Glucose-Normalized Standardized Uptake Value from ^{18}F-FDG PET in Classifying Lymphomas

Ching-yee Oliver Wong, MD, PhD[1]; Joseph Thie, PhD[2]; Kelly J. Parling-Lynch, MD[3]; Dana Zakalik, MD[3]; Jeffrey H. Margolis, MD[3]; Marianne Gaskill, CNMT[1]; Jack Hill, PhD[4]; Feng Qing, MD, PhD[1]; Darlene Fink-Bennett, MD[1]; and Conrad Nagle, MD[1]

[1]Nuclear Medicine, William Beaumont Hospital, Royal Oak, Michigan; [2]Graduate School of Medicine, University of Tennessee Medical Center, Knoxville, Tennessee; [3]Medical Oncology, William Beaumont Hospital, Royal Oak, Michigan; and [4]Medical Information Service, William Beaumont Hospital, Royal Oak, Michigan

Our objective was to derive the best glucose sensitivity factor (g-value) and the most discriminating standardized uptake value (SUV) normalized to glucose for classifying indolent and aggressive lymphomas. **Methods:** The maximum SUV obtained from ^{18}F-FDG PET over the area of biopsy in 102 patients was normalized by serum glucose ([Glc]) to a standard of 100 mg/dL. Discriminant analysis was performed by using each SUV_{100} (SUV $\times$ {100/[Glc]}^{g}, calculated using various g-values ranging from -3.0 to 0, one at a time) as a variable against the lymphoma grades, and plotting the percentage of correct classifications against g (g-plot) to search for the best g-value in normalizing SUV_{100} for classifying grades. To address the influence of the extreme glucose conditions, we repeated the same analyses in 12 patients with [Glc] $\leq$ 70 mg/dL or [Glc] $\geq$ 110 mg/dL. **Results:** SUV_{100} correctly classified lymphoma grades ranging from 62% to 73% ($P < 0.0005$), depending on the g-value, with a maximum at a g-value of -0.5. For the subgroup with extreme glucose values, the g-plot also revealed higher and more optimal discrimination at a g-value of -0.5 (92%) than at a g-value of 0 (83%) ($P = 0.03$). The discrimination deteriorated at $g < -1$ in both analyses. The box plot for all cases using a g-value of -0.5 showed little overlap in classifying lymphoma grades. For a visually selected threshold SUV_{100} of 7.25, the sensitivity, specificity, and accuracy of identifying aggressive grades were 82%, 79%, and 81%, respectively. **Conclusion:** The results suggest that metabolic discrimination between lymphoma grades using a glucose-normalized SUV from ^{18}F-FDG PET is improved by introducing g-value as an extra degree of freedom.

Key Words: PET; lymphoma; SUV normalization; glucose sensitivity

J Nucl Med 2005; 46:1659–1663

In oncologic evaluation, the use of PET with ^{18}F-FDG as its analog tracer is now widespread because it can take advantage of diverse capabilities among cell types in metabolizing serum glucose. The modeling of glucose metabolic rate (MRglc) with ^{18}F-FDG (*1,2*) follows the solid foundation from ^{14}C-deoxyglucose (*3*):

$$MRglc = \{[Glc]/LC\}\{K_1k_3/(k_2 + k_3)\}, \quad \text{Eq. 1}$$

where the rate constants K_1 and k_2 are the forward and reverse transport across the capillary/cellular membrane, k_3 is the phosphorylation of ^{18}F-FDG to ^{18}F-FDG-6-P in cells, and LC is a lumped constant that accounts for the transport and phosphorylation difference between ^{18}F-FDG and glucose, assuming a small dephosphorylation rate constant (k_4) for ^{18}F-FDG-6-P. It has been shown that as ^{18}F-FDG competes with serum glucose, the standardized uptake value (SUV), but not the MRglc, in certain tumors is influenced by the serum glucose concentration ([Glc]) (*4*). Assuming negligible free ^{18}F-FDG at the time of imaging, it has been theorized that SUV is related to MRglc (*5*):

$$MRglc = G \times [Glc] \times SUV \quad \text{Eq. 2A}$$

or

$$MRglc/G = [Glc] \times SUV, \quad \text{Eq. 2B}$$

where G is a product of LC and a subject-independent proportional constant relating the term of dose per body weight to the time integral of total plasma ^{18}F-FDG activity (*5*).

