Prognostic Value of Fluorodeoxyglucose Positron Emission Tomography in Non-small Cell Lung Cancer: A Review

Giancarlo Pillot, MD, * Barry A. Siegel, MD, †‡ and Ramaswamy Govindan, MD*‡

Abstract: Non-small cell lung cancer (NSCLC) remains a leading cause of cancer morbidity and mortality. Adjuvant chemotherapy improves survival in resected early-stage NSCLC. However, a significant proportion of patients with early-stage lung cancer are cured by surgery alone. There are no reliable clinical or molecular markers to predict outcomes after surgery in early-stage NSCLC. Positron emission tomography with 2-[18F]fluoro-2-deoxy-D-glucose (FDG-PET) improves the accuracy of staging work-up in NSCLC. The standardized uptake value, a commonly used semiquantitative measure of FDG uptake, correlates with tumor doubling time and indices of cell cycling. Therefore, FDG-PET may be a useful predictor of outcome independent of its role in tumor staging. In this review, we critically examine the published studies on the utility of FDG-PET as a prognostic tool in patients with NSCLC and provide direction for future research.

Key Words: Non-small cell lung cancer, 2-[18F]fluoro-2-deoxy-D-glucose positron emission tomography, Prognosis.

Lung cancer is the leading cause of cancer-related death in both men and women in the United States, with an estimated 172,570 new cases and 163,510 patients dying from this disease in 2005. More than 87% of patients with lung cancer are diagnosed with non-small cell lung cancer (NSCLC); of these, approximately 30% will present with resectable disease. It has now become an accepted practice to offer adjuvant chemotherapy for patients with resected stage IB and II NSCLC based on the results of several prospective studies that have demonstrated a statistically and clinically significant survival advantage with this approach. However, these studies have also demonstrated that nearly 60% of patients with resected stage IB NSCLC and 45% of patients with stage II NSCLC will not have recurrent disease after resection even without adjuvant therapy. Unfortunately there are no reliable clinical, radiological, or molecular predictive markers to identify those patients at high risk of recurrent disease.

A number of factors other than tumor stage have been shown to have prognostic significance in patients with resected NSCLC. However, the available data from retrospective studies are hard to interpret as they are derived from a mixed group of patients with differing stages of disease who undergo various treatments. Specifically, among patients with stage I disease, clinical and pathologic features such as tumor size, extent of surgery, and factors such as visceral pleural invasion or high mitotic rate have demonstrated some correlation with prognosis. Other biological markers, such as Ki-67 labeling index (a cell proliferation marker), and p53 mutation status, have also provided correlative information. However, the published studies have not yielded consistent results for all markers and have been often hampered by small sample size. Moreover, many individual markers lose their unique prognostic significance when they are subjected to multivariate analysis. Although these molecular approaches hold promise for prognostic risk stratification of patients with early-stage NSCLC, none has yet been validated prospectively. Additionally, the need for further testing of original tumor tissue would add further expense, as well as a clinical delay while awaiting the results. An easy-to-use, readily available clinical variable or set of variables thus may be more useful and more quickly incorporated into practice both in academic centers and in the community.

Positron emission tomography (PET) with the glucose analogue 2-[18F]fluoro-2-deoxy-D-glucose (FDG) is a useful imaging method for diagnosing, staging, and monitoring therapy of several malignancies. FDG-PET enhances the accuracy of clinical staging and excludes patients with distant metastasis or advanced nodal disease from futile surgery. Although most studies of FDG-PET in NSCLC have focused on its use in staging, the intensity of FDG uptake in an individual tumor also may serve as a surrogate marker of the biological behavior of that tumor. The purpose of this study was to review the role of the semiquantitative assessment of tumor FDG uptake as a prognostic indicator for NSCLC.

Background

The interested reader is referred to other reviews of the methodology for FDG-PET. Briefly, FDG-PET is performed by the injection of FDG, a radiopharmaceutical that is taken up by most cancer cells to a greater extent than by most normal tissues. FDG-PET images are acquired approximately 60 minutes after injection, when the uptake of FDG is at its peak. The uptake of FDG is then quantified and compared with a standardized uptake value (SUV), which is calculated by dividing the radiotracer activity in the tissue by the radiotracer activity in the blood. The SUV is a commonly used semiquantitative measure of FDG uptake, and it correlates with tumor doubling time and indices of cell cycling. Therefore, FDG-PET may be a useful predictor of outcome independent of its role in tumor staging.

