# Comparison of SUV and Patlak slope for monitoring of cancer therapy using serial PET scans

Nanette M. T. Freedman[1], Senthil K. Sundaram[2], Karen Kurdziel[2], Jorge A. Carrasquillo[2], Millie Whatley[2], Joann M. Carson[2], David Sellers[2], Steven K. Libutti[2], James C. Yang[2], Stephen L. Bacharach[2, 3]

[1] Hadassah University Hospital, Jerusalem, Israel
[2] National Institutes of Health, Bethesda, MD, USA
[3] Bldg. 10, Room 1C401, NIH, Bethesda, MD 20892-1180, USA

Received: 18 April 2002 / Accepted: 18 July 2002 / Published online: 23 October 2002


**Abstract.** The standardized uptake value (SUV) and the slope of the Patlak plot (*K*) have both been proposed as indices to monitor the progress of disease during cancer therapy. Although a good correlation has been reported between SUV and *K*, they are not equivalent, and may not be equally affected by metabolic changes occurring during disease progression or therapy. We wished to compare changes in tumor SUV with changes in *K* during serial positron emission tomography (PET) scans for monitoring therapy. Thirteen patients enrolled in a protocol to treat renal cell carcinoma metastases were studied. Serial dynamic fluorodeoxyglucose (FDG) PET scans and computed tomography (CT) and magnetic resonance (MR) scans were performed once prior to treatment, once at 36±2 days after the start of treatment, and (in 7/13 subjects, 16/27 lesions) a third time at 92±9 days after the start of treatment. This resulted in a total of 33 scans, and 70 tumor Patlak and SUV values (one value for each lesion at each time point). SUV and *K* were measured over one to four predefined tumors/patient at each time point. The input function was obtained from regions of interest over the heart, combined, if necessary, with late blood samples. Over all tumors and scans, SUV and *K* correlated well ($r$=0.97, $P$<0.0001). However, change in SUV with treatment over all tumor scan pairs was much less well correlated with the corresponding change in *K* ($r$=0.73, $P$<0.0001). The absolute difference in % change was outside the 95% confidence limits expected from previous variability studies in 6 of 43 pairs of tumor scans, and greater than 50% in 2 of 43 tumor scan pairs. In four of the six cases, the two indices predicted opposing therapeutic outcomes. Similar results were obtained for SUV normalized by body weight or body surface area and for SUVs using mean or maximum count. Changes in CT and MR tumor cross-product dimensions correlated poorly with each other ($r$=0.47, $P$=NS), and so could not be used to determine the "correct" PET index. Absolute values of SUV and *K* correlated well over the patient population. However, when monitoring individual patient therapy serially, large differences in the % changes in the two indices were occasionally found, sometimes sufficient to produce opposing conclusions regarding the progression of disease.

*Keywords:* PET – FDG – SUV – Patlak

**Eur J Nucl Med (2003) 30:46–53**
DOI 10.1007/s00259-002-0981-4

Stephen L. Bacharach (✉)
National Institutes of Health, Bethesda, MD, USA
e-mail: steve-bacharach@nih.gov
Tel.: +1-301-4025686, Fax: +1-301-4960114

## Introduction

The ability to monitor a tumor's response to therapy is of great importance. If success or failure of response could be assessed early in the course of treatment, the treatment could possibly be altered accordingly. Positron emission tomography (PET) with fluorodeoxyglucose (FDG) is particularly suitable for therapy monitoring since it can quantify changes in metabolism with time, and such changes are thought to precede changes in lesion morphology as measured by conventional imaging modalities like computed tomography (CT) and magnetic resonance imaging (MRI) [1, 2].

The two most widely used quantitative indices of FDG metabolism are the standardized uptake value (SUV) [3, 4, 5] and the rate of FDG uptake as measured by the Patlak slope, *K* [6, 7]. SUV is a simple semiquantitative index, calculated by measuring the activity concentration in the tumor during a short-duration (typically 10–15 min) static scan acquired late (typically 45–60 min) after injection, and then normalized for the injected dose and either patient weight or lean body mass

[8, 9] or body surface area [10, 11]. To calculate the Patlak slope, *K*, dynamic imaging is required. Both a tumor time-activity curve and an arterial input function (i.e., the arterial FDG blood concentration as a function of time) must be measured. Patlak analysis is therefore much more demanding than SUV in both the acquisition and the computation stage of the process. For these reasons, patient throughput can be much higher if SUV rather than Patlak analysis is used, accounting for the wide use of the SUV method. Unfortunately, the two methods do not measure exactly the same quantity. There are three important differences between SUV and the Patlak slope. First, SUV measures the total activity in the tumor, and includes both metabolized FDG and any unmetabolized FDG in the blood, in the intracellular spaces, or in the cell (e.g., unphosphorylated FDG). Patlak analysis separates these two components out – the Patlak slope is determined only by metabolized FDG. The second important difference is that SUV often depends strongly on how long after injection the static scan is acquired. Patlak slope avoids this time dependence. Finally, Patlak slope uses the integral under the arterial input function (i.e., the sum of all the FDG available to the tumor – the "available dose") for normalization. SUV approximates this integral by the injected dose divided by the body weight (or lean body mass or body surface area).

