Spatial Heterogeneity in Sarcoma ^{18}F-FDG Uptake as a Predictor of Patient Outcome

Janet F. Eary[1], Finbarr O'Sullivan[2], Janet O'Sullivan[2], and Ernest U. Conrad[1]

[1]Division of Nuclear Medicine, Departments of Radiology and Orthopedic Surgery, University of Washington, Seattle, Washington; and [2]Department of Statistics, University College Cork, Cork, Ireland

^{18}F-FDG PET images of tumors often display highly heterogeneous spatial distribution of ^{18}F-FDG–positive pixels. We proposed that this heterogeneity in ^{18}F-FDG spatial distribution can be used to predict tumor biologic aggressiveness. This study presents data to support the hypothesis that a new heterogeneity-analysis algorithm applied to ^{18}F-FDG PET images of tumors in patients is predictive of patient outcome. **Methods:** ^{18}F-FDG PET images from 238 patients with sarcoma were analyzed using a new algorithm for heterogeneity analysis in tumor ^{18}F-FDG spatial distribution. Patient characteristics, tumor histology, and patient outcome were compared with image analysis results using univariate and multivariate analysis. Cox proportional hazards models were used to further analyze the significance of the data associations. **Results:** Statistical analyses show that heterogeneity analysis is a strong independent predictor of patient outcome. **Conclusion:** The new ^{18}F-FDG PET tumor image heterogeneity analysis method is validated for the ability to predict patient outcome in a clinical population of patients with sarcoma. This method can be extended to other PET image datasets in which heterogeneity in tissue uptake of a radiotracer may predict patient outcome.

Key Words: ^{18}F-FDG; PET; sarcoma; image analysis

One of the common features of malignant tumors is biologic heterogeneity. This aspect of malignancy has long been known and has classically been described with the histologic features of cellular proliferation; necrosis; noncellular accumulations, such as matrix material and fibrous tissue; and physiologic characteristics, such as differences in blood flow, cellular metabolism, oxygenation, and expression of specific receptors. Tissue heterogeneity in malignant neoplasms has become increasingly apparent in recent years, as more refined molecular and genetic probes designed to assay tumors for common features show diversity in tissues of origin, clonality, changes in expression and regulation in genomic and proteomic pathways, and biomarker expression

(1–7). All of these tumor characteristics have implications for tumor biologic behavior in the patient, which manifest as prognosis and treatment-response variability. A common goal of clinical tumor diagnosis and staging practices is to predict the biologic behavior for local invasiveness and spread to distant sites to plan treatment and response-evaluation schedules.

In addition to its use as a tumor diagnosis and staging modality, PET imaging with ^{18}F-FDG is under intense investigation for use in predicting tumor behavior (8). In the case of sarcomas, the level of tumor ^{18}F-FDG uptake is reflective of tumor grade or aggressiveness determined by standard histopathology (9–12). However, there is more information in the spatial and quantitative data generated in the standard clinical ^{18}F-FDG PET image. This knowledge led us to ask what additional biologic information is contained in each tumor image. Because of the inherent high spatial resolution in PET images, the whole tumor image has an exquisite capability to reflect tumor spatial heterogeneity in ^{18}F-FDG uptake or tumor metabolism. We hypothesized that heterogeneity in tumor metabolism recorded by ^{18}F-FDG uptake is reflective of tumor biologic heterogeneity and could be used as a parameter to predict patient outcome. In this study, we analyzed our patient image dataset with a new image-analysis statistical tool designed to report tumor heterogeneity in ^{18}F-FDG uptake in the entire tumor image volume. This technique assesses the degree to which the ^{18}F-FDG uptake has an ellipsoidal structure whose intensity is greatest at the center and progressively diminishes toward the boundary. The data from this analysis were compared with patient outcome to determine the ability of the method to independently predict patient outcome on the basis of the tumor heterogeneity in ^{18}F-FDG metabolism at presentation. The sarcoma study population was used to evaluate the ability of this analysis to predict patient outcome.

