important point. The standardized uptake value is not standardized between institutions. It is just standardized within an institution. Therefore a tumor measured to have an SUV of 4.3 in one institution might have a different measurement at another institution. There is interest in the nuclear medicine field for trying to come up with equipment similar to the phantom used in CT scanning that would allow standardization of SUV measurements between institutions, but thus far, there is no means of standardizing, and this just should be taken as a relative value.

The second question is about SUV being a continuous variable. I believe that there are good reasons to consider treating SUV as a binary variable. We have observed that patients who have an SUV of less than 5 almost never have nodal metastases. In our earlier article with 100 patients, there are 33 patients who had an SUV of less than 5, and only one of them had a lymph node metastasis. There are 67 patients with an SUV of greater than 5, and the incidence of nodal metastases was 33%. We have made a similar observation in patients with esophageal cancer that prognosis is best defined by dichotomizing around an SUV of 5. Therefore there might be a binary cut-off point, but this will require further research.

Dr Reed. You say that 3 clinical variables are available preoperatively to give 3 prognostic categories, yet many of us do not currently perform biopsies on small, highly PET-positive peripheral lesions before surgical intervention. You also used pathologic T size and not a CT estimate. Do you now recommend biopsy of all lesions to adequately prognosticate and perhaps, in addition, supply tissue for other markers?

Dr Downey. Again, this is a very important point. To determine the prognostic value of PET SUV, what I did was to compare PET with what would be the optimal combination of standard tests, which would be if the fine-needle aspiration preoperatively told you everything about the histology that the final pathology was going to show you, and also radiographically they were able to determine the same size that was determined by means of pathology. Despite this being the best standard that could possibly be obtained, PET still turned out to be an independent predictive variable. Obviously, there is going to be degradation from that high standard in common practice because if we try to measure things radiographically, it will only approximate pathologically measured size, and histology defined from a fine-needle aspiration is never going to be as good as the final pathology. Therefore I believe that the PET SUV might turn out to be the best independent prognostic variable that we have available in the preoperative period. I am not advocating needle biopsies in everyone.

Dr Reed. It was hoped that FDG-PET could be a useful tool in identifying patients at high risk of recurrence with each stage, particularly resected stage I and II disease. Because of your findings, would you still favor a prospective multi-institutional study looking at this issue in a homogeneous population of patients; that is, those with surgically resected stage I and II disease?

Dr Downey. There are about 10 to 12 articles that have been published looking at the prognostic significance of PET SUV. Most recently there was Dr Cerfolio's article from Alabama in which he found results that were different from ours. He found that the PET did predict survival independent of pathologic staging. It was a smaller group of patients but still a very substantial number. I have not talked to him to figure out why we have the results that we have and he has results that he has. Therefore I would still consider this an open question for development. I do not know whether it requires a multi-institutional study.

Dr Reed. Finally, how would you use PET to stratify for neoadjuvant therapy? Tell us what you would recommend for a patient with clinical T2 N0 squamous cell carcinoma with an SUV_{MAX} of 10.

Dr Downey. We have a very low threshold for recommending patients for neoadjuvant therapy; however, I think we might have identified a group of patients who might not benefit. Although it is somewhat of a philosophical question, it is not clear whether it would be worth giving induction therapy to the group with the best prognosis, who have a 3-year survival of 86%. We might be able to reduce the number of patients who get referred for induction therapy.

Dr Reed. I believe, finally, that this is a very important paper because it refines the use of PET, and I want to congratulate Dr Downey on a very nice presentation. Thank you very much.

Dr Downey. Thank you.

Dr Robert J. Cerfolio *(Birmingham, Ala)*. Dr Miller said I can only ask one quick question, and therefore I will try to ask you 2 quick questions.

First of all, I congratulate you on your findings, and I think our findings are relatively similar, and we are corroborating one another. There are some specifics. You grouped stage II and stage III disease together. That might have been a factor, and we can talk

about that later. But I think we do need a multi-institutional prospective study, and I think we should look at neoadjuvant therapy for the patients with stage IB disease. The patients with stage IB disease are the ones we all know do poorly, and I think the SUV_{MAX} helps you identify the patients with stage IB disease who do poorly. Therefore this is my chance to try to get all the world's experts who are here to participate in a study like that, which I think we really need to do.

The second thing is that we need to get the SUV_{MAX} to be the same at different institutions. It is not that hard to do. The nuclear radiologists tell us they can do it. Then an SUV_{MAX} in Brazil will be the same as an SUV_{MAX} in New York or Birmingham.