References

normal cells. Intracellularly, FDG is phosphorylated in a manner similar to glucose; however, the resultant FDG-6-phosphate is not further metabolized and is retained (“metabolically trapped”) in the cell. When the F-18 radionuclide decays, the resulting annihilation photons are detected by the PET scanner; the data from many detected decay events are transformed into a three-dimensional representation of signal intensity. Areas in the body that produce significant signal as a result of increased FDG uptake are visible on FDG-PET images. However, this signal is attenuated by body tissues, which absorb or scatter the annihilation photons as they pass through the patient. Correction of the emission data for attenuation is performed with a separate transmission scan. This can be obtained either with an extrinsic radiation source (such as germanium-68), as is typically done in a conventional PET scanner, or with X-rays, as is done with the computed tomographic (CT) component of the study from a dedicated PET/CT scanner. However, these different methods of attenuation correction can also contribute to differences in the perceived PET signal.24

Although tumor FDG uptake can be quantified by several different methods,25 the one most commonly used in clinical practice is the standardized uptake value (SUV).26 The SUV is defined as:

\[
SUV = \frac{C_{ROI}}{(A/Wt)}
\]

where \(C_{ROI}\) is the decay-corrected tracer concentration in \(\mu Ci/cm^3\) measured on the PET image within a region of interest (ROI); \(A\) is the total injected activity of the radiopharmaceutical in \(\mu Ci\); and \(Wt\) is the patient’s body weight in grams. The denominator of this equation (\(A/Wt\)) is the average concentration of the tracer within the body (assuming that there has been no elimination of the tracer). The SUV for any ROI can be either the mean value within the region (SUVmean) or the value for the voxel within the region having the maximal signal (SUVmax). The SUVmean is highly dependent on the method used for ROI placement. The ROI is drawn around a lesion that may not be sharply demarcated, portions of which may include an admixture of tumor and normal tissue that cannot be distinguished because of the limited resolution of the PET scanner; inclusion of this lower activity tissue in the ROI will artificially decrease the measured signal. Additionally, the SUVmean reflects the variability in tracer uptake from one region to another within the tumor, including areas with low activity as a result of poor blood supply or necrosis. The SUVmax, as the result for a single pixel, is subject to greater statistical variability because of the noise in the reconstructed PET image. With small lesions, both measures will be reduced by the effects of partial volume averaging, the error introduced when the lesion size is less than twice the spatial resolution of the scanner.27

The SUV has a number of other limitations as a measure of tracer uptake in a tumor28 and can be affected by several factors not specifically related to the intrinsic metabolic activity of the tumor (Table 1). One important factor in the variability of SUV is the interval from the injection of FDG to imaging. Although imaging was initiated 40-60 minutes after injection of FDG in most of the studies that we review herein, the tumor uptake of the tracer is still increasing at this time, and the plateau of uptake does not typically occur for several hours. In one series of patients with untreated NSCLC, the 60-minute SUV was 46% lower than the value measured at the plateau of the uptake phase.29 The plasma glucose level during the period of FDG uptake can also affect the SUV because glucose competes with FDG for uptake into cancer cells. Insulin levels are also important because insulin increases FDG uptake in skeletal muscle and adipose tissue, making less FDG available for uptake in the tumor; this is why oncological FDG-PET studies are performed after 4 or more hours of fasting. The SUV is also affected by body habitus.26 Because of the relatively low uptake of FDG in adipose tissue, the apparent SUVs of tumors (and of other tissues) will be higher in obese patients than in those of normal body weight. To address this problem, it has been suggested that the SUV should be normalized to the patient’s lean body mass or body surface area.30 Physical effects related to scan acquisition also affect the SUV. Because data acquisitions for PET typically take several minutes at each scanner position, the respiratory motion of lung tumors is an additional basis for partial volume averaging; again, this effect is greatest with small tumors. With PET/CT, this respiratory variation of tumor location has been demonstrated to produce additional variability in SUV of up to 30%, because of respiratory misregistration of the CT and PET data sets, which results in incorrect attenuation correction of the emission PET data.24 Finally, the method of image reconstruction affects SUV, with certain methods of reconstruction leading to underestimates of SUV by up to 30%.31 Accordingly, SUVs determined with different scanners, acquisition protocols, and reconstruction methods may not be comparable.

Despite all these challenges, the SUV can be used as a reasonably reliable semiquantitative measurement of tumor uptake. Variability in uptake related to fasting blood glucose level, time from FDG injection to imaging, and method of SUV determination can be standardized. SUVmax normalized to lean body mass has been shown to be reproducible in patients with lung cancer across repeat determinations; repeated values were within a mean of 11.3% of each other.32 It should be noted, however, that this reproducibility was only tested in tumors larger than 2 cm in size.

