
SUV: Standard Uptake or Silly Useless Value?

John W. Keyes, Jr.

PET Center, Bowman Gray School of Medicine, Winston-Salem, North Carolina

Key Words: standard uptake values; positron emission tomography; tracer quantitation

J Nucl Med 1995; 36:1836–1839

Quantification has always held a pervasive allure for nuclear medicine. There is a sense that anything that can be quantified should be. Once we have the numbers in hand, there is a mysterious tendency to accord them special value. Why this should be the case is unclear. Perhaps it is because of the functional nature of nuclear medicine studies, which seem as if they should be quantified. Perhaps it is because it is so easy to extract numbers from our studies. Whatever the cause, the phenomenon has been particularly strong in PET, in which we hear at every turn that PET is “quantitative.”

DiChiro and Brooks commented on this phenomenon years ago (1), noting that, in many applications, subjective interpretation of PET images is actually superior to the use of quantitative data. Undaunted, however, the field has continued to lean heavily toward numbers as the final arbiter of correctness.

A common and particularly egregious misapplication of quantification of PET data occurs in the use of the SUV (standardized uptake value) in cancer studies. As generally used, this value is defined as the tissue concentration of tracer as measured by a PET scanner divided by the activity injected divided by body weight (2). This measure is also referred to variously as the DUR [differential uptake ratio (3) or dose uptake ratio (4)], the DAR [differential absorption ratio (5) or dose absorption ratio (6)] and several other similar terms. Although one might be immediately suspicious of anything with so many aliases, these values now appear in many reports on the use of PET in cancer. It is often used as a measure to characterize the malignancy versus benignancy of lesions. Indeed, it is almost necessary to include this measure to get a paper past initial review. Yet when viewed objectively, the SUV as it is now currently used is so flawed as a quantitative measure as to be virtually worthless for the purpose for which it is usually used.

FACTORS AFFECTING SUV

The primary problem with the SUV is that it is subject to too many sources of variability which are not controlled or even taken into account in most reported studies. Table 1 lists several of the most important sources of error in the SUV. This table is not meant to be comprehensive but only to outline the major sources of variability in the SUV independent of actual lesion characteristics.

Patient Size

Body composition and habitus are a source of variability because the conventional SUV normalizes for body weight. Fat, however, has a much lower uptake of FDG (and probably other PET tracers used in oncology) than other tissues. Consequently, the SUV for many tissues shows a strong positive correlation with weight (2,7). Zasadny and Wahl (2) found, for example, that the SUV of normal liver varied by almost 50% over a range of body weights from 50 to 110 kg. Corrections for this effect have been proposed by Zasadny and Wahl (2) and Kim et al. (7), but such corrections are not used in the majority of published papers that use the SUV.

Standardized Measurement Times

The time from tracer injection to the time of PET scanning (uptake period) has been shown to have an extreme effect on the SUV. Hamburg et al. (4) have clearly shown that the FDG uptake in lung carcinoma does not plateau for several hours. Scans obtained at the usual imaging times of 45–60 min postinjection occur during a rapid uptake phase for FDG and hence are subject to great variability. They note for example “The difference between the plateau DUR and the 60 min value was $46\% \pm 6\%$ pretreatment These data indicate that the DUR can vary widely with the time of measurement. . . .” (4). The plateau value was not reached for almost 5 hr in the pretreatment group. Furthermore, these investigators show that the slope of tracer uptake varies considerably before and after treatment. Since a common use of the SUV is to quantify tumor response to treatment, this adds another potential source of variability if imaging times are not carefully controlled. They advise delaying scan times for several hours to ensure that measurements are made at plateau values. No reports have been published to date which utilize such prolonged imaging delays.

The time after injection is such an important variable that it can be significant even during scanning on a single patient. In our laboratory, we have observed that the SUV

Apr. 24, 1995; Jan. 21, 1995.

For correspondence or reprints contact: John W. Keyes, Jr., MD, PET Center, Bowman Gray School of Medicine, Winston-Salem, NC 25175-1088.