These three differences make it possible that changes observed in SUV with therapy may not agree with the changes observed in the Patlak slope with therapy. It is the purpose of this paper to investigate whether such discordances occur in practice, and if so, to examine their causes and to determine whether the magnitudes of the discordances might affect the ability of PET to accurately monitor therapy.

To do this, we performed both SUV and Patlak analysis on 13 patients undergoing therapy for metastases from clear cell renal carcinoma. Each subject was studied at baseline and at one or two times during the course of therapy. We compared the changes in SUV with therapy to the changes in Patlak *K* with therapy.

Note that the intent of this paper was not to determine which method is "best" for monitoring changes. Rather it was to assess whether differences between the two methods occur, and to gain insight into the reasons behind those differences. This information would permit a better understanding of recent [12, 13] and future data assessing the relative clinical efficacy of the two methods.

## Materials and methods

*Patient population.* Serial FDG PET scans were acquired on 13 patients as part of a larger protocol to monitor the effects of therapy on metastases from clear cell renal carcinoma. All patients had undergone nephrectomy for removal of the primary tumor. Therapy consisted of either anti-vascular endothelial growth factor (anti-VEGF) or placebo. All subjects had previously received interleukin-2 therapy (IL-2), and had failed to respond. IL-2 was discontinued at least 4 weeks prior to the baseline PET scan. In this protocol a single index lesion (>2 cm longest dimension) was selected for study, prior to the first PET scan, based on CT and MRI data. However, often other lesions were also in the field of view. For the purposes of the present study all additional lesions in the field of view at the time of the first PET scan were also studied. A total of 27 lesions were followed over time, ranging in size at baseline (as determined by CT) from 7.0 cm to 1.9 cm (longest dimension). Six patients had only one tumor, three had two tumors, one had three tumors and three had four tumors. Fourteen lung, seven mediastinal, and six retroperitoneal metastases were studied. Dynamic FDG PET scans as well as CT and MR scans were performed prior to treatment, at ~5 weeks after the start of treatment, and (in 7/13 subjects, 16/27 lesions) at ~13 weeks after the start of treatment. This resulted in a total of 33 scans, and 70 tumor Patlak and SUV values (one value for each lesion at each time point). This protocol was approved by the National Cancer Institute's Institutional Review Board, and written informed consent was obtained from all patients before entry.

*Image acquisition.* All PET scans were acquired on a GE Advance PET scanner [14] producing 35 slices over an approximately 15-cm axial field of view. In-plane and axial reconstructed resolution was ~6.5 mm full-width at half-maximum (at the center of the field of view), with a slice separation of 4.25 mm. Images were reconstructed into a 256 by 256 array (2 mm/pixel), using filtered back-projection for dynamic data and iterative reconstruction (OSEM – 28 iterations, 4 subsets) for static images; reconstructed resolution was the same for both reconstruction methods. Patients fasted for at least 6 h. Approximately 10 mCi (370 MBq) fluorine-18 FDG was injected over a 2-min period using a constant infusion pump, and dynamic PET scans were acquired as follows. If both the heart and the tumor could be included in the same single field of view then the dynamic acquisition began at injection, with scan times of 30 s/frame for the first 4 min, 3 min/frame for the next 18–21 min and 5 min/frame thereafter for an average of 58 (±3.6) min total. In three of the 13 patients, the heart was not in the field of view of the tumor. In this case the initial 30 min of data were acquired with the patient's heart in the field of view (scan times as described above). The bed was then moved so that the tumor was in the field of view, and dynamic scanning was resumed (with the same scan times as described above). This permitted the early portion of the FDG input function always to be obtained from the cardiac image data regardless of whether the tumor was in the field of view of the heart or not. Venous blood samples were acquired at an average of 17.6, 20.8, 24.0, 30.2, 39.1, and 55.1 min. Samples from 25 min on were used to continue the image-based input function when the tumor was not in the field of view of the heart. Samples from <25 min were used for quality control to ensure venous blood samples agreed with image-based values. Absolute activity concentrations were determined from these weighed blood samples using a well counter. All data were corrected for attenuation using an 8-min measured transmission scan. The late dynamic images were summed, yielding a static image with an average duration of 18.4±4.3 min, and a scan mid time of 48.9±2.0 min post injection.

Regions of interest (ROIs) were drawn over the tumors on each patient from the static scan using MedX (Sensor Systems, Inc, Reston, Va.) and a three-dimensional, automatic threshold-based, region growing program. All ROIs were confirmed visually. These three-dimensional regions were then used both for calculation of the SUVs and to generate time-activity curves for the

Patlak analysis. We deliberately used the same ROIs for both SUV and Patlak in order to eliminate any effect of variability due to ROI size or placement on the comparison of serial changes in SUV and Patlak. Small ROIs (1.68 $cm^2$ average) were drawn manually on the cardiac left atrial cavity (visualized from the early FDG images) to obtain an image-based input. The regions were small and far from the myocardial wall, to minimize spillover and partial volume effects.