**TABLE 1. Factors Affecting SUV Determination**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood glucose level at time of FDG injection</td>
<td>The level of blood glucose can affect the uptake of FDG into cancer cells.</td>
</tr>
<tr>
<td>Plasma glucose level at time of FDG injection</td>
<td>Glucose competes with FDG for uptake into cancer cells.</td>
</tr>
<tr>
<td>Body habitus</td>
<td>Insulin levels are important because insulin increases FDG uptake.</td>
</tr>
<tr>
<td>Time of image acquisition after injection of FDG</td>
<td>Time from FDG injection to imaging affects SUV.</td>
</tr>
<tr>
<td>Lesion characteristics</td>
<td>Variability in uptake related to lesion characteristics.</td>
</tr>
<tr>
<td>Size/partial volume effects</td>
<td>Size/partial volume effects can affect SUV.</td>
</tr>
<tr>
<td>ROI placement</td>
<td>ROI placement can affect SUV.</td>
</tr>
<tr>
<td>Respiratory variation of position of lesion of interest</td>
<td>Respiratory variation of tumor location can affect SUV.</td>
</tr>
<tr>
<td>Method of image reconstruction and attenuation correction</td>
<td>Reconstruction methods may affect SUV.</td>
</tr>
</tbody>
</table>

SUV, standardized uptake value; FDG, 2-[18F]fluoro-2-deoxy-D-glucose; ROI, region of interest.
Biological Correlates of FDG-PET Signal Intensity

The results of FDG-PET have been correlated with several presumed markers of tumor aggressiveness. In one study of patients referred to Duke University for FDG-PET evaluation of indeterminate focal pulmonary abnormalities, serial chest radiographs or CT scans were assessed for doubling time of the lesion of interest and correlated with the SUVmax. In this study, an SUVmean more than or equal to 2.5 was found in 49 of 53 lesions proved to be malignant. Additionally, the SUVmean was found to correlate with lesion doubling time. The SUV for FDG has also been shown to correlate with lung cancer proliferation as measured by indices of cell proliferation. Higashi et al.34 demonstrated a correlation between SUV (measured within a ROI encompassing all pixels with activity >90% of the maximal pixel activity in the primary tumor) and tumor cell proliferation of subsequently resected samples as measured by proliferating cell nuclear antigen labeling. These investigators found a much weaker correlation between tumor cell density and SUV and reported that bronchoalveolar carcinomas tended to have lower SUVs than other histological types of NSCLC. Vesselle et al.35 also studied the correlation of FDG uptake with tumor proliferation in a series of patients, most of whom went on to surgical resection (35 of 39 study patients). They found a positive correlation between SUVmax and tumor cell proliferation as measured by Ki-67 labeling, with the correlation being most robust in patients with stage I NSCLC. They also noted an association between degree of cellular differentiation as determined by pathologic grading by light microscopy and SUVmax. In another study of pulmonary adenocarcinomas, Higashi et al.36 also demonstrated a correlation between SUV for FDG (measured within a ROI encompassing all pixels with activity >80% of the maximal pixel activity in the primary tumor) and perceived indices of aggressiveness, such as pleural involvement, vascular invasion, or lymphatic invasion. It should be noted, however, that PET scanning with another radiopharmaceutical, 3’deoxy-3’-18F-fluorothymidine may be better correlated with tumor proliferation (as assessed by Ki-67 staining) than FDG.37

In aggregate, these data demonstrate that the SUV measured by FDG-PET correlates with several indices of tumor growth and aggressiveness. It is a reasonable postulate, therefore, that SUV may also be associated with outcome.

Summary of Published Reports Correlating SUV with Patient Outcome

Ten published studies from several countries have correlated prognosis with SUV in NSCLC. All have been single-institution, retrospective studies. One of these studies failed to demonstrate that SUV correlated with prognosis, whereas the remaining nine demonstrated a statistically significant result. Tables 2 and 3 summarize the technical aspects of scan acquisition, patient populations investigated, and results of these studies.