Could you comment on the role of a prospective study looking at neoadjuvant chemotherapy for a patient with stage IB disease and how we would identify the patients with stage IB disease with integrated PET/CT and not dedicated PET?

Dr Downey. I did not mention it, but in this group we transitioned from just PET to PET/CT in the middle of this study at our own institution. I think one of the most interesting findings was the statistical phenomenon of an interaction that was seen in the patients with clinical stage IB disease, such that if it is a discriminant, it is maybe most effective in telling you who is a good- and a poor-risk patient with clinical stage IB disease. I am not sure that I would use it to start off a multi-institutional trial involving induction therapy at this point because there is no defined role for induction therapy. I think that we need to simply validate that we can identify preoperatively patients with something that correlates with pathologic stage, and that would be best done across institutions. What I think we have here is that it does show there is no relationship between PET and pathology and therefore no additional benefit of adding PET to pathology; therefore we can use pathologic stage as the gold standard, and it would be a relatively straightforward study to do to define prognostic variables preoperatively, have surgeons assign a clinical stage, and then just find out what they find at the operation a week later and see how close we can get. Then after we have done that, shown what variables work, then we can decide who should get induction therapy.

Dr Daniel L. Miller *(Atlanta, Ga)*. Although I enjoyed your presentation, I think the most important point about this is when we look at visceral pleural invasion because we are getting to these smaller tumors—2 cm, 1.5 cm—and they are peripheral, and they are going to have a significant amount of visceral pleural invasion. Your greatest crossover group was stage IA to IB disease, and I think that probably was not measurement but was probably related to visceral pleural invasion. We are looking back now to our own data. There are actual lung cancers less than 3 cm with visceral pleural invasion that are actually doing worse than those of 4 cm without visceral pleural invasion. Did you look at visceral pleural invasion at all?

Dr Downey. I do have a slide that showed that, but I did not display it. The 2 areas, sort of the move from clinical to pathologic stage, if we break it down by T, N, and M, it was not in the N and the M category, it was in the T category, and it was not on measured size. The majority of them, as you pointed out and you picked up on, were going from T1 to T2 or T1 to T4. Less so, but still those are the major areas that fell off.

Dr Miller. Well, I think it is more important now with our pathologists with determining visceral pleural invasion at the time of resection because they are all getting adjuvant treatment now for stage IB disease, and if you are going to do a stage IB disease study preoperatively, we are going to miss some patients because we are not going to be able to identify that visceral pleural.

Dr Anthony P. Yim *(Shatin, Hong Kong)*. Dr Downey, I really enjoyed your paper.

Did you look at delayed SUV values apart from the SUV_{MAX}, and do they affect the outcome?

Dr Downey. No. Our nuclear medicine physicians have been very interested, and I have been sort of peripheral to that work, in trying to see whether there are any other parameters that work better than the SUV_{MAX}. They have tried integrating the SUV across the tumor and various things, and they have never found anything that they think is better in any of these data sets than the SUV_{MAX} of the primary site of disease.

Dr Yim. Just to comment, in my locality we still see a lot of inflammatory disease, such as tuberculosis, and we pay a lot of attention to the delayed value.

Dr Frank C. Detterbeck *(New Haven, Conn)*. Rob, I have just a comment. Maybe you know the answer to this better than I do, but my understanding is that there are a lot of factors that go into SUV, and I do not think it is as simple as just standardizing the machine at one institution versus that at another. It has to do with the activity of the FDG, the timing between the injection, and when you scan. It also has to do with the size of the lesion. Once the lesion is less than 2 cm, you have to correct for the size of the lesion because otherwise you will underestimate it. It also has to do with where in the lung the lesion is because if there is a lot of movement, then you do the volume averaging on the PET scan just like you do on the CT scan. Therefore there are a lot of factors. My understanding is that determining a reproducible and reliable SUV value is not at all straightforward or simple, but perhaps you know more about this than I do or perhaps Rob Cerfolio does.

Dr Downey. I think actually Dr Cerfolio said it would be easier to standardize between institutions. I do not think it will be particularly that easy. It is the time between the injection versus when it is scanned, the dose that is given. Your point about the tumor size, though, I think is a very important one that has nothing to do with the institution. The smaller the lesions are, the more there is going to be a volume averaging. The CT scan is obtained over the length of time that the PET scan lasts, and there will be sort of a movement up and down, both smearing out the sort of SUV value over a larger volume—relatively larger volume—but also degrading the quality of the CT scan. Therefore these are all things that have to be taken into account as we work on the preoperative clinical variables.