MATERIALS AND METHODS

A total of 238 patients with sarcoma diagnosed by biopsy was imaged with ^{18}F-FDG PET before either neoadjuvant chemotherapy or surgical resection (Table 1); these patients were seen in a sarcoma clinic. The patient datasets in this study are from groups whose ^{18}F-FDG PET data were reported elsewhere in clinical research reports (10,11,13). Consequently, the patients' clinical

TABLE 1
Patient Characteristics for Variables Considered in This Analysis (n = 234)

Variable	Level	Frequency	Percentage
Sex	M	124	54
	F	106	46
Metastatic disease at diagnosis	Yes	41	18
	No	189	82
Prestudy tumor grade*	Low	36	16
	Intermediate	72	32
	High	118	52
Tumor subtype	Bone	63	27
	Cartilage	18	8
	Soft tissue	149	65

*Tumor grade not available for 4 patients.

Sixteen patients were removed for lack of complete data availability.

course and other pertinent associated clinical data were recorded in our sarcoma patient database.

18F-FDG PET

All patients underwent standard 18F-FDG PET on an Advance scanner (GE Healthcare), according to a previously described protocol (13). Briefly, patients received 18F-FDG (259–370 MBq [7–10 mCi]) intravenously 45–60 min before emission and attenuation-scan acquisition of the tumor fields of view. Images were reconstructed using standard filtered backprojection algorithms. Image regions of interest of the entire tumor volume were hand-drawn around the tumor area of increased 18F-FDG uptake on axial planes for subsequent uptake and heterogeneity analysis. The tumor 18F-FDG uptake was reported as the maximum standardized uptake value (SUVmax) for the entire tumor volume. This value was derived from application of the standard clinical SUV analysis program to the PET image. Only primary tumor masses were analyzed.

Anatomic Imaging Analysis

CT and MR images for 77 patients were available for review of the presence or absence of suspected tumor necrosis. For the remainder of the patient group, these images were not in the archives or had been obtained at referring institutions and were not available for review. The data unavailable for review were obtained by image review or by report.

Heterogeneity Analysis

We used our image analysis algorithm described previously (14) for quantitation of tumor 18F-FDG spatial heterogeneity. Briefly, this method assesses the extent to which the spatial distribution of 18F-FDG uptake within the tumor follows a certain idealized pattern. The idealized pattern has solid ellipsoidal or football-shaped contours in which the uptake progressively declines from a central area of increased activity within the volume (Fig. 1).

Mathematically, the idealized pattern is defined by an equation representing the 18F-FDG use, U(x), as a function of a 3-dimensional spatial coordinate x:
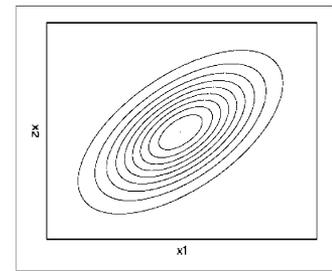

FIGURE 1. Three-dimensional ellipsoidal model for homogeneous tissues.

$$U(x) = g[(x - \mu)' \textstyle\sum^{-1} (x - \mu)], \qquad \text{Eq. 1}$$

where μ represents the location of the greatest use, $\sum$ is a positive-definite–shape matrix, and g is a nonnegative monotone decreasing–level function defined over the positive real line. In heterogeneity analysis, the values of μ, $\sum$, and g are adapted for each tumor dataset using a residual sum-of-squares criterion. The heterogeneity variable (HET) is defined as the percentage of variability in the voxel-level 18F-FDG uptake that is not explained by the optimized ellipsoidal use pattern. Thus, HET is 100 minus the percentage of R^2 associated with the optimized fit of the elliptically contoured structure specified by Equation 1. As HET is scale-invariant, it complements the scale-dependent SUVmax variable.

It is apparent that the HET is focused on discriminating a biologic structure that would be missed by consideration (e.g., relative dispersion) of the marginal distribution or histograms of uptake. Figure 2C shows a histogram from 2 hypothetical distributions with markedly different pixel spatial distributions and HET values close to the extremes but the same histogram of pixel distribution intensity. This figure demonstrates that quantitation based on the simple histogram as shown in this example cannot be used to characterize the spatial pattern within the volume. In our study, the tumor volume 18F-FDG datasets consisting of the 3-dimensional spatial coordinates and 18F-FDG uptake values for each voxel in the tumor volume were exported to a stand-alone workstation (PC) for heterogeneity analysis.