In a study of FDG-PET in mediastinal staging of NSCLC at the University of Michigan,38 SUVmax of the primary tumor was measured and compared with outcome. Of a total of 97 identified patients who underwent FDG-PET, 38 met enrollment criteria of histologically proven NSCLC that had been previously untreated. Patients who died of non-cancer causes or were lost to follow-up were excluded. Approximately 31% of these patients had early-stage (I or II) disease, and the median follow-up period was 26.5 months. These investigators chose the median SUV of their patient population (8.72) as the cutoff to discriminate between high- and low-uptake groups. These investigators found no statistically significant difference between the survival of patients in the high- and low-uptake groups. They did, however, note a nonsignificant trend toward better survival in patients with SUVmax less than or equal to 8.72 who also were free of N2 or M1 disease. This study is limited by its small size and by the broad ranges of both stage and histological subtype among the patients evaluated; therefore, few conclusions can be drawn from its failure to demonstrate a statistically significant association between SUV and stage. The squamous cell carcinomas had a higher SUV than the adenocarcinomas, although it is not clear whether this SUV difference correlates with tumor size.

Vansteenkiste et al.39 analyzed the outcome of Belgian patients assessed by FDG-PET who had been enrolled in protocols investigating lymph node staging and the evaluation of solitary pulmonary nodules. They analyzed the course of 125 patients, almost all of whom were surgically staged. The mean follow-up time of patients in this study was 19 months. When more than one lesion was noted on PET, the SUVmax of the primary tumor was used. In this study, approximately half of the patients had early-stage (I or II) disease, and most of these were treated with surgery alone. The authors chose a SUVmax cutoff value of 7 by testing several different cutoffs for statistical significance and using the value with the greatest statistical significance. They did find a continuum of statistically significant cutoffs between SUV values of 6 and 11. In addition to finding that survival was related to performance status, stage, and tumor histology, they demonstrated a statistically significant difference in survival among patients based on SUVmax. In a subgroup analysis of patients of all stages who were treated with surgery at some point in their clinical course, the survival difference based on SUVmax persisted. In a multivariate analysis, stage, performance status, and SUVmax were independent prognostic variables.

Investigators at Duke University also reviewed the prognostic significance of SUV.40 They reviewed all 155 patients seen at their institution between 1992 and 1996 with a new diagnosis of NSCLC and an FDG-PET study. The median follow up of the surviving patients was 20.9 months. A mean SUV was determined on the nodule on the FDG-PET image with the maximal uptake, and the ROI used was a circular area adjusted to include 80% of peak counts in the nodule. Approximately 45% of patients had early-stage (I or II) disease. The investigators found that stage, lesion size, and SUV greater than 10 were correlated with poor prognosis (median survival of 11.4 months versus 24.6 months for those with SUV ≤10). In a multivariate analysis, the significant association of SUV with poorer outcome persisted.
TABLE 2. Technical Methods of Scanning

<table>
<thead>
<tr>
<th>Author, year</th>
<th># scanners used</th>
<th>Fasting time/glucose</th>
<th>Time from injection to imaging</th>
<th>Attenuation correction method</th>
<th>Reconstruction method</th>
<th>Method of SUV determination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sugawara et al.</td>
<td>2</td>
<td>≥4 hrs ≤126 mg/dl</td>
<td>50-70 min</td>
<td>Pre-injection transmission scan</td>
<td>NR</td>
<td>SUV max of maximal region of uptake</td>
</tr>
<tr>
<td>Vansteenkiste et al.</td>
<td>1</td>
<td>≥6 hrs NR</td>
<td>50-70 min</td>
<td>Pre-injection transmission scan</td>
<td>Filtered back projection</td>
<td>SUV max</td>
</tr>
<tr>
<td>Ahuja et al.</td>
<td>2</td>
<td>≥4 hrs NR</td>
<td>At least 30 min</td>
<td>Pre-injection transmission</td>
<td>NR</td>
<td>SUV mean of ROI containing 90% of peak counts</td>
</tr>
<tr>
<td>Dhital et al.</td>
<td>1</td>
<td>≥6 hrs NR</td>
<td>NR (mean 81 min)</td>
<td>Transmission; unclear timing</td>
<td>NR</td>
<td>SUV mean corrected for lean body mass</td>
</tr>
<tr>
<td>Higashi et al.</td>
<td>1</td>
<td>≥4 hrs ≤120 mg/dl</td>
<td>40 min</td>
<td>Pre-injection transmission scan</td>
<td>NR</td>
<td>SUV max</td>
</tr>
<tr>
<td>Jeong et al.</td>
<td>1</td>
<td>≥6 hrs NR</td>
<td>60 min</td>
<td>Pre-injection transmission scan</td>
<td>NR</td>
<td>SUV max</td>
</tr>
<tr>
<td>Downey et al.</td>
<td>2</td>
<td>≥6 hrs NR</td>
<td>NR</td>
<td>Post-injection transmission scan</td>
<td>Iterative</td>
<td>SUV max</td>
</tr>
<tr>
<td>Sasaki et al.</td>
<td>1</td>
<td>≥6 hrs ≤150 mg/dl</td>
<td>60 min</td>
<td>Post-injection transmission scan</td>
<td>Iterative</td>
<td>SUV max</td>
</tr>
<tr>
<td>Borst et al.</td>
<td>1</td>
<td>≥6 hrs glucose measured</td>
<td>60 min</td>
<td>Post-injection transmission scan</td>
<td>Iterative</td>
<td>SUV max</td>
</tr>
<tr>
<td>Cerfolio et al.</td>
<td>2</td>
<td>≥4 hrs NR</td>
<td>60 min</td>
<td>Pre-injection transmission scan for PET</td>
<td>Iterative</td>
<td>SUV max</td>
</tr>
</tbody>
</table>