Dr John Howington *(Cincinnati, Ohio)*. Rob, nice presentation. There was discussion about needle biopsies and no-needle biopsies, and then the discussion led over into doing induction therapy for lesions with a high SUV. Being in Cincinnati, in the histoplasma belt, I just cringe at the thought of a patient with a solitary pulmonary nodule, with SUVs of 8, 9, and 10 that are histoplasma, being mistakenly treated for cancer. The idea of doing induction therapy in a patient without a tissue diagnosis is hard for us to swallow.

Dr Downey. I was not advocating or arguing against fine-needle aspirations. That was not the point of the paper.

Dr Cerfolio. Well, since you invited me to answer that, Frank, I would say that the formula that goes into the SUV_{MAX} is not dependent on where the nodule is located. It is not dependent on the size of the nodule. It is a very specific formula that looks at the activity at a pixel that can add some of those variables. Actually, those are probably less than 10% differences from talking to the GE guys who do this every day. It has to do with the weight of the patient and the injected dose, and those are very easily corroborated among centers and organized. Therefore I do not think it is that hard to get an SUV_{MAX} to be the same across the world. I will shut up now.

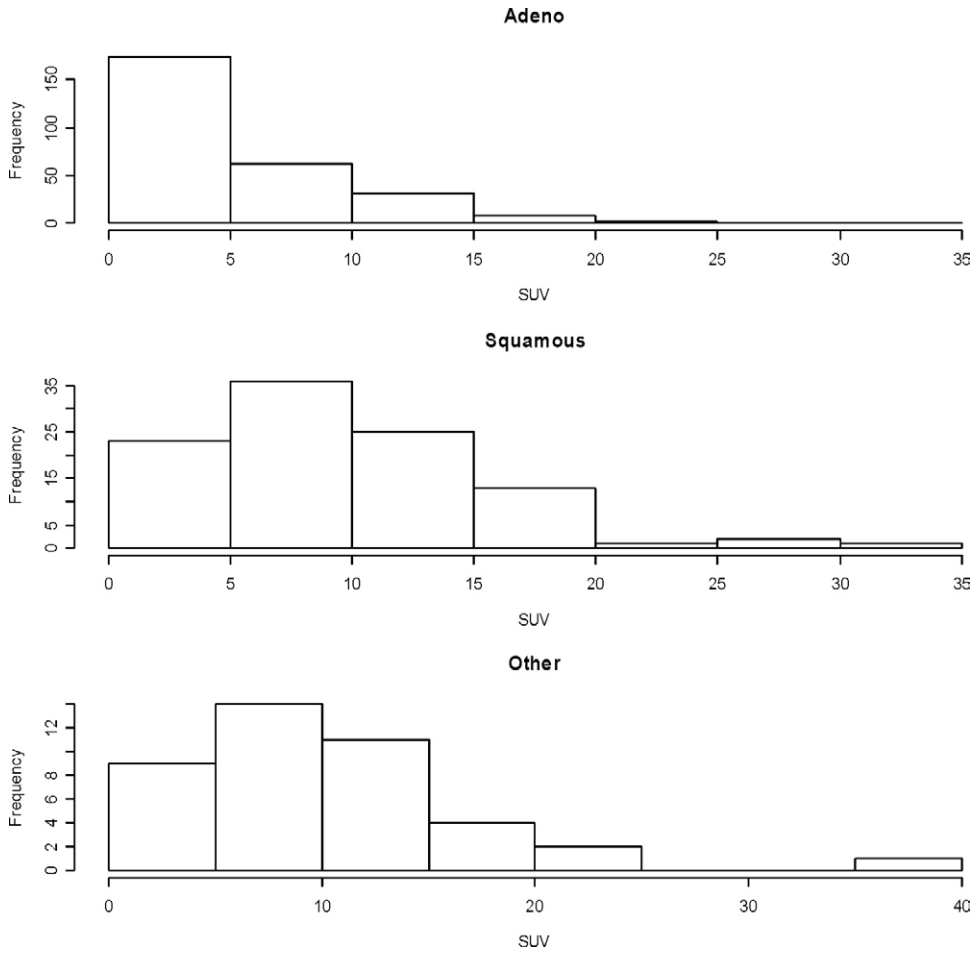

Figure E1. Histograms of distribution of standardized uptake value *(SUV)* within each histologic type. The median SUV for adenocarcinoma *(Adeno)* was significantly different from the squamous and other histologies ($P < .01$). The median values for squamous carcinoma and histologies categorized as "other" were not different ($P = .69$).