Thus, by taking the differentials on both sides of Equation 2B with respect to [Glc] and treating MRglc/G as a single term, one can introduce a proposed glucose sensitivity factor, or g-value (*6*):

$$d\,(SUV)/d\,[Glc] = g \times \{SUV/[Glc]\}, \quad \text{Eq. 3}$$

where in terms of MRglc/G, $g = \{d\,[(MRglc/G)]/d\,[Glc]\} \times \{[Glc]/(MRglc/G)\} - 1$. For cases in which G may be assumed to be independent of [Glc] (*5*), g is simplified to $\{d\,(MRglc)/d\,[Glc]\} \times \{[Glc]/(MRglc)\} - 1$. In certain special situations, assuming that changes in MRglc are negligible with respect to [Glc] (*4*) or that MRglc/G is

Received Apr. 12, 2005; revision accepted Jun. 30, 2005.

For correspondence or reprints contact: C. Oliver Wong, MD, PhD, Positron Diagnostic Center and Medical Cyclotron, Department of Nuclear Medicine, William Beaumont Hospital, 3601 W. Thirteen Mile Rd., Royal Oak, MI 48073-6769.

E-mail: owong@beaumont.edu

constant, then the change in SUV with respect to [Glc] is given by:

$$d\,(SUV)/d\,[Glc] = -SUV/[Glc]. \quad \text{Eq. 4}$$

This explains the inverse relationship between [Glc] and ^{18}F-FDG uptake in many experimental tumors (*7*). Despite these attempts to theorize the meaning of the proposed g-value with few basic assumptions from the literature (*5*), Equation 3 can also be treated as just an empiric formula and, by being written in the logarithmic form, becomes:

$$g = d\,\{\ln\,(SUV)\}/d\,\{\ln\,[Glc]\} = \{d(SUV)/d[Glc]\} \times \{[Glc]/SUV\}. \quad \text{Eq. 5}$$

Thus, g-value may be regarded as the percentage change in SUV per percentage change in [Glc] for any type of functional dependence of SUV on [Glc], as proposed by a prior preliminary investigation (*6*). When Equation 5 is written in the functional forms, it becomes simple practical ratios:

$$g = \{\ln\,(SUV_1) - \ln\,(SUV_2)\}/\{\ln\,[Glc_1] - \ln\,[Glc_2]\} \quad \text{Eq. 6A}$$

or

$$SUV_1/SUV_2 = \{[Glc_1]/[Glc_2]\}^g, \quad \text{Eq. 6B}$$

where subscripts 1 and 2 denote SUV at 2 different [Glc] conditions.

Although little attention has been given to this newly defined g-value, the theoretic values of g appear to range from −1, as in Equation 4 when the change in MRglc with respect to [Glc] is negligible (*4*), to 0, as in Equation 3 when the percentage change in MRglc/G equals that in [Glc], because some investigators do assume what would be g = −1 or g = 0 (*4,5*). The objective of this study was to derive the best g-value and the most discriminating SUV normalized to glucose for classifying indolent and aggressive lymphomas, using an empiric g-value ranging from −3 to a practical upper limit of 0 when there is no correction for [Glc]. Thus, within this range falls a g-value of −1, which is popularly assumed as a universal value in correcting SUV for serum glucose to a standard [Glc] (*5,8,9*) and is included in this study. Many previous studies sought the best way to normalize SUV. A previous study suggested that glucose correction using a g-value of −1 reduces the variability of SUV calculations (*9*). Instead of choosing discriminant analysis, another prior study chose the receiver operating characteristic curve for optimization (*10*). Those investigators did not consider fractional g-value as a degree of freedom in optimization. The current study introduces an extra dimension, g-value for SUV. Part of this portion of the investigation has been reported in abstract form (*11*).