SUV, standardized uptake value; ROI, region of interest; CT, computed tomography; PET, positron emission tomography; NR, not reported.

*Difluous.

Dhital et al. 41 studied the predictive value of FDG-PET in 77 consecutive patients with biopsy-confirmed NSCLC referred to a single surgeon. The interpretation of this study is limited by the lack of clear reporting of variables such as the range of time from FDG injection to image acquisition and the fasting glucose levels. The investigators seem to use a SUV mean corrected for lean body mass. The investigators reported on several different SUVs as a discriminatory factor between patient with poor and good prognoses. They found a difference in 12-month survival based on a SUV cutoff of 20, although the proportion of patients surviving at this SUV cutoff at 1 year is not clearly reported. The stage distribution of the patients, treatments provided (presumably primarily surgical), and corrections for other known prognostic factors were not provided.

A Japanese study by Higashi et al. 42 evaluated 57 patients with NSCLC who had been surgically treated and had also undergone FDG-PET as part of their initial evaluation. This small patient population was a pure surgical group specifically excluded patients with metastasis at presentation and those who had received adjuvant therapy. Similar to the Duke study, the investigators used the SUV mean for an irregular ROI including values that were more than 90% of the maximal uptake in the ROI on the FDG-PET image with the most intense uptake. The median follow-up interval was 33.5 months. These investigators found a broad range of statistically significant SUV cutoff values and used the most significant value at an SUV cutoff of 5. There was a statistically significant overall and disease-free survival advantage for patients with SUV values below this cutoff value. Specifically, patients with stage I disease and an SUV value of less than or equal to 5 had a 5-year expected disease-free survival of 88%, whereas those with an SUV value greater than 5 had an expected disease-free survival of less than or equal to 17%. These investigators found SUV to be the most significant independent factor affecting prognosis; it was a better predictor than tumor size (greater than versus less than 3 cm) or stage (stage I versus stage II or III), although it should be noted that patients with stage II or III disease represented only approximately 20% of the total population in this study.

A Korean study by Jeong et al. 43 reviewed 73 patients who had a diagnosis of NSCLC and had FDG-PET performed. Both SUV mean and SUV max of the primary lesion were determined, although only the SUV max was used for correlation with prognosis. The authors state that surgical staging was performed in all but two patients. They used a SUV max value of 7 as the cutoff value but did not comment on how this cutoff value was chosen. Approximately two thirds of their patients had stage I or II disease. They found differences in the means of SUV max for different histological types of NSCLC, with squamous cell carcinoma having a
higher $\text{SUV}_{\text{max}}$ than adenocarcinoma, and bronchioalveolar carcinoma having the lowest mean $\text{SUV}_{\text{max}}$; however, the investigators did not independently analyze histology-specific SUV. They found a statistically significant difference in overall survival between the high-$\text{SUV}_{\text{max}}$ ($\geq 7$) and low-$\text{SUV}_{\text{max}}$ ($>7$) groups. They also found differences in survival with respect to stage (stage I to IIIA versus IIIB plus IV) and histological subtype. However, on multivariate analysis, only stage and SUV persisted as independent prognostic variables.