GTS

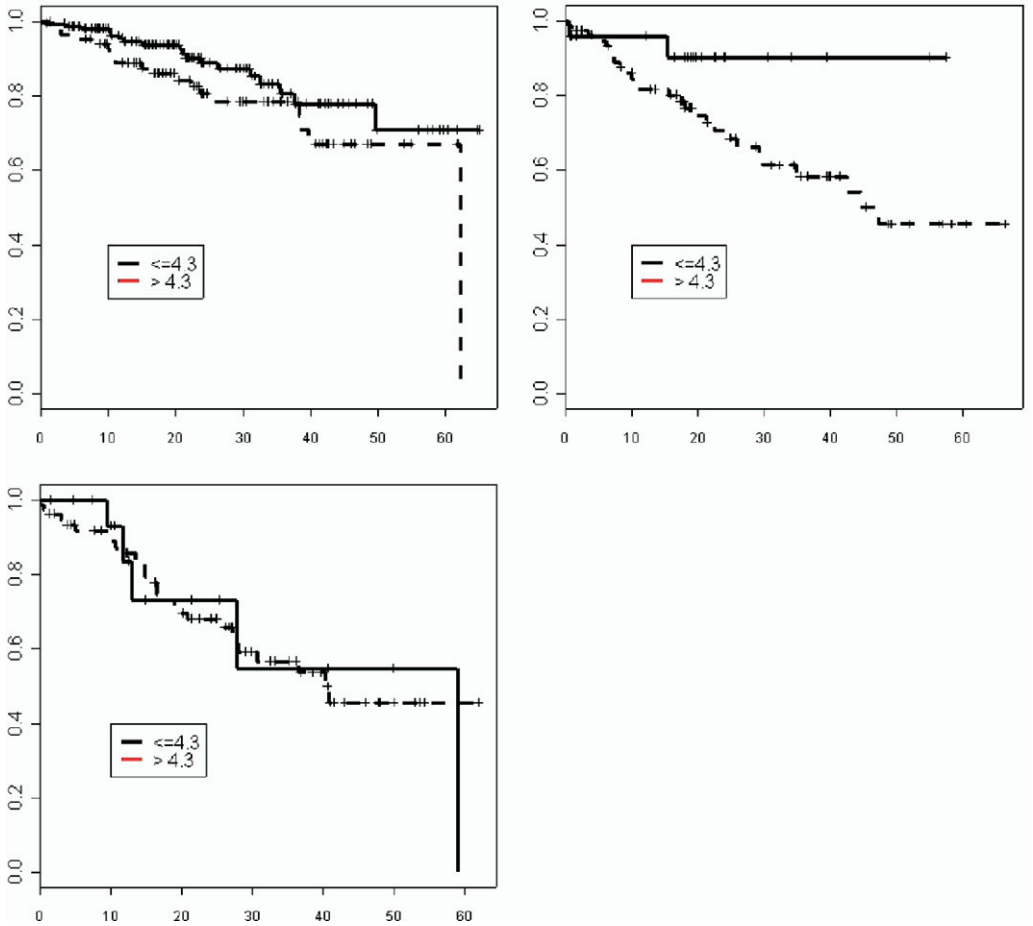

Figure E2. Survival curves for patients with clinical stage IA (top left), IB (top right), and II to IV (bottom left) disease stratified by the optimal value for standardized uptake value (SUV; 4.3) demonstrating interaction between SUV and clinical TNM stage.