Statistical Analysis

Death and disease reappearance (local recurrence, metastasis, or death) were considered. Time to either an event or the last follow-up was measured in months from the initial PET study. These times to death (denoted "survival") or times to disease reappearance (denoted "disease-free survival") were evaluated using standard survival analysis techniques (15). Prognostic factors that were considered included patient age, sex, tumor subtype, tumor grade, tumor SUVmax, and the 18F-FDG tumor uptake HET. HET and SUVmax were found to have skewed distributions, and a square-root transformation was applied to both variables. Unless otherwise noted, all references to these variables are to the transformed versions. Univariate analyses consisted of statistical summaries of variables and univariate Cox proportional hazards regression for each variable to assess its importance in relation to survival and disease-free survival (15). For HET and SUVmax, log-rank tests were also used to compare survival curves for those patients above and below the median values. In multivariate analyses, the relationship between the time to outcome and full set of measured prognostic factors was evaluated using the standard multivariate Cox proportional hazards analysis. This
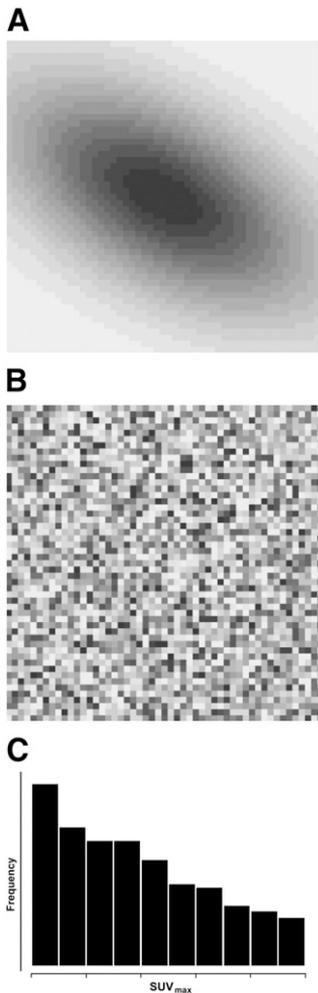
FIGURE 2. (A) Spatially coherent pixel-distribution model. (B) Spatially incoherent pixel-distribution model. (C) Histogram result for models shown in A and B.

8.2 (Table 2). These values are similar to previous results (*14*). Table 2 also shows the HET values for the study group. The median HET value was 5.2, with a mean of 7.1.

Univariate Analysis

Univariate statistics (frequency tables) for all variables considered in this analysis are shown in Table 3. Kaplan–Meier survival curves, for both SUVmax and heterogeneity, are shown for disease-free and overall survival in Figures 4 and 5. Log-rank tests were used to compare the survival curves associated with patients above and below the median values for both SUVmax and heterogeneity (*15*). Disease-free survival was calculated for only those patients who entered the study free of metastatic disease (193). The *P* values for the log-rank tests are shown with each plot. Univariate Cox proportional hazards regression was used to assess the importance of SUVmax, tumor heterogeneity, and tumor grade. SUVmax and heterogeneity were significant predictors of overall patient survival (Table 3). Tumor grade and size were also significant for disease-free survival.

Multivariate Analysis

The relationship between the time to death (or recurrence or relapse) and the full set of measured prognostic factors was evaluated using the standard multivariate Cox proportional hazards regression analysis. This analysis permitted an examination of the influence of the PET measures and controlled for the impact of other variables. The data were

analysis permits an examination of the influence of the PET measures while controlling for the impact of other variables. A full model (including all factors) was fit for each outcome.

RESULTS

A total of 238 patients was included in the analysis. Four patients with extreme SUVmax values (more than 5 SDs above the median) were excluded. Median follow-up was 31.3 mo. For 41 patients who entered the study with evidence of metastatic disease, median follow-up was 18.3 mo. For patients who entered the study with only local disease (193), median follow-up was 34.4 mo. During the study, 125 patients in the analysis group of 234 patients experienced local recurrence events, and metastatic disease appeared in 56 patients. The median overall survival for the entire group was 52.8 mo, and the median progression-free survival was 37 mo.