SUVs were calculated using both the mean activity concentration within the ROIs and (after performing a 5-mm three-dimensional gaussian smooth to avoid bias due to image noise) using the maximum activity concentration within the ROI. Activity concentrations within the ROI were normalized by the injected dose, and by using each of three different corrections for body habitus: lean body mass (LBM), body surface area (BSA) and weight (WT) in kg. LBM was calculated for male and female patients from the formulae:

$$\text{LBM (m)} = 1.1 \times \text{weight} - 120\ (\text{weight}/\text{height})^2$$

$$\text{LBM (f)} = 1.07 \times \text{weight} - 148\ (\text{weight}/\text{height})^2$$

where the patient weight is given in kg, and height in cm [15].

BSA was given by the following formula [16]:

$$\text{BSA}\ (\text{m}^2) = 0.007184 \times \text{weight(kg)}^{0.425} \times \text{height}\ (\text{cm})^{0.725}$$

SUVs computed from mean or maximum counts and normalized to body weight, BSA, or LBM are referred to as $SUV_{mean\text{-}WT}$ or $SUV_{max\text{-}WT}$, $SUV_{mean\text{-}BSA}$ or $SUV_{max\text{-}BSA}$, and $SUV_{mean\text{-}LBM}$ or $SUV_{max\text{-}LBM}$.

Patlak analysis was performed using the input function and tumor time-activity curves. The data from 10 min (to allow for equilibration) to the end of acquisition were used to obtain the slope $K$ and intercept $V_0$ (the FDG volume of distribution), as well as to determine the estimated errors in those two parameters. All computations and linear regressions were performed using IDL [Research Systems Inc., Boulder, Colo.].

## Results

At the times of the three serial scans, means (and standard deviations) of $SUV_{mean\text{-}LBM}$ were 4.58 (2.74), 4.66 (2.94), and 3.43 (1.29) respectively, while corresponding values for Patlak slope, $K$, were 0.029 (0.022), 0.029 (0.022), and 0.022 (0.014) respectively ($P$=NS for all pairs of study times). Table 1 summarizes $SUV_{mean\text{-}LBM}$ and Patlak slope values for all tumor studies, pre and post therapy. Differences between pre- and post-therapy values were not significant; differences between post-placebo and post-anti-VEGF were also not significant. Figure 1 compares the Patlak slopes, $K$, with the corresponding $SUV_{mean\text{-}LBM}$ values for all 70 tumor studies, demonstrating a good correlation (r=0.97, SEE=0.0053, $P$<0.0001) over a wide range of $SUV_{mean\text{-}LBM}$ and $K$ values. Similar results ($r$=0.93, SEE=0.0078, $P$<0.0001) were found when $SUV_{max\text{-}LBM}$ rather than $SUV_{mean\text{-}LBM}$ was used. The correlations remained high when the analysis was done on a per patient basis, rather than per tumor (13 index lesions, $r$=0.98, $P$<0.0001 for $SUV_{mean\text{-}LBM}$). Correlation was not significantly different for pre- and post-therapy studies with placebo or with anti-VEGF. Despite overall good agreement, Fig. 1 shows that some data points do not lie near the predicted straight line and a few of the measured SUVs deviate by as much as 50% from the predicted values.

**Table 1.** SUV and Patlak $K$: summary: (no significant differences were observed between columns)

| | Before therapy (baseline + placebo) | Post anti-VEGF therapy |
|---|---|---|
| SUV, mean±SD | 4.25±2.55 | 4.44±2.66 |
| SUV, range | 1.49–9.81 | 1.64–11.04 |
| Patlak $K$, mean±SD | 0.026±0.021 | 0.029±0.020 |
| Patlak $K$, range | 0.005–0.074 | 0.005–0.074 |
| SUV vs $K$ correlation coefficient, $r$ | 0.977 | 0.960 |
| No. of observations | 34 | 36 |

All SUV values shown are $SUV_{mean\text{-}LBM}$

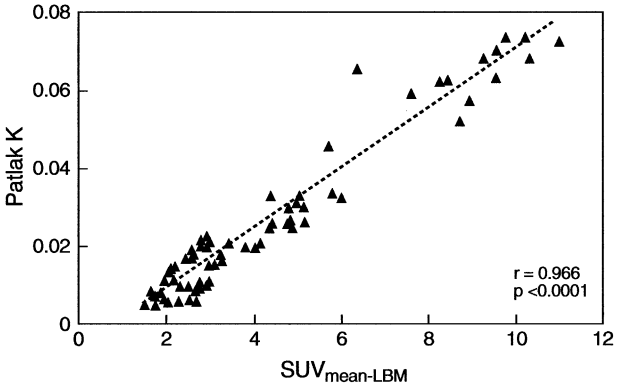


**Fig. 1.** Regression line ($K$=0.0077×SUV–0.006) for Patlak slopes, $K$, vs $SUV_{mean\text{-}LBM}$ values for all the 70 tumor studies (27 tumors studied two to three times each)

Mean absolute percent changes in SUV and Patlak $K$ between serial scans for individual tumors were 19.7% (SD 14.8%) and 26.0% (SD 22.1%) respectively. Figure 2 shows these percent changes in SUV (%ΔSUV) between serial scans for each tumor, compared with the corresponding changes in $K$ (%Δ$K$). Although the correlation in %Δ$SUV_{mean\text{-}LBM}$ versus %Δ$K$ is still significant, it is much poorer ($r$=0.73) than the correlation between $K$ and $SUV_{mean\text{-}LBM}$ (Fig. 1), with a large scatter of the data about the line (SEE=0.23). When only one lesion per patient was considered, the correlation remained poor ($r$=0.68). Examining Fig. 2, occasional large differences are seen between the serial changes in FDG uptake predicted by $SUV_{mean\text{-}LBM}$ and those predicted by the Patlak slope. The average absolute % difference was 15.2% over all tumor pairs.