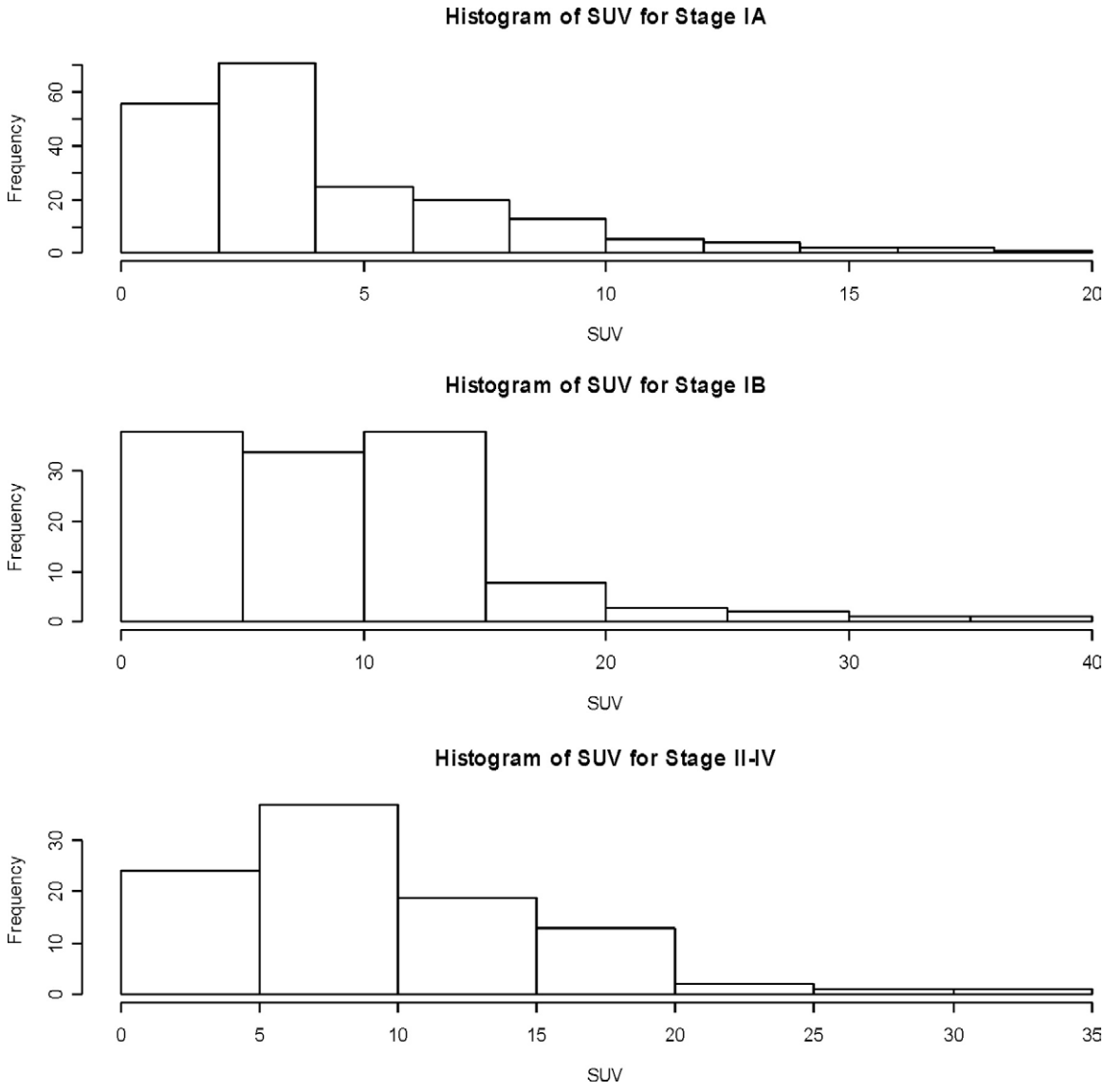

Figure E3. Histograms of the distribution of standardized uptake value *(SUV)* within each stage group. SUV in stage IA disease is significantly lower than that in stage IB disease ($P < .01$) and stage II to IV disease ($P < .01$). SUV distribution in stage IB and stage II to IV disease was not different ($P = .39$).

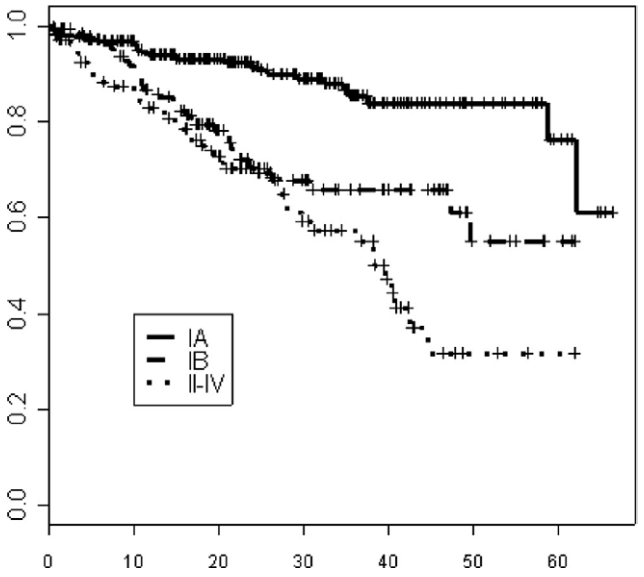

Figure E4. Survival after resection stratified by pathologic TNM stage (IA vs IB vs II-IV).