Heterogeneity analysis was completed for 234 patients. As expected, the image appearance for the range of HET values was highly variable. Figure 3 shows an image of a patient with a high-grade tumor. For the study group, the median tumor ^{18}F-FDG SUVmax was 6.2, with a mean of
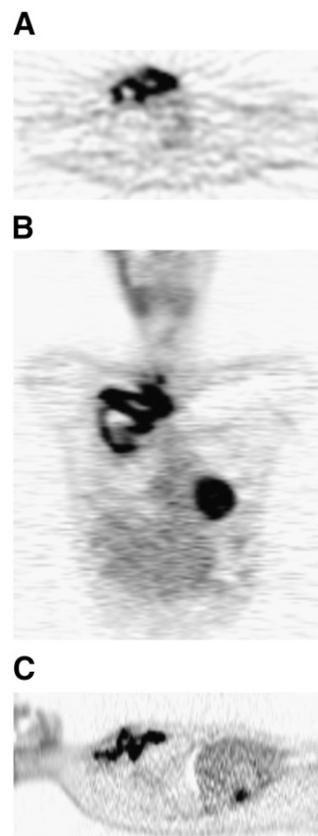

FIGURE 3. ^{18}F-FDG PET images of heterogeneous uptake in patient with sarcoma. Transverse (A), coronal (B), and spatial (C) images showing large, poorly differentiated sarcoma in right upper chest.

TABLE 2
Patient Characteristics (Continuous Variables, *n* = 230)

Variable	Mean	SD	Minimum	25th percentile	Median	75th percentile	Maximum
Age (y)	46.6	16.8	17.7	33.8	45.3	59.1	86.3
Size (voxels)	5148	5394	156	1644	3180	7017	34360
Square-root size	64.1	32.3	12.5	40.6	56.4	83.8	185.4
Tumor SUVmax	8.2	6.0	1.5	4.0	6.2	11.3	40.3
Square-root SUVmax	2.7	.94	1.2	2.0	2.5	3.4	6.4
Tumor heterogeneity	7.1	6.3	.43	2.3	5.2	10.1	44.6
Square-root heterogeneity	2.4	1.1	.65	1.5	2.3	3.2	6.7

stratified into 2 groups for the analysis: those patients who entered the study with primary tumors and those who had metastases at the time of presentation. Stratification allowed separate baseline hazards for the groups to be applied. Age, sex, and tumor subtype, despite showing little prognostic potential in the univariate analyses, were included in the multivariate model to assess any potential interactions or confounding factors.

Age, SUVmax, and heterogeneity were entered as continuous variables. Tumor subtype, tumor grade, and sex were entered as nominal variables. Frequencies of occurrence for prognostic variables considered are shown in Table 1. In the analysis for overall survival, SUVmax, age, sex, and tumor type had no significant contribution toward patient prognosis (Table 4). However, tumor grade and heterogeneity were significant predictors of survival. An increase of 6.45% (i.e., 1 SD) in heterogeneity is associated with a 65% increase in the risk of death; risk of death was more than 4 times greater for patients with high-grade tumors than for those with low-grade ones.

The multivariate analysis results for disease-free survival are shown in Table 5. Tumor grade is again a significant

TABLE 3
Univariate Analyses Results for Patient Survival

	Hazard ratio (*P*)	
Variable	Overall survival	Disease-free survival
Age	1.14 (0.18)	1.04 (0.69)
Sex	0.84 (0.38)	0.87 (0.49)
Tumor grade		
High vs. low	1.90 (**0.005**)	1.79 (**<0.01**)
Intermediate vs. low	0.29 (**0.02**)	0.39 (**0.03**)
Tumor subtype		
Cartilage vs. bone	2.44 (0.09)	1.93 (0.14)
Soft tissue vs. bone	1.90 (0.22)	1.44 (0.39)
Tumor size (voxels)	1.17 (0.06)	1.26 (**<0.01**)
Tumor SUVmax	1.46 (**<0.001**)	1.31 (**<0.01**)
Tumor heterogeneity	1.81 (**<0.001**)	1.83 (**<0.001**)