Downey et al.\textsuperscript{44} published a report on the prognostic value of FDG-PET in 100 patients surgically treated for NSCLC at Memorial-Sloan Kettering Cancer Center. The authors specifically excluded patients treated with any modality other than surgery and included only patients who underwent a R0 resection (complete resection with no resid-

---

**TABLE 3. Summary of Studies Relating SUV to Prognosis in Patients with NSCLC**

<table>
<thead>
<tr>
<th>Author, year</th>
<th>$N$</th>
<th>Histology</th>
<th>Stage</th>
<th>Treatment</th>
<th>Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sugawara et al.\textsuperscript{38}</td>
<td>38</td>
<td>Squamous Adenocarcinoma Large cell</td>
<td>I 50% II 32% III 18% IV</td>
<td>Resection Nonsurgical</td>
<td>Median survival (mo)$^a$ 76% SUV=8.72 NR SUV≤8.72 25.8</td>
</tr>
<tr>
<td>Vansteenkiste et al.\textsuperscript{39}</td>
<td>125</td>
<td>Squamous Adenocarcinoma Large cell</td>
<td>I 54% II 25% III 21%</td>
<td>Resection Nonsurgical</td>
<td>2-year survival 73% SUV&gt;7 43% SUV≤7 83%</td>
</tr>
<tr>
<td>Ahuja et al.\textsuperscript{40}</td>
<td>155</td>
<td>Squamous Adenocarcinoma Large cell</td>
<td>I or II 37% III 34% IV 7%</td>
<td>unspecified</td>
<td>Median survival (mos.) SUV&gt;10 11.4 SUV≤10 24.6</td>
</tr>
<tr>
<td>Dhital et al.\textsuperscript{41}</td>
<td>77</td>
<td>Squamous Adenocarcinoma Large cell</td>
<td>Not reported. I 58% II 23% III 13% Other</td>
<td>Presumably surgical, specifics not reported</td>
<td>Median survival (mos.) SUV&gt;20 6 SUV&lt;20 32$^b$</td>
</tr>
<tr>
<td>Higashi et al.\textsuperscript{42}</td>
<td>57</td>
<td>Squamous Adenocarcinoma Large cell</td>
<td>14% I 60% IA 41% II 23% III</td>
<td>All surgical Non-surgical</td>
<td>5-year survival$^b$ SUV&gt;5 20% SUV≤5 90%</td>
</tr>
<tr>
<td>Jeong et al.\textsuperscript{43}</td>
<td>73</td>
<td>Squamous Adenocarcinoma Large cell</td>
<td>51% I 41% II 3% III</td>
<td>Resection Non-surgical</td>
<td>67% Not specifically reported Kaplan Meyer curves reported.</td>
</tr>
<tr>
<td>Downey et al.\textsuperscript{44}</td>
<td>100</td>
<td>Squamous Adenocarcinoma Large cell</td>
<td>24% I 41% II 3% III</td>
<td>All surgical, with R0 resection</td>
<td>2-year survival SUV&gt;9 68% SUV≤9 96%</td>
</tr>
<tr>
<td>Sasaki et al.\textsuperscript{45}</td>
<td>162</td>
<td>Squamous Adenocarcinoma NOS</td>
<td>43% I 46% II 10% III</td>
<td>Resection Definitive XRT</td>
<td>57% 2-year survival SUV=5 94% SUV&gt;5 65%</td>
</tr>
<tr>
<td>Borst et al.\textsuperscript{46}</td>
<td>51</td>
<td>Squamous Adenocarcinoma large NOS</td>
<td>33% I or II 25% III 20%</td>
<td>All definitive XRT</td>
<td>2-year survival SUV&lt;15 60% SUV≥15 27%</td>
</tr>
<tr>
<td>Cerfolio et al.\textsuperscript{47}</td>
<td>315</td>
<td>Squamous Adenocarcinoma Other</td>
<td>54% Ia 33% Ib 13% II</td>
<td>Resection Nonsurgical</td>
<td>Mean survival 71% SUV&lt;10 3.2 yrs SUV≥10 1.6 yrs</td>
</tr>
</tbody>
</table>

$^a$Not statistically significant.

$^b$Not specifically reported. Estimated from Kaplan-Meyer curve ROI, region of interest; NR, not yet reached at time of publication; NOS, not otherwise specified; BAC, bronchioalveolar carcinoma; SUV, standardized uptake value. Not all percentages total 100% because of rounding.
Squamous cell carcinoma versus adenocarcinoma was also found on univariate analysis to be a statistically significant prognostic variable, but on multivariate analysis, only size, SUV_{max}, and their interaction significantly predicted survival. Using both a cutoff size greater than 3 cm and SUV_{max} greater than 9, the authors were able to determine a particularly poor prognosis subgroup with a 3-year survival of only 47%. Few other authors have reported the use of significant discriminators on multivariate analysis in tandem to create a more clinically robust prognostic score.