TABLE 1
Factors Affecting the SUV

Body composition and habitus
 Length of uptake period
 Plasma glucose
 Recovery coefficient and partial volume effects

in a lung tumor increased from 5.5 to 7.7 (40%) between 30 and 60 min postinjection. This interval is well within the time required to scan multiple levels within the body and suggests that significant differences in SUV values might be found in a single tumor during a single scanning session depending on where in the body and, hence, when in the scanning sequence, the tumor is visualized.

Given the magnitude of these time effects, it is surprising that they have not been better recognized and controlled. In many published studies, the actual time of imaging is not even described. When this information is available, it is often evident that significant variability occurs in this parameter from patient to patient.

Plasma Glucose Levels

The plasma glucose level at the time of study also has a major effect on the SUV. Langen et al. (6) studied a group of patients with bronchogenic cancer fasting and repeated the measurements after infusing sufficient glucose to approximately double the plasma glucose level (avg 84.6 mg/100 ml versus 168.3 mg/100 ml). The tumor DAR (same as the SUV) fell 41.8% from a fasting mean of 5.07 to a hyperglycemic mean of 2.84. Lindholm et al. (8) found changes of similar magnitude in a group of patients with head and neck cancer using oral glucose loading. It is clear that intrasubject variations in plasma glucose can have highly significant effects on measured SUVs. There appears to be no reason to believe that intersubject variations are not just as great. Although most published reports indicate that patients were studied in the fasting state, the actual plasma glucose values are rarely mentioned and never corrected.

Despite the tendency to lower the SUV, the effect of raising plasma glucose on actual tumor visualization can be difficult to predict because the tumor-to-background ratio can change favorably due to changes in normal tissue FDG uptake. For example, Ishizu et al. (9) found that the tumor-to-cortical gray matter ratio in brain tumors increased 27% after glucose loading, although the actual percent uptake of FDG (and by inference the SUV) in the tumors declined by factors of 2 to 3. In this setting, subjective visualization of the tumor is actually improved with a lower SUV.

Recovery Coefficients and Partial Volume Effects

Finally, the problem of recovery coefficient and partial volume effects must be considered. These two problems are related but also distinctly different. The recovery coefficient can be defined as the ratio of the measured activity in a lesion divided by the true activity. Partial volume refers to portions of a tomographic image containing part of

TABLE 2
Factors Affecting the Recovery Coefficient

Imaging system parameters
 In-plane resolution
 Z-axis resolution
 Reconstruction filter
 Object size
 Object geometry
 Region of interest
 Size
 Shape
 Placement
 Average or maximum value

one anatomic structure and part of another, mixed together, so that they are indistinguishable. The clearest example is a tomographic section of finite thickness, say 1 cm, that is positioned so that the top 5 mm of a tumor extends into the section, while the other 5 mm of thickness contains normal tissue. In the final image, the tumor is seen but with reduced counts due to the admixture of normal background. Partial volume can affect a single pixel or many.

Table 2 lists factors that affect the recovery coefficient and influence the magnitude of partial volume effects in most cases. Hoffman et al. (10) addressed these problems in detail and showed the strong interdependence of system resolution, reconstruction filter and object size. In general, objects smaller than twice the resolution of the imaging system will show recovery coefficients substantially less than one. For example, Hoffman et al. show that the recovery coefficient was only 50% for a cylinder equal to the FWHM of their imaging system. Kessler et al. (11) have shown that three-dimensional recovery coefficients can be even more severely affected. For a sphere diameter equal to the FWHM of their imaging system, the recovery coefficient was only 31.6%.

The work of both Hoffman et al. (10) and Kessler et al. (11) indicate a strong dependence of the recovery coefficient on object geometry. Kessler et al., for example, note that “. . . the dimension necessary for full recovery is larger for cylinders than for slabs and larger still for spheres.” This dependence on object geometry is emphasized in real structures by the work of Mazziotta et al. (12) who note, “Partial volume effects . . . were largest for small, thin, irregularly shaped objects whose pixel values were most different from neighboring structures.”

These recovery coefficient effects have been recognized by a number of investigators and attempts have been made to correct for this effect. Usually, such corrections are made by empirically measuring the recovery coefficient for various size spheres in a phantom and applying these corrections to actual patient data. The problem with such corrections is that they require accurate estimates of lesion size to know what correction to apply and, more importantly, they always fail to consider the object geometry dependence of the recovery coefficient.