MATERIALS AND METHODS

Patient Group and PET Imaging

The study included 102 patients with newly diagnosed non-Hodgkin's lymphoma ($n = 60$) or Hodgkin's disease ($n = 42$) but without diabetes or other cancers and who underwent PET within 3 mo of biopsy. Lymphomas were classified using the World Health Organization criteria and graded clinicopathologically as indolent or aggressive, by pathologists who were unaware of the PET results. Scans were obtained on a dedicated whole-body PET scanner (Advance; GE Healthcare) 1 h after injection of ^{18}F-FDG (370 MBq, on average) and after the patients had fasted about 4 h. [Glc] was recorded for each patient just before injection of the tracer. PET images were reconstructed using an iterative reconstruction algorithm with segmented attenuation correction. Using the maximum SUV obtained over the area of biopsy and setting the standard [Glc] equal to 100 mg/dL, we defined a standardized SUV, called SUV_{100}, from Equation 6B:

$$SUV_{100} = SUV \times \{100/[Glc]\}^g. \quad \text{Eq. 7}$$

Data Analysis

To determine the sign of the g-value, we first calculated the g-value using the regression method, ln (SUV) versus ln ([Glc]), which yields a negative value of −0.3 for the entire population, as reported in the literature (*4,6*). Because of the various potential experimental errors, g-values more negative than −1 were considered to be beyond those predicted by theory (Eq. 3). Thus, the discriminant analysis was performed for each independent variable SUV_{100} by separately using various negative g-values to calculate SUV_{100}s ranging from −3.0 to 0 at −3, −2.5, −2, −1.5, −1.125, −1, −0.875, −0.75, −0.625, −0.5, −0.25, and 0, one at a time. Each generated set of SUV_{100}s is then a variable in discriminant analysis to use against the dichotomous value of lymphoma grades (indolent or aggressive types). The cutoff value of SUV_{100} for each g-value used in discriminant analysis was chosen automatically by computer algorithm (SPSS Inc.) to avoid threshold selection bias. The box plots were later examined to determine the discrimination by each derived SUV_{100}. Finally, the percentage of correct classifications using various g-values ranging from −3.0 to 0 was plotted against g-value to search for that which best normalized SUV_{100} for classifying lymphoma grades. To address the influence of the extreme glucose conditions, we repeated the same analyses on 12 patients with [Glc] ≤ 70 mg/dL or [Glc] ≥ 110 mg/dL.

RESULTS

Patients with aggressive lymphoma did not statistically differ from patients with indolent lymphoma in age (51 ± 19 y vs. 54 ± 19 y), height (170 ± 10 cm vs. 165 ± 10 cm), weight (77 ± 16 kg vs. 73 ± 19 kg), or [Glc] (90 ± 13 mg/dL vs. 89 ± 16 mg/dL). SUV was, however, significantly higher in aggressive lymphomas than in indolent lymphomas (15.7 ± 10.2 vs. 5.9 ± 2.9, $P < 0.0005$). SUV_{100} correctly classified lymphoma grades ranging from 62% to 73% ($P < 0.0005$), depending on the g-value, with a maximum at a g-value of −0.5 (Fig. 1A). The plot for percentage correct classification against g-value or the g-plot for all cases rose monotonically to a peak as g-value increased from −3.0 to −0.5 and decreased when g-value was more than −0.5. For the subgroup with extreme glu-