Sasaki et al. reviewed 162 consecutive patients at M.D. Anderson Cancer Center who had at least 6 months of follow-up and had undergone FDG-PET before potentially curative treatment for NSCLC, although only approximately half the patients had early-stage (I or II) versus locally advanced (stage III) disease. Surgery was the primary modality in 57% of the patients, with the rest receiving radiotherapy as the backbone of their definitive therapy. The median follow-up duration was only 17 months in this study. The SUV_{max} was determined for both the primary lesion and the regional lymph nodes. A SUV_{max} cutoff value was determined by testing several different values for statistically significant discriminatory ability, and the investigators were able to separate groups with good (94% 2-year overall survival) and poor (65% overall survival) survival groups (p = 0.02). They found that this significant difference in outcomes based on SUV persisted regardless of therapeutic modality (surgery based versus radiation based). They also found an SUV_{max} cutoff value of 5 of the primary tumor predicted differences in metastasis-free survival and loco-regional disease-controlled survival. They did not find the SUV_{max} of any involved lymph nodes to predict outcome at any SUV_{max} tested. On multivariate analysis, primary tumor SUV_{max}, age, N stage, and performance status remained predictive of overall survival. Although limited by short follow-up, this study is strengthened by its large size, which allowed this multivariate analysis controlling for several known prognostic features, as well as the demonstration that the prognostic value of SUV persists regardless of treatment modality.

Borst et al. evaluated a series of 51 patients diagnosed with NSCLC who were treated with definitive radiation therapy between January 1999 and November 2001 who underwent PET scan as part of staging evaluation. SUV_{max} and SUV_{mean} were both determined in a manually drawn ROI around the primary lesion. They also performed a correction to each SUV based on the measured pre-scan blood glucose concentration. The authors investigated and did not find a correlation between SUV_{max} and tumor volume (measured by CT) and lymph node status. The investigators did find SUV_{max} to be associated with a complete response by receiver operating characteristic analysis. Multivariate analysis for response found SUV_{max} as a continuous parameter was independently associated with response, as were stage and performance status. A median SUV_{max} value of 15 was used as the cutoff for univariate survival analysis and was associated with both disease-specific survival and overall survival. Patients with SUV_{max} less than 15 had a 2-year survival of 60%, whereas those with SUV_{max} values greater than or equal to 15 had a 2-year survival of only 27%. The small size of this study and heterogeneity of stage again limit the conclusions that can be drawn, but this study is interesting as it specifically demonstrates that SUV is associated with prognosis in a population solely treated by a radiation-based approach, implying that its prognostic value persists despite treatment modality used.

The largest patient series is from Cerfolio et al. and consists of 315 patients at the University of Alabama who presented between January 2001 and June 2004 with either biopsy-proven NSCLC or an indeterminate pulmonary nodule. This study included patients with stage I to IV disease and with a median follow-up of 26 months. FDG-PET was performed on one of two scanners (one dedicated PET and one PET/CT) after a 4-hour fast. Fasting glucose levels were not described. The SUV_{max} within an ROI drawn around the primary tumor was used for this study. A higher SUV_{max} was found in the entire population to be independently correlated to higher disease stage (III or IV), moderate to poor differentiation on pathology, and patients who would subsequently have a less than a complete resection. For the entire population, an SUV_{max} value greater than or equal to 10 was associated with a poor survival (mean survival 1.6 years) compared with those with a SUV_{max} value less than 10 (mean survival 3.2 years). This difference persisted on multivariate analysis, as did survival based on stage and resection status. The survival differences for stage Ib, II, and IIIa as predicted by mean SUV_{max} in each stage also were statistically significant. Specifically, among the 82 patients with stage IB disease, the 4-year disease-free survival was 51% for those with SUV_{max} greater than the median value of 10.3 for this group versus 92% for those with SUV_{max} below the median value. The large size of this study allowed the authors to demonstrate SUV as a correlate of prognosis, even in a narrow early staged group. It also should be noted that a correlation was found even though the patient population was evaluated by two different scanners, with two different methods of attenuation correction. Although this information is provocative, the relatively short duration of follow-up in this patient population limits the conclusions that can be drawn.