Columns 2 and 3 are hazard ratios for variable when it is sole variable in model. For continuous variables, hazard ratio should be interpreted as increase in survival risk associated with increase of 1 SD of covariate. Significant variables are in bold type.

predictor, whereas all other variables are not. This dataset is restricted to those patients who had no metastatic disease at the start of the study. Increasing SUVmax does not significantly increase the risk of recurrence. This association of tumor heterogeneity risk and tumor grade is a powerful predictor of patient outcome. An increase of 6.4% (i.e., 1 SD) in heterogeneity is associated with a 60% increase in the risk of recurrence; high-grade tumors are 3 times more likely to recur than are low-grade ones.

In the subset of 77 patients analyzed for the presence of necrosis on anatomic imaging, 13% were described as positive. The relationship of the association of anatomic necrosis with heterogeneity was weak, with a *P* value of 0.74. The multivariate analysis of these data indicates a potential prognostic role for anatomic necrosis, with a *P* value of 0.09, and in combination with the HET value (*P* = 0.02) and tumor grade (*P* = 0.07).

DISCUSSION

In this study, we demonstrated that heterogeneity image analysis is a strong independent predictor of survival in patients with sarcoma. In this mixed group of soft-tissue and bony sarcomas, the SUVmax is somewhat less predictive of tumor-grade survival. This is likely due to the nonsimilarities in behaviors between these 2 major histologic groups. Our previous studies in more homogeneous sarcoma histologic tumor types show a more significant relationship between overall survival and tumor ^{18}F-FDG SUVmax (*9,11,13*). Tumor histologic grading schemes are designed to predict patient outcome via an assessment of the aggressiveness of the tumor cytologic features. The concept of tumor histologic grade is used to evaluate tumor metastatic potential (*16,17*). A sarcoma with a highly heterogeneous ^{18}F-FDG uptake distribution suggests multiple cell populations of differing growth rates, vascularity, necrosis, ground substance, and fluid collections. Histologic examination of these tumors has long shown us that these tumor features also imply aggressive behavior. In fact, the histologic hallmark of a high-grade sarcoma is the presence of necrosis; however, in this analysis there is only a weak association of HET with the presence of anatomic imaging findings in a subset of the patients in this study. This feature is one of several tumor characteristics that is likely quanti-
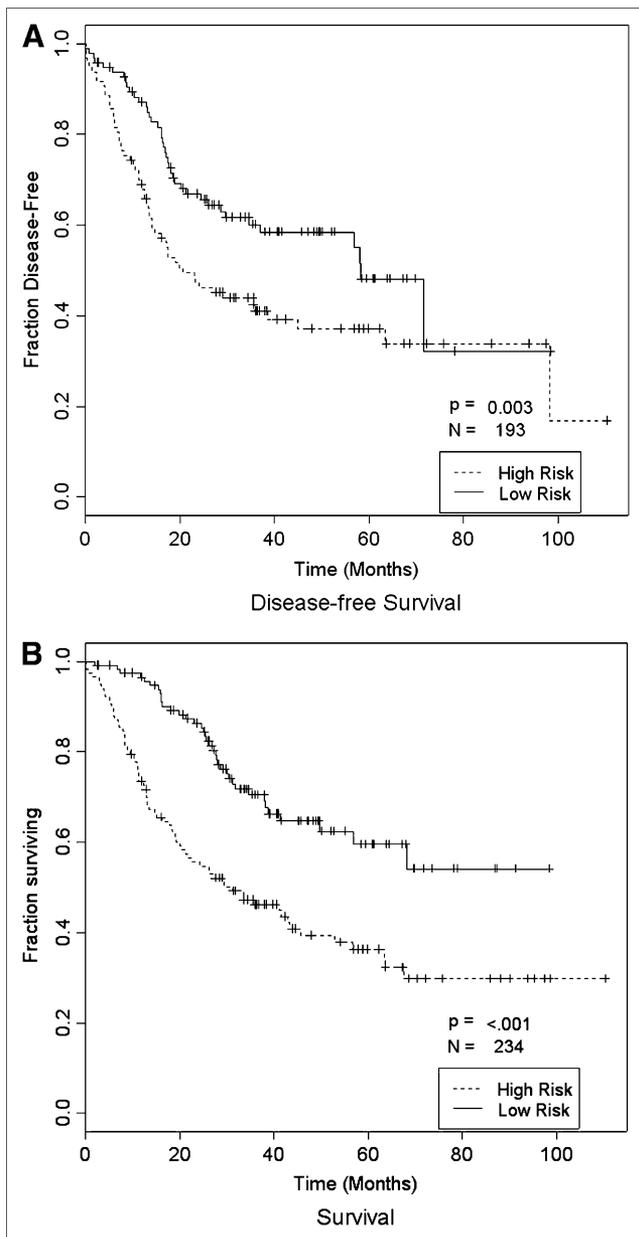
FIGURE 4. Kaplan–Meier survival curves, for SUVmax, are shown for disease-free (A) and overall (B) survival. Log-rank tests were used to compare survival curves associated with patients above and below median values. *P* values for log-rank tests are shown with each plot.