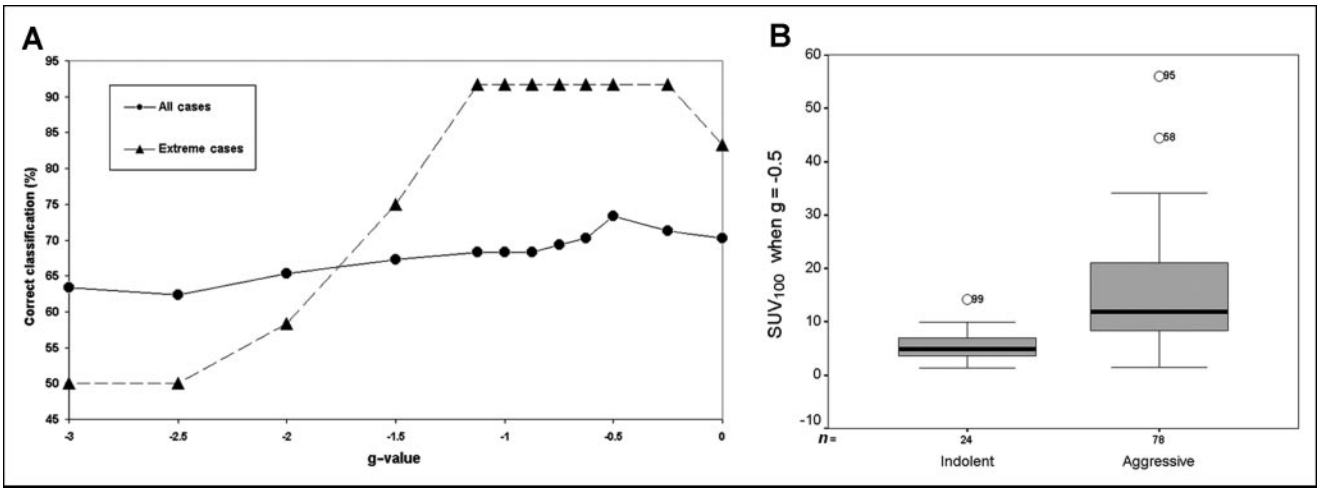

FIGURE 1. (A) g-plot: percentage of correctly classified pathologically proven indolent and aggressive lymphoma grades against g-values, including the entire group and the subgroup with extreme [Glc]. (B) Box plot for entire group ($n = 120$) for g-value of -0.5. Outliers are identified by ○ with case numbers. Indolent lymphomas are widely separated from aggressive lymphomas using SUV_{100} at g-value of -0.5, which may be considered a boundary zone.

cose values, the g-plot also revealed higher and more optimal discrimination when the g-value was -0.5 (92%) than when it was 0 (83%) ($P = 0.03$). The discrimination deteriorated at $g < -1$ in both analyses. The box plot for all cases using a g-value of -0.5 showed little overlap in classifying lymphoma grades (Fig. 1B). If a cutoff SUV_{100} of 7.25 was manually chosen near the separation line in the box plot, the accuracy of identifying the aggressive grade was 81%, with a sensitivity of 82% and a specificity of 79% (Eq. 9). The SUV_{100} was significantly higher in aggressive lymphomas than in indolent lymphomas (14.8 ± 9.6 vs. 5.6 ± 2.9, $P < 0.0005$). The means for the 2 groups were lower than the uncorrected SUVs because for the entire group the global mean [Glc] was slightly lower (90 ± 13 mg/dL) than the standard [Glc] (100 mg/dL). The regression of ln (SUV) versus ln [Glc] yielded an estimated g-value of -0.3, which was close to the empiric results reported here. The plot of SUV at a g-value of 0 versus the [Glc] of all 102 patients showed 2 populations: one representing indolent lymphomas and one representing aggressive lymphomas (Fig. 2A). When there was no glucose correction, the area between the mean SUV lines of these 2 populations represented the boundary zone between them. This boundary zone was refined by plotting ln (SUV) against ln [Glc] or the logarithmic form of SUV against $[Glc]^g$ (Fig. 2B). The areas between the mean regression lines or their 95% confidence intervals might be considered examples of the boundary zones. Straight lines (in the logarithmic form) with various slopes (g) derived from Equation 7 could also serve as the discriminating lines between these 2 populations:

$$\ln (SUV) = \ln (SUV_{100}/100^g) + g \times \ln [Glc]. \quad \text{Eq. 8}$$

Parallel lines with different cutoff SUV_{100}s but the same slope at a g-value of -0.5, obtained from discriminant analysis crossing the boundary zone, yielded the various threshold values for achieving clinically appropriate and useful sensitivity, specificity, and accuracy, as in the following special formula from Equation 8 with the data from the current study:

$$\ln (SUV) = \ln (SUV_{100\,cutoff}/100^{-0.5}) - 0.5 \times \ln [Glc] =$$
$$-0.5 \times \ln [Glc] + \{\ln (SUV_{100\,cutoff}) + \ln (10)\}. \quad \text{Eq. 9}$$

DISCUSSION

PET is now a well-established diagnostic tool, with visual interpretations of images dominating the diagnostic process. However, quantitative techniques, including the so-called semiquantitative SUV analyses, supplement PET. These SUV analyses take advantage of the separation in mean SUVs between 2 populations—in this study, between patients with indolent lymphomas and patients with aggressive lymphomas (Figs. 1B and 2). The impartiality of quantitative classification makes it a valuable adjunct to visual analyses.

An underlying difficulty in this classification process, caused by dispersion about the means, is the presence of overlap (Figs. 1B and 2): SUVs are lower for some aggressive types than for some indolent types, leading to incorrect discrimination. Primarily responsible for these patient-to-patient variations in uptake within a disease class are cellular biologic factors, which the rate constants of Equation 1 reflect. Also partially responsible is the random distribution of [Glc] in patients. This work reduces that variability by optimally correcting all SUVs to a standard glucose level of 100 mg/dL. It is this optimal correction, with the best choice of g-value, that is proposed, in contrast to the tradition of using SUV with, usually, a g-value of 0 (no correction) or, occasionally, of -1 (a theoretic correction, which occurs only when there is no change in MRglc/G with respective to [Glc], as suggested by Equation 3).