Finally, there are data casting some doubt on the utility of FDG-PET as an independent prognostic factor for NSCLC. A series from the University of Washington of 156 patients with potentially resectable NSCLC was analyzed for correlation of SUV and pathologic stage. The FDG-PET technique was controlled to ensure acceptable fasting blood glucose levels and standardized timing of imaging after FDG injection. The authors measured the SUV_{max} and also calculated a
They compared the $SUV_{\text{max}}$ with the well-defined stage for each patient, and although they found a correlation between $SUV_{\text{max}}$ and T status by American Joint Committee on Cancer tumor, node, metastasis staging, this correlation did not persist when using the PVC-$SUV_{\text{max}}$. They specifically note that the $SUV_{\text{max}}$ was significantly different between T1 and T2, T1 and T3, and T1 and T4 tumors, but not between any of the other T stages. The authors did not find any correlation between N status and $SUV_{\text{max}}$ or M status and PVC-$SUV_{\text{max}}$, but they did find a correlation between $SUV_{\text{max}}$ and M status. They also demonstrated a correlation of primary tumor size with T stage and found that the groups of patients with and without metastasis differed significantly with respect to size and $SUV_{\text{max}}$ but not with respect to PVC-$SUV_{\text{max}}$. The authors interpret these data to indicate that $SUV_{\text{max}}$ does not, in fact, correlate with stage when a partial-volume correction is applied. The apparent correlation of the uncorrected $SUV_{\text{max}}$ between T1 and other T stages or with M status occurs because $SUV_{\text{max}}$ is underestimated without this correction. Thus, the SUV is simply a surrogate for size of the primary lesion, which is predictive for stage. This results of study show that there is some controversy regarding the true utility of the primary tumor SUV. In the studies reviewed herein correlating SUV with prognosis, no formal correction for partial-volume averaging was used. The partial-volume corrections used by Veselle et al. are based on measured recovery coefficients of spheres of known size and radioactivity. The validity of this approach can be challenged because primary tumors are rarely perfect spheres. However, these data only fail to demonstrate a correlation of PVC-$SUV_{\text{max}}$ and stage at presentation; in contrast, the other studies reported in this review investigated the correlation of $SUV_{\text{max}}$ with survival and recurrence of the primary tumor. As Veselle et al. note, their data do not rule out a correlation of $SUV_{\text{max}}$ and survival independent of correlation with stage.

**DISCUSSION**

The currently used tumor, node, metastasis staging system, despite its drawbacks, continues to be very useful for predicting outcome in resected stage I-III NSCLC. FDG-PET has the potential to identify patients with more aggressive tumors within each stage, given the correlation between the SUV and tumor growth rate. The challenge is to determine whether FDG-PET can be a useful tool in identifying patients at high risk of recurrence within each stage and, more specifically, in resected stage I and II NSCLC.

The studies presented herein do not, however, provide guidance regarding the use of the SUV measured by FDG-PET to predict outcomes in routine clinical practice. None of the reported studies were conducted prospectively. The variable methods of performing the FDG-PET studies, assigning ROIs, and determining the most appropriate SUV resulted in a variety of SUV cutoffs used to define groups with good and poor prognoses. It is notable that many authors found that SUV was a continuous variable in the separation of groups with good and poor prognoses, which suggests that a strict binary cutoff between groups with good and poor prognoses may not be appropriate. Moreover, all the studies suffer from small sample size and the inclusion of patients with various histologies and different stages of the disease with significantly varied outcomes. To complicate matters further, a significant number of patients had received postoperative therapy. We believe that the role of FDG-PET should be studied prospectively in a homogeneous population of patients, such as those with stage IB and II NSCLC who are scheduled to undergo surgical resection. Determining a method for quantifying FDG uptake in a way that is reproducible across institutions will likely add to the usefulness of this approach.

To validate the hypothesis that SUV can truly discriminate between groups with good and poor prognoses, a prospective trial observing the differences in survival would need to be undertaken. Future studies should take into consideration the influence of primary tumor size (within T stage) and histology on survival and recurrence. The increasing widespread use of adjuvant chemotherapy after resection, although appropriate, is likely to complicate the interpretation of the studies that examine the predictive utility of FDG-PET in NSCLC.

In conclusion, we believe that FDG-PET could be a potentially useful tool in determining prognosis and guiding therapy for patients with early-stage NSCLC. It is likely that a predictive score that includes tumor size, pleural invasion, and SUV may help to identify those at high risk of relapse after resection of early-stage NSCLC. Such a risk stratification approach will be useful for developing novel adjuvant therapies for the high-risk group. Conversely, those who have a very low risk of recurrence may be spared the side effects of adjuvant chemotherapy.

**REFERENCES**