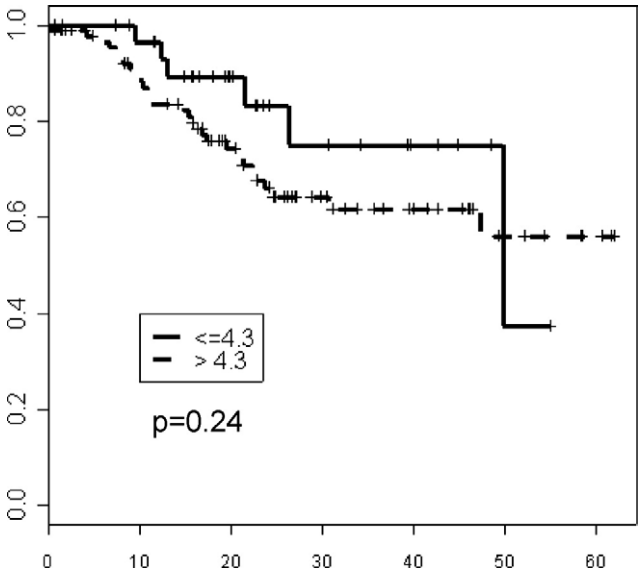

Figure E6. Pathologic stage IB disease stratified by optimal standardized uptake value.

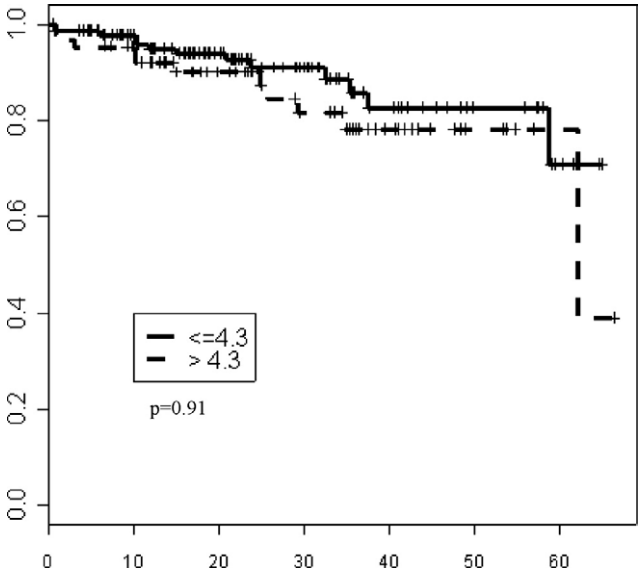

Figure E5. Pathologic stage IA disease stratified by optimal standardized uptake value.

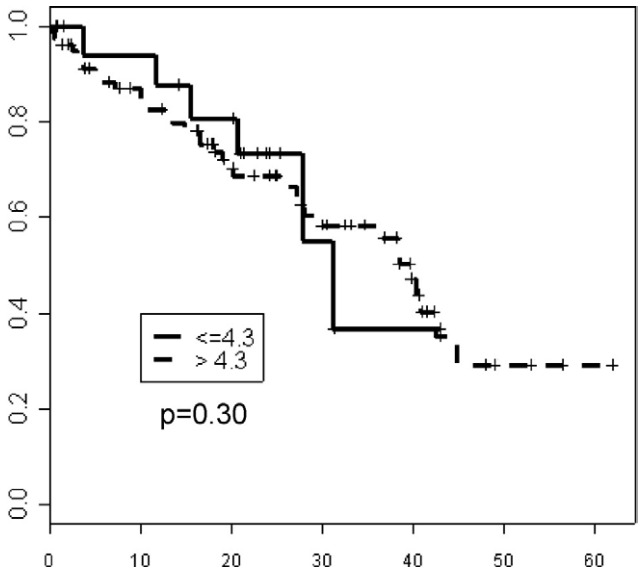

Figure E7. Pathologic stage II to IV disease stratified by optimal standardized uptake value.