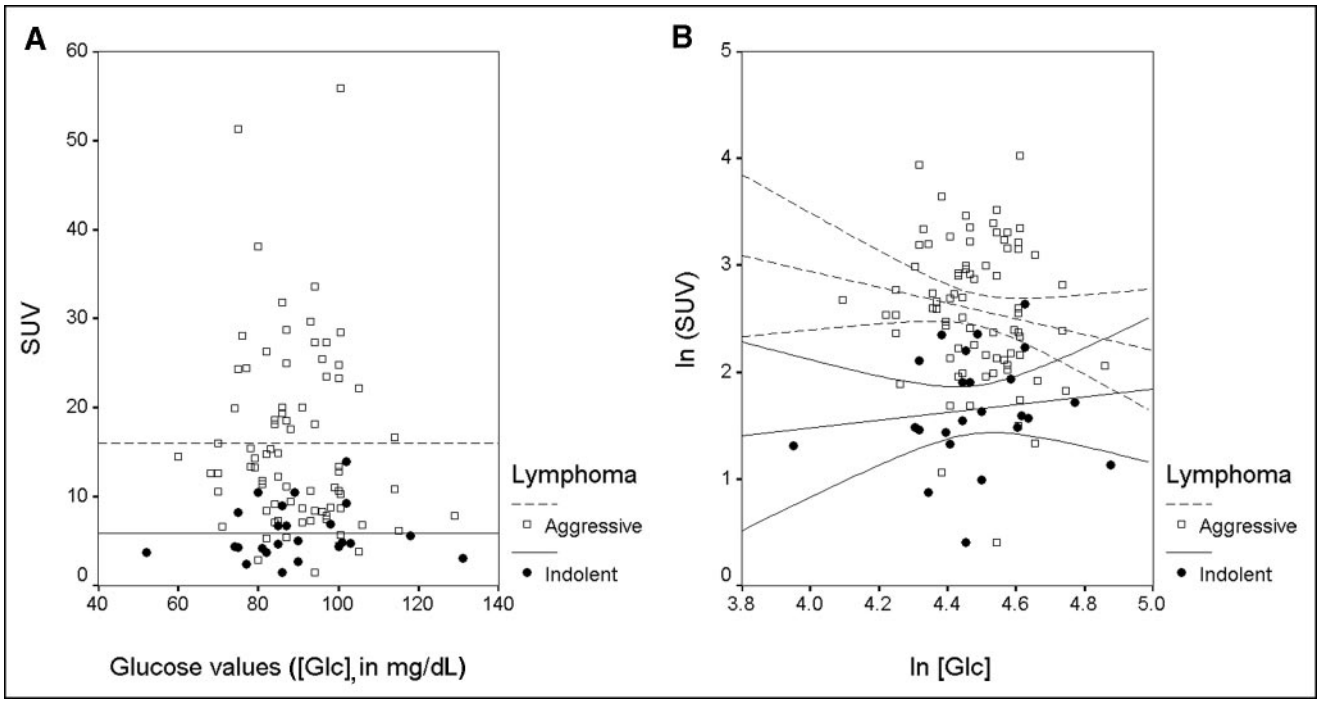

FIGURE 2. (A) Plot of original SUV without glucose correction against [Glc] for aggressive and indolent lymphomas, with mean reference lines. Boundary zone lies somewhere between these 2 reference lines. (B) Plot of ln (SUV) against ln [Glc], the logarithmic form of SUV against $[Glc]^g$ for aggressive and indolent lymphoma grades, with lines indicating mean regression lines and 95% confidence intervals. Graph shows refinement of boundary zone with proposed g-value.

Figure 1A shows that a g-value of −0.5 achieved the best classification rate. The reason for the substantial drop-off in this rate when much more negative (or higher absolute) g-values are used is that the randomness of [Glc] in the population has a greater influence on the corrected SUVs. The use of unsuitable g-values in the correction may cause greater variability than does the use of a g-value of −0.5, and the resulting larger overlap in SUVs between the 2 classes means less accurate classifications. Thus, the result of using an optimal g-value to reduce variability from [Glc] in the population is somewhat analogous to that of using clinically impractical clamps on [Glc]. The glucose-normalized SUV, or SUV_{100}, may be useful in oncologic PET measurements for comparisons and for monitoring treatment response when the [Glc] is more variable than in the current study.

Glucose sensitivity is only one of many correctable factors that influence SUVs (*5,12*) which, among other factors, has been somewhat neglected. Clinically, a glucose correction to SUV is rarely done, and even then it is universally done using a nominal g-value of −1. A special feature of the study here is that, after correction, it keeps variability to below that when no correction is used or when correction is used with a nominal g-value of −1.

This study used the maximum SUVs over the biopsy sites. Clinically, the decision on biopsy site is not based totally on metabolic status. For instance, mediastinal or spleen uptake may be more or less intense than supraclavicular or inguinal uptake. However, because of the accessibility of the superficial sites, biopsy was performed at these rather than at sites deep in the body. Thus, the study was originally designed to include this unavoidable overall systematic error. In actuality, the biopsy site chosen was that deemed to produce the least trauma. No major trauma was incurred by the minimally invasive procedures, and sufficient time elapsed after the PET scan to prevent significant inflammatory changes from being a confounding factor in the SUV calculation.

No single test is 100% accurate in clinical practice. Nonetheless, if SUV_{100} were to help better determine tumor biology and be used in conjunction with current staging and prognostic factors, one could potentially better determine which lymphomas are indolent and can be managed with watchful waiting and which are aggressive and require local or systemic therapy. In patients with disease that straddles the indolent and aggressive subtypes, such as follicular type II lymphomas, an SUV_{100} greater than 7.25 may prompt one to choose a more aggressive therapy. The SUV data might also spare a patient from multiple biopsies by directing the surgeon to an area more likely to exhibit more aggressive behavior.