FIGURE 5. Kaplan–Meier plots for disease-free (A) and overall survival (B) for heterogeneity measure for sarcomas.

tated with HET analysis but not the most influential factor contributing to this image-computed value (*16*).

The ^{18}F-FDG SUVmax reflects tumor biologic aggressiveness in sarcoma, but heterogeneity analysis adds a new dimension to this understanding (*14*). In this study, the heterogeneity analysis model for ^{18}F-FDG spatial distribution in tumors we developed generates a value that independently correlates with patient outcome. The model used is simple and is a comparison of the tumor volume pixel-uptake distribution with a smooth elliptic solid object with homogeneous density. This object was chosen to represent
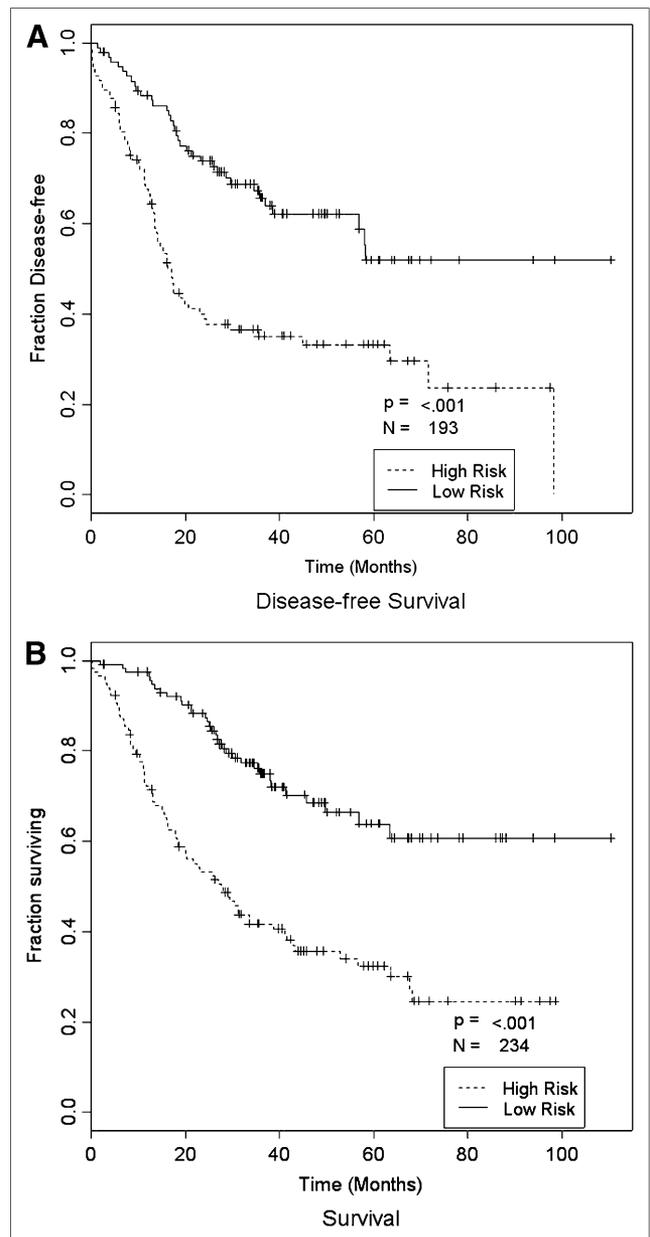
homogeneous tumor uptake. A sarcoma with this type of ideal uptake could presumably be either high or low grade. However, as histologic grade in tumor increases, so can heterogeneity in the tumor cellular makeup and viability. On the basis of these characteristics, it is probable that comparison of actual tumor volumes with the model can identify high-grade tumors of any type. A high level of heterogeneity in tumor metabolism immediately implies a high metastatic potential and a poor patient outcome.