Weber et al. (and others) [17, 18] estimated the variability in $K$ or SUV values of tumors by using the method of repeated measures. Figure 3 in Weber's paper indi-

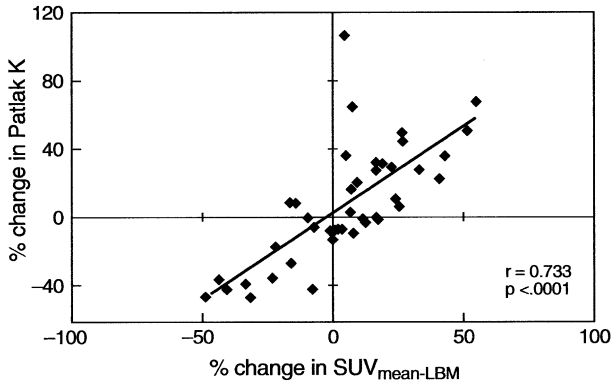


**Fig. 2.** Regression line ($\Delta K$=1.01×ΔSUV+0.034) for % change in $SUV_{mean\text{-}LBM}$ vs % change in $K$ from the first to the second scan in each scan pair for each tumor

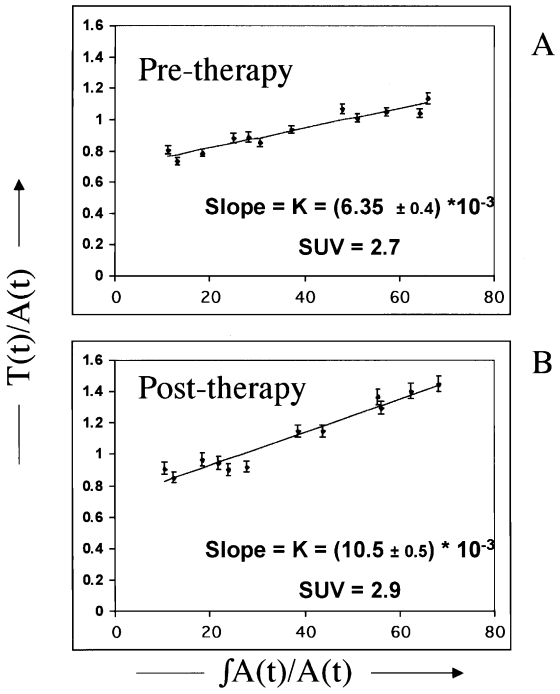


**Fig. 3.** Patlak plots for a tumor (**A**) pre-therapy and (**B**) during therapy; slopes ($K$ values) and corresponding $SUV_{mean\text{-}LBM}$ values are also indicated. $A(t)$ and $T(t)$ are the arterial concentration and tumor concentration of activity as a function of time

cates the % change in $K$ (or SUV), which could be considered a "true" change in glucose metabolism, at the 95% confidence level. In the present paper, when discrepancies between $\%\Delta K$ or %ΔSUV occurred, we defined those discrepancies as "large" if they were at least as big as the % changes indicated as significant in Weber's Fig. 3. Weber's Fig. 3 indicated that for low $K$ values, $\%\Delta K$ needed to be very high in order for a metabolic change to be considered a "true" change while for high $K$'s the $\%\Delta K$ could be much lower. To be conservative, we never counted a discrepancy of less than 20% to be a true discrepancy, but for tumors with lower metabolic rates we required the discrepancy to be much larger, just as specified in Weber's Fig. 3.

Defined in this way, "large" discrepancies between $\%\Delta K$ and $\%\Delta SUV_{mean\text{-}LBM}$ were found in 6/43 tumor scan pairs (average absolute difference 44.2%) and occurred in four of the 13 patients. The discrepancies were greater than 50% in 2 of 43 pairs (average absolute difference 79.9%), in two separate patients. In four of the six cases where the discrepancy was "large", the two indices predicted opposing therapeutic outcomes. Large discrepancies between $\%\Delta K$ and $\%\Delta SUV_{mean\text{-}LBM}$ were not confined to tumors with low initial uptake; range of initial $SUV_{mean\text{-}LBM}$ in these six tumors was 1.95–7.61. When changes in $SUV_{max\text{-}LBM}$, instead of $SUV_{mean\text{-}LBM}$, were compared to changes in $K$, similar results were found (Table 2).