Although the method is quite general, the specific data presented here pertain to indolent and aggressive lymphoma and may not apply to other types of cancer. Other neoplasms have other sensitivities to glucose (*6*) and intrinsic variations in rate constants, which dominate uptake variability.

Moreover, the scatter of the data points shown in Figure 2 indicates that other, unknown, factors also greatly affect SUV and lessen the significance of the plasma glucose adjustment. Nonetheless, with the model of SUV adjustment, one could allow greater glucose variability in patients when performing ^{18}F-FDG PET and thus relax the fasting or dietary requirements before the study. In any event, the model introduced in this study improves the classification of lymphomas, although the improvement may be only slight if there is little variability in glucose levels. It is the ability of the model to account for some unexpected variations in glucose levels during PET imaging that makes this correction useful. Further studies on other cancers and on applications after treatment are under way.

CONCLUSION

The results of this study suggest that the use of a glucose-normalized SUV from ^{18}F-FDG PET, with correction to a standard at 100 mg/dL, is a robust metabolic discriminator between lymphoma grades. This approach introduces an extra degree of freedom by seeking the best choice of g-value rather than using the customary g-value of 0 (no glucose correction) or, sometimes, −1 for the discrimination marker. A g-value less than −1.0 is most likely physiologically unrealistic. Use of an optimally chosen g-value may have varying degrees of success for various cancers under different conditions.

REFERENCES

1. Huang SC, Phelps ME, Hoffman EJ, Sideris K, Selin CJ, Kuhl DE. Noninvasive determination of local cerebral metabolic rate of glucose in man. *Am J Physiol.* 1980;238:E69–E82.
2. Phelps ME, Huang SC, Hoffman EJ, Selin CE, Kuhl DE. Tomographic measurement of regional cerebral glucose metabolic rate in man with (F-18) fluorodeoxyglucose: validation of method. *Ann Neurol.* 1979;6:371–388.
3. Sokoloff L, Reivich M, Kennedy C, et al. The [C-14]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. *J Neurochem.* 1977; 28:897–916.
4. Langen KJ, Braun U, Rota Kops E, et al. The influence of plasma glucose levels on fluorine-18-fluorodeoxyglucose uptake in bronchial carcinomas. *J Nucl Med.* 1993;34:355–359.
5. Huang SC. Anatomy of SUV. *Nucl Med Biol.* 2000;27:643–646.
6. Thie J, Smith GT, Hubner KF. Compilation of serum glucose sensitivity for FDG-PET with implications for quantitation [abstract]. *J Nucl Med.* 2004; 45(suppl):402P.
7. Wahl RL, Henry CA, Ethier SP. Serum glucose: effects on tumor and normal tissue accumulation of 2-[F-18]-fluoro-2-deoxy-D-glucose in rodents with mammary carcinoma. *Radiology.* 1992;183:643–647.
8. Lindholm P, Minn H, Leskinen-Kallio S, Bergman J, Ruotsalainen OU, Joensuu H. Influence of the blood glucose concentration on FDG uptake in cancer: a PET study. *J Nucl Med.* 1993;34:1–6.
9. Nakamoto Y, Zasadny KR, Minn H, Wahl RL. Reproducibility of common semi-quantitative parameters for evaluating lung cancer glucose metabolism with positron emission tomography using 2-deoxy-2-[^{18}F]fluoro-D-glucose. *Mol Imaging Biol.* 2002;4:171–178.
10. Graham MM, Peterson LM, Hayward RM. Comparison of simplified quantitative analyses of FDG uptake. *Nucl Med Biol.* 2000;227:647–655.
11. Wong C-YO, Thie J, Parling-Lynch KJ, et al. Glucose-normalized standard uptake value (SUV) from F-18 FDG PET in classifying lymphomas [abstract]. *J Nucl Med.* 2005;46(suppl):374P.
12. Thie JA. Understanding the standardized uptake value, its methods, and implications for usage. *J Nucl Med.* 2004;45:1431–1434.