However, heterogeneity models that can be even more accurate in identifying tumor biologic behavior characteristics may be possible, because this first version we report is based on only the heterogeneity in ^{18}F-FDG spatial distri-

TABLE 4
Multivariate Analyses for Patient Survival ($n = 225$, Events $= 102$)

Variable	Coefficient	Hazard ratio (confidence interval)	P
Age	0.004	1.07	0.61
Sex	−0.24	0.86 (0.68, 1.08)	0.28
Grade			
High vs. low	1.60	4.93 (1.66, 14.64)	**0.004**
Intermediate vs. low	1.03	2.8 (0.93, 8.40)	0.07
Subtype			
Cartilage vs. bone	0.06	1.06 (0.33, 3.46)	0.92
Soft tissue vs. bone	0.25	1.28 (0.44, 3.76)	0.65
Size (sqrt)	−0.005	0.84 (0.66, 1.09)	0.40
SUVmax (sqrt)	0.10	1.10 (0.88, 1.37)	0.40
Heterogeneity (sqrt)	0.49	1.72 (1.30, 2.28)	**<0.001**

Continuous variable, hazard ratio, and confidence intervals are interpreted as increase in survival risk associated with increase of 1 SD of covariate. Significant variables are in bold type. sqrt = square root.

bution characteristics. It may be that a model tumor that includes surface irregularities that closely reflect the extent to which the boundary of tumor tissue is infiltrated will have even stronger predictive abilities for tumor behavior assessment (*18*).

It is expected that the new HET variable has clinical validity for predicting patient outcome. The HET takes the idea of ^{18}F-FDG image analysis one step further than calculation of the SUVmax and analyzes the tumor area independent of region-of-interest generation. This application of HET analysis to a large dataset of patient images and their outcomes provides clinical validation for the technique. It also demonstrates proof of principle that an objective imaging analysis can provide important data for increased biologic insights into a pathologic process beyond simple observations and generation of tumor SUV.

Objective, statistically based imaging analysis algorithms may provide new answers to our continued questions: "What's in this image?" and "What insights can this quantitative 3-dimensional image of a pathologic process provide beyond simple identification of areas of abnormal tissue metabolism?" A great deal of tissue biologic information in the standard clinical ^{18}F-FDG PET image is currently underutilized. HET image analysis validated with clinical data can be rapidly translated to other disease processes in which deviations from normal distribution patterns of receptor or tissue processes are the basis for disease. In this initial clinical validation of HET analysis, we open statistically based image analysis to direct clinical application for understanding the disease process in a patient, using biologically specific and relevant radiopharmaceuticals.