Region of Interest Effects

In addition to these dramatic effects, region of interest (ROI) effects must be considered. The size, shape and placement of the ROI are all important, particularly if the average counts within the ROI are used as the measurement value. Placing a large ROI around an object and averaging the counts within the region will change the measured value because counts from the edge of the object are reduced because counts from surrounding tissues will be included, as emphasized by Kessler et al. (11).

The recovery coefficients measured by Hoffman et al. (10) and Kessler et al. (11) are based on the maximum pixel value within an ROI. This is the truest measure of the actual activity within the region. If average values are used, as is very common in calculating SUVs, the distortions introduced can be significant. For example, Kessler et al. (11) showed that if the average value within a region equal in diameter to the apparent object size is used, as opposed to the single highest pixel, for a sphere less than $2.7 \times \text{FWHM}$, there is a loss of 24% in apparent activity. The effect of ROI size was also studied by Kuwert et al. (13) who found that increasing the ROI size from 2 to 20 mm led to a 66% decrease in the apparent glucose metabolic rate in the caudate nucleus. They also emphasize the use of peak values within the ROI as the best method.

Given these limitations, it is reasonable to conclude that significant errors occur in most of the published SUV measurements, even if corrections for recovery coefficient losses have been made. Particular skepticism should be reserved for measurements based on ROI averaged SUVs.

As noted, recovery coefficient and partial volume effects are related but distinctly different. Partial volume is affected by changes in system resolution and is perhaps most affected by changes in slice thickness and spacing. Miller et al. (14) discussed changes in the axial recovery coefficient with changes in slice spacing and axial resolution. They found that "Recovery coefficient varies with the position of the object in relation to the slice. If the offset of the object from the slice center is unknown the recovery coefficient has an associated uncertainty." In effect, these authors are saying that even if you measure everything correctly, you still cannot be sure of the exact correction to apply, particularly for small objects.

IS THERE A ROLE FOR SUV IN PET?

As one reads the published reports of studies using the SUV, it becomes obvious that most investigators have either ignored or overlooked these error sources. Quite frequently, no mention is made of one or more of these factors, so that it is impossible to determine what was done. Common examples of problems include failure to standardize measurement times after injection, failure to correct for body habitus, use of averaged ROI values rather than maximum values and failure to correct for plasma glucose levels. I have yet to see an article in which all of these variables are appropriately described and

treated. Even the papers dealing specifically with some of these problems fail to discuss and manage some of the other factors.

At a meeting where I recently presented some of this material, an irate member of the audience rose during the comment session to state that this was wrong, that quantitation in the form of the SUV was vitally important in the marginal cases to separate benign from malignant lesions. It was clear that he had missed the point. These are not subtle effects, but factors that can cause potential errors of 50% or more. No one would question that an SUV of 11 or even 8 represents malignancy. Such cases rarely represent a diagnostic question after subjective visual interpretation. What is the significance, however, of an uncorrected SUV of 2.8 in an obese diabetic patient with a 1-cm lesion? It is precisely these marginal situations that the value becomes too uncertain to be useful.

Is there a glimmer of hope for the SUV? I believe yes. As in many situations in which true quantification is difficult, a measure of improvement can be achieved by using serial measurements with the individual patient serving as their own control. For example, the SUV of an individual tumor nodule could be measured before and after therapy and any change used as an index of therapeutic response. If careful attention is paid to proper technique this approach *should* work.

What kind of attention is needed? First, a lesion should be chosen that is large enough to obviate recovery coefficient errors and partial volume effects. For most modern PET scanners, this probably means lesions no smaller than 2 cm. When assessing the counts in the lesion, the ROI should be placed to encompass the entire lesion and the maximum pixel value, not the average, within the ROI should be used as the measurement value. The time from injection to measurement must be held constant to eliminate or at least minimize the time dependency of the SUV. Finally, appropriate corrections for body habitus and plasma glucose should be applied. Given these caveats the SUV *may* provide a reliable index of changes in tumoral FDG uptake over time in any individual patient.