Changes in CT and MR tumor cross-product dimensions were computed in one tumor (the index lesion) from each patient. Changes in CT size correlated poorly with both $\%\Delta K$ and $\%\Delta SUV_{mean\text{-}LBM}$ ($r$=0.49, $P$=0.06 and $r$=0.45, $P$=0.08). The changes in CT tumor cross-product size ranged from a decrease of 54% to an increase of 20%, but the average absolute change was only 16.9%. The changes in CT and MR dimensions also correlated poorly with one another ($r$=0.47, $P$=NS).

In keeping with recommendations [5] for obtaining reproducible measures of SUV, the static scan time was kept quite constant (48.9±2 min) for all studies. For some but not all the tumors, FDG uptake was still rising at that time. However, there was no significant relationship between whether the curve was rising or not and magnitude of difference between %ΔSUV and $\%\Delta K$. Mean discrepancies were 18.9% and 10.8% for rising

**Table 2.** Effect of normalization on correlation of %ΔSUV with $\%\Delta$Patlak $K$

| | $\%\Delta SUV_{mean\text{-}LBM}$ | $\%\Delta SUV_{max\text{-}LBM}$ | $\%\Delta SUV_{mean\text{-}WT}$ | $\%\Delta SUV_{mean\text{-}BSA}$ |
|---|---|---|---|---|
| Correlation coefficient, $r$ | 0.732 | 0.718 | 0.737 | 0.745 |
| SEE | 0.230 | 0.235 | 0.228 | 0.225 |
| $P$ | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| No. of "large" discrepancies | 6 | 7 | 6 | 6 |

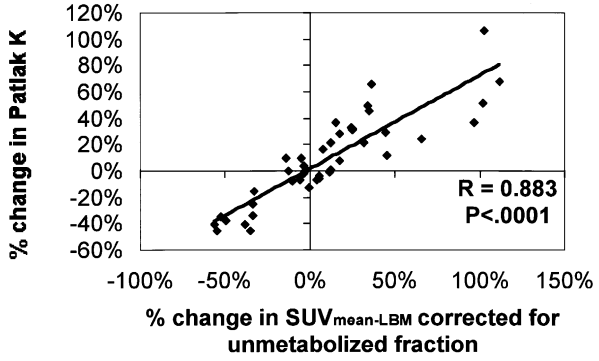


**Fig. 4.** Regression line ($\Delta K$=0.708×ΔSUV+0.012) for % change in SUV corrected for unmetabolized activity vs % change in $K$

and flat curves respectively ($P$=NS), and both discrepancies >50% occurred in cases where the curves were flat. Nor was there any relationship between the magnitude of the discrepancy and tumor size on PET ($r$=0.097, $P$=NS).

We also examined the effect of the error caused by statistical fluctuation when estimating the Patlak slope from relatively noisy dynamic data. The average error in slope was 4.1%±4.7%, too low to explain the observed discrepancies. Figure 3 shows the Patlak plots, before and during therapy, for a tumor with a large discrepancy between %ΔSUV and %Δ$K$. The Patlak slope increased by more than 50%, from 6.35 (±0.4)×$10^{-3}$ to 10.5(±0.5)×$10^{-3}$, while the SUV increased by only 7%, from 2.7 to 2.9. The uncertainties in fitting a line to the data, as seen in Fig. 3, are far too small to explain the discrepancy between the observed %Δ$K$ and %ΔSUV, even though the errors in fitting the Patlak slope in this case were larger than the average error. Note that while this large discrepancy occurred for low $K$, some discrepancies also occurred at higher $K$.

Having investigated these possible "technical" reasons for discrepancy between %ΔSUV and %Δ$K$, we next examined the relative contributions of two of the fundamental differences between SUV and Patlak $K$, as described in the Introduction. The Patlak slope, $K$, indicates only the rate of uptake of metabolized FDG, while SUV includes metabolized and unmetabolized FDG. The fraction of unmetabolized activity can be calculated from the intercept of the line on the Patlak plot. This "unmetabolized fraction" varied considerably between tumors, ranging from 6% to 67% at the time of the static scan. The unmetabolized fraction also often changed between serial scans, with a mean change of 9.1% (SD 7.1%). Correcting SUV values for the unmetabolized activity reduced some but not all discrepancies, and markedly improved the correlation between %ΔSUV and %Δ$K$, from $r$=0.73 to $r$=0.88 (Fig. 4). The average absolute difference in % discrepancy was reduced slightly from 15.2% to 14.5%, but with considerably reduced variation (Table 3).

The last fundamental difference between Patlak analysis and SUV is that the former properly accounts for the FDG available to the tumor, the so-called available dose, through the $\int A(t)$ term in the Patlak equation. In the $SUV_{mean\text{-}LBM}$ the available dose is assumed to be 1/[(injected dose)×LBM]. Correction of the SUVs for the available dose [using the measured $A(t)$] also improved the correlation ($r$ increased from 0.73 to 0.85), and reduced the average % discrepancy (from 15.2% to 12.4%).

To determine whether these two factors together might explain the majority of the discrepancies, we also looked at the result of correcting both for unmetabolized fraction and for the available dose. The result was a dramatic improvement (Table 3), with correlation of serial changes up from 0.73 to 0.97, mean % discrepancy down from 15.2% to 7.2%, standard deviation now only 5.3%, and no "large" discrepancies. The results of all the corrections attempted are summarized in Table 3. Of course, neither of these corrections is practical without a dynamic acquisition, but making them gives insight into the cause of the discrepancies.