TABLE 5
Multivariate Analyses for Disease-Free Survival
($n = 188$, Events $= 95$)

Variable	Coefficient	Hazard ratio (confidence interval)	P
Age	0.01	1.13 (0.89, 1.44)	0.30
Sex	−0.23	0.79 (0.52, 1.21)	0.29
Grade			
High vs. low	0.59	1.80 (0.80, 4.05)	0.15
Intermediate vs. low	0.26	1.3 (0.59, 2.86)	0.52
Subtype			
Cartilage vs. bone	0.16	1.18 (0.42, 3.32)	0.75
Soft tissue vs. bone	−0.02	0.98 (0.40, 2.41)	0.97
Size (sqrt)	−0.00	0.93 (0.71, 1.21)	0.6
SUVmax (sqrt)	−0.05	0.96 (0.74, 1.23)	0.72
Heterogeneity (sqrt)	0.51	1.75 (1.28, 2.40)	**<0.001**

As in Table 4, continuous variable, hazard ratio, and confidence intervals are interpreted as increase in survival risk associated with increase of 1 SD of covariate. Significant variables are in bold type. sqrt = square root.

CONCLUSION

The new ^{18}F-FDG PET tumor image heterogeneity analysis method is validated for its ability to predict patient outcome in a clinical population of patients with sarcoma. This method can be extended to other PET image datasets in which heterogeneity in tissue uptake of a radiotracer may have patient outcome predictability.

REFERENCES

1. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000;100:57–70.
2. Bissell MJ, Radisky D. Putting tumours in context. *Nat Rev Cancer.* 2001;1: 46–54.
3. Thiagalingam S. A cascade of modules of a network defines cancer progression. *Cancer Res.* 2006;66:7379–7385.
4. Loeb LA, Loeb KR, Anderson JP. Multiple mutations and cancer. *Proc Natl Acad Sci USA.* 2003;100:776–778.

5. Auerbach R, Auerbach W. Regional differences in the growth of normal and neoplastic cells. *Science*. 1982;215:127–134.

6. Kinzler KW, Vogelstein B. Cancer-susceptibility genes: gatekeepers and caretakers. *Nature*. 1997;386:761–763.

7. Cahill DP, Kinzler KW, Vogelstein B, Lengauer C. Genetic instability and Darwinian selection in tumours. *Trends Cell Biol*. 1999;9:M57–M60.

8. Shankar LK, Hoffman JM, Bacharach S, et al. National Cancer Institute. Consensus recommendations for the use of ^{18}F-FDG PET as an indicator of therapeutic response in patients in National Cancer Institute Trials. *J Nucl Med*. 2006;47:1059–1066.

9. Eary JF, Conrad EU. Positron emission tomography in grading soft tissue sarcomas. *Semin Musculoskelet Radiol*. 1999;3:135–138.

10. Eary JF, Mankoff DA. Tumor metabolic rates in sarcoma using FDG PET. *J Nucl Med*. 1998;39:250–254.

11. Folpe AL, Lyles RH, Sprouse JT, Conrad EU III, Eary JF. (F-18) fluorodeoxy-glucose positron emission tomography as a predictor of pathologic grade and other prognostic variables in bone and soft tissue sarcoma. *Clin Cancer Res*. 2000;6:1279–1287.

12. Schwarzbach MH, Dimitrakopoulou-Strauss A, Willeke F, et al. Clinical value of [18-F] fluorodeoxyglucose positron emission tomography imaging in soft tissue sarcomas. *Ann Surg*. 2000;231:380–386.

13. Eary JF, O'Sullivan F, Powitan Y, et al. Sarcoma tumor FDG uptake measured by PET and patient outcome: a retrospective analysis. *Eur J Nucl Med Mol Imaging*. 2002;29:1149–1154.

14. O'Sullivan F, Roy S, Eary J. A statistical measure of tissue heterogeneity with application to 3D PET sarcoma data. *Biostatistics*. 2003;4:433–448.

15. Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Time Failure Data*. New York, NY: Wiley; 1980.

16. Guillou L, Coindre JM, Bonichon F, et al. Comparative study of the National Cancer Institute and French Federation of Cancer Centers sarcoma group grading systems in a population of 410 adult patients with soft tissue sarcoma. *J Clin Oncol*. 1997;15:350–362.

17. Unni KK. Osteosarcoma of bone. *J Orthop Sci*. 1998;3:287–294.

18. O'Sullivan F, Roy S, O'Sullivan J, Vernon C, Eary J. Incorporation of tumor shape into an assessment of spatial heterogeneity for human sarcomas imaged with FDG-PET. *Biostatistics*. 2005;6:293–301.