It is worth noting that many of these problems, particularly those relating to recovery coefficients and partial volume effects apply to all attempts at quantification, not just measurements of the SUV. Absolute determinations of tracer uptake, be it FDG or a specific neuroreceptor ligand, should be viewed with a close and critical eye. Where structures are small and the anatomic geometry is complex, as in the brain, it is doubtful if any of the published quantitative data that we have seen can really be considered to represent "truth."

Otherwise, what conclusions can we draw? First, and most importantly, most of the currently published data on SUVs in tumors are of little or no value to investigators outside the laboratory where the investigation was conducted. Second, our journal editors and reviewers need to be much more critical of submitted quantitative results. Minimally, they should require full methodological de-

scriptions of how the results were achieved at a level of detail sufficient for another laboratory to duplicate the experiment. Where obvious flaws in methodology are obvious, such as widely varying times from injection to measurement, results should be rejected or a full discussion of potential errors should be required.

CONCLUSION

As currently applied, the SUV is, in fact, a “silly useless value” and its continued application as a quantitative index for malignancy per se should be discouraged. It is interesting to note in follow-up to this recommendation that Lowe et al. (15) found no difference in the accuracy of separating benign from malignant pulmonary nodules when they compared visual interpretation with the SUR (same as SUV), which harkens back to the pleas of DiChiro and Brooks (1).

ACKNOWLEDGMENTS

The author is indebted to J. Gordon Chan at The Austin Hospital in Heidelberg, Australia for pointing out to him the true meaning behind the initials SUV.

This work was presented in part at the Pre-Congress Symposium “PET and SPECT: Controversies and Future Directions,” Cairns, Australia, October 19–20, 1994.

REFERENCES

1. DiChiro G, Brooks RA. PET quantitation: blessing and curse. *J Nucl Med* 1988;29:1603–1604.
2. 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.
3. Minn H, Joensuu H, Ahonen A, Klein P. Fluorodeoxyglucose imaging: a method to assess the proliferative activity of human cancer in vivo: comparison with DNA flow cytometry in head and neck tumors. *Cancer* 1988; 61:1776–1781.
4. Hamburg LM, Hunter GT, Alpert NM, et al. The dose uptake ratio as an index of glucose metabolism: useful parameter or over simplification? *J Nucl Med* 1994;35:1308–1312.
5. Strauss LG, Clorius JH, Schlag P, et al. Recurrence of colorectal tumors: PET evaluation. *Radiology* 1989;170:329–332.
6. Langen K-J, Braun U, Kops ER, et al. The influence of plasma glucose levels on fluorine-18-fluorodeoxyglucose uptake in bronchial carcinomas. *J Nucl Med* 1993;34:355–359.
7. 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.
8. Lindholm P, Minn H, Leskinen-Kallio S, et al. Influence of the blood glucose concentration on FDG uptake in cancer—a PET study. *J Nucl Med* 1993;34:1–6.
9. Ishizu K, Sadato N, Yonekura Y, et al. Enhanced detection of brain tumors by [¹⁸F]fluorodeoxyglucose PET with glucose loading. *J Comput Assist Tomogr* 1994;18:12–15.
10. Hoffman EJ, Huang S-C, Phelps ME. Quantitation in positron emission computed tomography: I. Effect of object size. *J Comput Assist Tomogr* 1979;3:299–308.
11. Kessler RM, Ellis JR, Eden M. Analysis of emission tomographic scan data: limitations imposed by resolution and background. *J Comput Assist Tomogr* 1984;8:514–522.
12. Mazziotta JC, Phelps ME, Plummer D, Kuhl DE. Quantitation in positron emission computed tomography: V. Physical-anatomical effects. *J Comput Assist Tomogr* 1981;5:734–743.
13. Kuwert T, Ganslandt T, Jansen P, et al. Influence of size of regions of interest on PET evaluation of caudate glucose consumption. *J Comput Assist Tomogr* 1992;16:789–794.
14. Miller TR, Wallis JW, Grothe Jr RA. Design and use of PET tomographs: the effect of slice and spacing. *J Nucl Med* 1990;31:1732–1739.
15. Lowe VJ, Hoffman JM, DeLong DM, et al. Semiquantitative and visual analysis of FDG-PET images in pulmonary abnormalities. *J Nucl Med* 1994;35:1771–1776.