## Discussion

FDG PET is increasingly used to monitor changes in tumor FDG metabolism with therapy. Such changes are a possible surrogate marker for tumor response and could perhaps assist in making early decisions regarding the

**Table 3.** Effect of correction of SUV for unmetabolized FDG and available dose on correlation between changes in SUV and Patlak $K$

| | $SUV_{mean\text{-}LBM}$ uncorrected | $SUV_{mean\text{-}LBM}$ corrected for unmetabolized fraction | $SUV_{mean\text{-}LBM}$ corrected for available dose | $SUV_{mean\text{-}LBM}$, both corrections |
|---|---|---|---|---|
| Correlation coefficient, $r$ | 0.733 | 0.884 | 0.849 | 0.975 |
| Mean discrepancy | 15.2% | 14.5% | 12.5% | 7.2% |
| SD discrepancy | 17.0% | 13.3% | 15.6% | 5.3% |
| No. of "large" discrepancies | 6 | 6 | 5 | 0 |

course of therapy. The SUV is an objective quantitative measure and is easy to determine. However, this simple measure of activity concentration in a tumor at an arbitrary time post injection, normalized for injected tracer dose/patient weight, may not always accurately reflect tracer metabolism [4, 5]. Figure 1 echoes the findings of others [19, 20, 21] that SUV correlates well but not perfectly with glucose metabolic rate, as estimated by Patlak analysis. The rationale for our study was to determine whether the changes in SUV with therapy agree with changes in tumor metabolism as measured by Patlak analysis, and if not, to determine why, and whether these discrepancies would affect decisions regarding therapy regimen.

Despite the good correlation between SUV and Patlak $K$ for single studies (Fig. 1), when we compared serial changes in these two indices following therapy, we found occasional large discrepancies between the changes in SUV compared with the changes in Patlak $K$. Such discrepancies occurred both for small and for large initial SUVs, and were not confined to one or two patients, but affected measurements from 4 of the 13 patients in this study. They were sufficiently large that they could have resulted in different conclusions about the response to therapy, according to the criteria of Weber et al. (and others) [17, 18].

Weber et al. [17] indicated that the principal sources of variability in measurement of SUV and Patlak $K$ were methodological issues such as accuracy of placement of ROIs and uncertainty in fitting the Patlak curve, as well as the day to day physiologic fluctuations in FDG uptake. While accurate placement of ROIs may affect the absolute values of SUV and Patlak $K$, it will have no impact on our comparison of the changes in these two indices, since exactly the same ROIs and the identical FDG data set were used for both. In addition, if physiologic variability affected both Patlak and SUV in the same way, then our comparisons of the change in SUV with the change in Patlak slope would be unaffected by such variability. Our variability, then, might be even smaller than the values reported by Weber et al.

Having identified clinically meaningful discrepancies between %ΔSUV and %Δ$K$ in 4/13 subjects, we asked what factors might cause these discrepancies. We noted, interestingly, that a discrepancy which occurred in one tumor in a subject, did not necessarily occur in other tumors in the same subject. This implies that errors in the input function, or in the LBM normalization, or in other patient-wide measures, could not be the sole cause of the discrepancies. We further investigated the effect both of patient-wide factors and of factors which might affect individual tumors, even in a single scan.

Within the category of patient-wide factors, we found that normalization of SUV for body weight or BSA rather than LBM did not significantly alter the discrepancies. This may be because neither changes in body weight nor changes in LBM nor changes in BSA (all used in the ΔSUV calculation) are likely to consistently represent the changes in the dose available to the tumor as described by the area under the input function, $\int A(t)$ (used in the Δ$K$ calculation), that occur from one serial scan to the next. We often found substantial changes in available dose between serial scans, as estimated from the input function, when there was no substantial change in (injected dose)/LBM. Changes in available dose can occur as a consequence of changes in a subject's physiologic state and metabolic status from one study to the next. Such changes would not be unexpected as a result of disease progression and/or effects of therapy in cancer patients. Correcting the SUV values for the available dose obtained from the integral of the input function resulted in a significantly improved correlation (from $r$=0.73 to $r$=0.85, $P$=0.078) and reduced average discrepancy. This suggests that an important contribution to the observed discrepancies may be the fact that SUV does not account for available dose while Patlak $K$ does.

Among the factors which are unique to each tumor, we investigated the effects of using mean versus maximum counts for SUV, slope of tumor time-activity curve, partial volume effect (i.e., tumor size), accuracy of Patlak fit, and fraction of unmetabolized tracer in the tumor ROI.

Use of maximum rather than mean counts for the SUV did not alter the correlation between changes in SUV and Patlak $K$. This was true despite the fact that the time-activity curves used for the Patlak analysis were based on mean counts.

The attraction of SUV, i.e., the simplicity of using only one "static" image, is also its limitation. Often the tumor time-activity curves were still rising even at 1 h. The rate of this rise may differ for the same tumor pre and post treatment. Unless the tumor time-activity curve has reached a plateau, this will cause errors in the change in SUV measured at a fixed time post injection. Despite this, we found no consistent relationship between whether the curves were rising or flat, and the discrepancy between changes in SUV and Patlak $K$.

Likewise, we could find no relationship between the discrepancies and the tumor sizes, which might have been caused by the partial volume effect. Presumably this was because the loss in counts produced by the blurring out of counts from the ROI would have exactly the same effect on the mean SUV values as on the Patlak slopes. It is possible, however, that blurring of counts into the tumor ROI might be time dependent, thereby affecting the $K$ value and the SUV in different ways.

There is always some error in estimating the Patlak slope using points calculated from relatively noisy dynamic data. However, these errors were found to be very small, and could not therefore explain the observed discrepancies.

On the other hand, the fraction of unmetabolized tracer in tumors was seen to have a substantial effect on the discrepancies. Possibly this was due to changes in tumor

structure with time or therapy. Correction for the unmetabolized FDG reduced some though not all discrepancies, and markedly improved the correlation of the % serial changes (from $r$=0.73 to $r$=0.88, $P$=0.020), indicating that unmetabolized FDG might be one of the major factors causing the discrepancies between %ΔSUV and %ΔPatlak for predicting response.

Since some patients in our study were treated with anti-VEGF, while others received placebo, we wondered whether the possible effects of anti-VEGF on body distribution of FDG uptake (and therefore on the input function) and/or on the unmetabolized fraction of FDG in tumors might give rise to increased discrepancies between %ΔSUV and %ΔPatlak. However some large discrepancies also occurred in the placebo patients, and there was no difference between the two groups in the correlation between %ΔSUV and %ΔPatlak. While a treatment such as anti-VEGF might potentially increase such discrepancies, it appears that they can also occur without such a specific cause, presumably as a result of metabolic variation or disease progression.

In summary, two effects stood out as having a major impact on the discrepancies between SUV and Patlak $K$ for predicting response – the available dose and the unmetabolized fraction. Correcting the SUV for each of these factors separately improved the correlation between %ΔSUV and %Δ$K$, and correcting for both of them together resulted in a dramatic improvement, with excellent correlation of serial changes ($r$=0.97), much reduced average discrepancy, and no clinically significant discrepancies. This finding is not surprising since these factors both appear in the Patlak equation, while the SUV equation ignores unmetabolized FDG and assumes available dose can be approximated with injected dose and body composition. These findings support the hypothesis that these two factors together may account for most of the discrepancies. Since both of these factors are in theory essential to accurate determination of glucose metabolic rates, we hypothesize that the Patlak values would be the better of the two indices for monitoring therapy. Unfortunately, we had no definitive data to verify this hypothesis. Comparison with size changes seen on CT and MR did not provide clarification, since changes on CT correlated equally poorly with changes in Patlak $K$ and in SUV, and there was sometimes disagreement between MR and CT, even in cases where one modality indicated quite a large change. Discordance between change in tumor size on MR and CT may have been due to the fact that many changes were small, or may have resulted from the difficulty in defining exact tumor extent on MR. Discordance between FDG PET and CT or MR may have arisen because the follow-up studies were performed early during therapy (there was only 5 weeks and 8 weeks between studies), when morphologic changes might not yet have occurred.

Another limitation of the study was its relatively small sample size, and its restriction to one tumor type. Despite the small size, however, the finding of such large discrepancies in four of the 13 subjects is strong evidence that the discrepancies may be clinically important. We cannot, of course, be certain that similar discrepancies would be seen for all tumor types, but at least one of the underlying causes of the discrepancies (the available dose) might well be independent of tumor type.

It is unfortunate that occasional discrepancies occur between serial change measured by SUV and by Patlak $K$. If both indices were equally reliable for measuring serial change in FDG tumor uptake, then one could safely use the much simpler SUV method. It is also unfortunate that no easy method comes to mind that would allow us to correct the SUV (assuming it is the SUV that is incorrect) without a dynamic acquisition. Our data investigating the causes of the discrepancy can at least guide us in exploring this possibility in the future. Presumably waiting a considerably longer time before static imaging would minimize the unmetabolized FDG component of the discrepancy. Proposed "simplified kinetic analysis" methods [22, 23] might also be able to correct for some but not all the causes of discrepancy, and perhaps some combination of these methods may be found which can compensate for changes in the available dose and minimize unmetabolized FDG without a full dynamic study.

### *Conclusion*

Absolute values of SUV and the Patlak slope $K$ correlated well over the patient population. However, when using SUV and Patlak $K$ to monitor therapy, the changes observed in the two indices were occasionally discrepant. These discrepancies were often sufficient to produce differing conclusions regarding the progression of disease. Our data suggest the discrepancies are primarily due to the unmetabolized FDG measured by SUV, and the inability of SUV to account for the available dose. Both these factors are in theory important to the accurate determination of glucose metabolic rates, suggesting (but not confirming) that Patlak slope is the more accurate index. Further research would be necessary to determine whether this increased accuracy translates into improved clinical efficacy. Our results suggest that further efforts in improving SUV calculations should focus on better approximations for available dose, or on methods to reduce the effects of unmetabolized FDG.

## References

1. Price P, Jones T. Can positron emission tomography (PET) be used to detect subclinical response to cancer therapy? *Eur J Cancer* 1995; 31A:1924–1927.
2. Young H, Baum R, Cremerius U, Herholz K, Hoekstra O, Lammerstsma AA, Pruim J, Price P. Measurement of clinical and subclinical tumour response using [$^{18}$F]-fluorodeoxyglucose and positron emission tomography: review and 1999

EORTC recommendations. *Eur J Cancer* 1999; 35:1773–1782.
3. Huang SC. Anatomy of SUV. Standardized uptake value. *Nucl Med Biol* 2000; 27:643–646.
4. Hamberg LM, Hunter GJ, Alpert NM, Choi NC, Babich JW, Fischman AJ. The dose uptake ratio as an index of glucose metabolism: useful parameter or oversimplification? *J Nucl Med* 1994; 35:1308–1312.
5. Keyes JWJ. SUV: standard uptake or silly useless value? *J Nucl Med* 1995; 36:1836–1839.
6. Patlak CS, Blasberg RG, Fenstermacher JD. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *J Cereb Blood Flow Metab* 1983; 3:1–7.
7. Patlak CS, Blasberg RG. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. Generalizations. *J Cereb Blood Flow Metab* 1985; 5:584–590.
8. Zasadny KR, Wahl RL, Standardized uptake values of normal tissues at PET with 2-[fluorine-18]-fluoro-2-deoxy-D-glucose: variations with body weight and a method for correction. *Radiology* 1993; 189:847–850.
9. Sugawara Y, Zasadny KR, Neuhoff AW, Wahl RL. Reevaluation of the standardized uptake value for FDG: variations with body weight and methods for correction. *Radiology* 1999; 213:521–525.
10. Kim CK, Gupta NC, Chandramouli B, Alavi A. Standardized uptake values of FDG: body surface area correction is preferable to body weight correction. *J Nucl Med* 1994; 35:164–167.
11. Schomburg A, Bender H, Reichel C, Sommer T, Ruhlmann J, Kozak B, Biersack HJ. Standardized uptake values of fluorine-18 fluorodeoxyglucose: the value of different normalization procedures. *Eur J Nucl Med* 1996; 23:571–574.
12. Wu H, Dimitrakopoulou-Strauss A, Heichel TO, Lehner B, Bernd L, Ewerbeck V, Burger C, Strauss LG. Quantitative evaluation of skeletal tumours with dynamic FDG PET: SUV in comparison to Patlak analysis. *Eur J Nucl Med* 2001; 28:704–710.
13. Nitzsche EU, Hoegerle S, Mix M, Brink I, Otte A, Moser E, Imdahl A. Non-invasive differentiation of pancreatic lesions: is analysis of FDG kinetics superior to semiquantitative uptake value analysis? *Eur J Nucl Med* 2002; 29:237–242.
14. DeGrado TR, Turkington TG, Williams JJ, Stearns CW, Hoffman JM, Coleman RE. Performance characteristics of a whole-body PET scanner. *J Nucl Med* 1994; 35:1398–1406.
15. Morgan DJ, Bray KM. Lean body mass as a predictor of drug dosage. Implications for drug therapy. *Clin Pharmokinetics* 1994; 26:292–307.
16. DuBois D, DuBois EF. A formula to estimate the approximate surface area if height and weight be known. *Arch Intern Med* 1916; 17:863–871.
17. Weber WA, Ziegler SI, Thodtmann R, Hanauske AR, Schwaiger M. Reproducibility of metabolic measurements in malignant tumors using FDG PET. *J Nucl Med* 1999; 40:1771–1777.
18. Romer W, Hanauske AR, Ziegler S, Thodtmann R, Weber W, Fuchs C, Enne W, Herz M, Nerl C, Garbrecht M, Schwaiger M. Positron emission tomography in non-Hodgkin's lymphoma: assessment of chemotherapy with fluorodeoxyglucose. *Blood* 1998; 91:4464–4471.
19. Minn H, Leskinen-Kallio S, Lindholm P, Bergman J, Ruotsalainen U, Teras M, Haaparanta M. [$^{18}$F]fluorodeoxyglucose uptake in tumors: kinetic vs. steady-state methods with reference to plasma insulin. *J Comput Assist Tomogr* 1993; 17:115–123.
20. Minn H, Zasadny KR, Quint LE, Wahl RL. Lung cancer: reproducibility of quantitative measurements for evaluating 2-[F-18]-fluoro-2-deoxy-D-glucose uptake at PET. *Radiology* 1995; 196:167–173.
21. Lodge MA, Lucas JD, Marsden PK, Cronin BF, O'Doherty MJ, Smith MA. A PET study of $^{18}$FDG uptake in soft tissue masses. *Eur J Nucl Med* 1999; 26:22–30.
22. Hunter GJ, Hamberg LM, Alpert NM, Choi NC, Fischman AJ. Simplified measurement of deoxyglucose utilization rate. *J Nucl Med* 1996; 37:950–955.
23. Graham MM, Peterson LM, Hayward RM. Comparison of simplified quantitative analyses of FDG uptake. *Nucl Med Biol* 2000; 27:647–655.