Intra- and Interobserver Agreement of SUV SPECT Quantitative SPECT/CT Processing Software, Applied in Clinical Settings for Patients with Solid Renal Tumors

Antonios Tzortzakakis[1,2], Maria Holstensson[3], Eva Hagel[4], Mattias Karlsson[3], and Rimma Axelsson[1,3]

[1]Division of Radiology, Department for Clinical Science, Intervention, and Technology, Karolinska Institutet and Medical Physics and Nuclear Medicine, Karolinska University Hospital, Huddinge, Sweden; [2]Division of Medical Imaging and Technology, Department of Neuroradiology, Karolinska University Hospital, Huddinge, Sweden; [3]Medical Radiation Physics and Nuclear Medicine, Functional Unit of Nuclear Medicine, Karolinska University Hospital, Huddinge, Sweden; and [4]Learning, Informatics, Management, and Ethics, Medical Statistics Unit, Karolinska Institutet, Stockholm, Sweden

Quantification tools for SPECT/CT are a field of ongoing research among hybrid imaging techniques. A recent multicenter study evaluating SUV SPECT software on phantoms showed that quantitative SPECT/CT is reproducible in Hybrid Viewer software. The aim of our study was to evaluate the intra- and interobserver agreement of quantitative SUV SPECT measurements in clinical settings for patients with solid renal tumors. **Methods:** The evaluation was part of a study that examined the role of ^{99m}Tc-sestamibi SPECT/CT in the characterization of solid kidney tumors and the differentiation of renal oncocytomas from renal cell carcinomas. Quantitative evaluation of SUV measurements was performed in Hybrid Viewer PDR, version 2.5. Forty-eight renal lesions were identified and examined twice by 2 independent readers. The agreement of the SUV measurements, concerning renal tumors and adjacent renal parenchyma, was evaluated using the intraclass correlation coefficient (ICC). **Results:** ICC for SUV_{max} measurements by the same reader was 97%–99% for solid renal tumors and 92%–98% for ipsilateral healthy renal parenchyma. ICC for SUV_{max} measurements between readers was 87%–89% for solid renal tumors and 72%–73% for ipsilateral healthy renal parenchyma. Estimated ICC for SUV_{mean} measurements of solid renal tumors was 95%–98% for the same reader and 86%–89% between readers. Similar results were found for SUV_{peak} measurements. **Conclusion:** The high ICCs indicate a strong agreement among SUV measurements for patients with solid renal lesions undergoing ^{99m}Tc-sestamibi SPECT/CT, not only for measurements by the same reader but also for measurements between 2 different readers.

Key Words: SUV SPECT; SPECT/CT; ^{99m}Tc-sestamibi SPECT/ CT; quantitative SPECT/CT; SUV SPECT/CT

For many years, SPECT was undervalued in comparison to PET (*1*), not only because of the higher spatial resolution and sensitivity of PET but also because PET/CT was enabling quantification of tissue radioactivity concentrations (*2*). However, quantitative SPECT/CT is used routinely in some applications, such as targeted radionuclide therapy imaging (*3*). The demand for quantitative SPECT for diagnostic purposes is an old one (*4*). The introduction of SPECT/ CT cameras and continued reduction of the absorbed dose from the CT component of this hybrid imaging method have revived interest in quantitative SPECT (*5*). This hybrid imaging technique is increasing the diagnostic accuracy of various nuclear medicine examinations in daily clinical practice (*6*). The introduction of new reconstruction methods with proper modeling of photon attenuation, scatter, and collimator resolution degradation has resulted in better image quality and improved quantification. Quantification with SPECT/CT is especially of interest when the use of longer–half-life radionuclides not available in PET/CT is required (*7*).

Recently, a commercial software product called SUV SPECT (Hermes Medical Solutions) was developed and tested in a multicenter study on phantoms (*8*). The study showed that quantitative SPECT is reproducible in this setting. On that basis, our institution conducted the MIDOR study (molecular imaging in differentiation of renal oncocytomas from renal cell carcinoma [RCC] by means of ^{99m}Tc-sestamibi SPECT/CT) on patients (*9*). ^{99m}Tc-sestamibi acts as a mitochondria-imaging agent that can trace possible renal oncocytomas, since such benign solid kidney tumors contain an increased number of mitochondria. The visual evaluation was of high diagnostic yield, with 11 of 12 correctly identified renal oncocytomas showing ^{99m}Tc-sestamibi uptake. However, the presence of false-positive results, namely RCC exhibiting ^{99m}Tc-sestamibi uptake, forced our group to explore whether quantitative tools could contribute to a more accurate characterization of the solid renal masses.

The MIDOR study is a part of a strategic innovation program founded by the Swedish Government that supports collaboration within academia, industry, and health care. Participation in this program gave our group access to SUV SPECT software designed for quantification of SPECT/CT studies. Other published studies using the SUV SPECT software demonstrated prominent results supporting its quantitative role for an accurate diagnostic approach in the evaluation of myocardial viability, in the absolute quantification of lumbar herniation, and in the evaluation of therapy response in prostate cancer patients with bone metastasis (10–13).

The aim of our present study was to evaluate the intra- and interobserver agreement of quantitative SUV SPECT measurements in clinical settings for patients with solid renal tumors undergoing ^{99m}Tc-sestamibi SPECT/CT. This is the first paper from our group that evaluates the consistency of SUV SPECT measurements in patients with primary solid kidney tumors.

MATERIALS AND METHODS

The Regional Ethical Review Board and the local Radiation Safety Committee approved the MIDOR study (*diarienummer* 2015/923-31/4). Written informed consent was acquired from all patients who participated in our study.

Forty-one patients with 48 solid renal tumors were included in this evaluation between September 2015 and May 2017. Before any invasive procedures (renal biopsy or surgical excision), all study participants underwent ^{99m}Tc-sestamibi SPECT/CT. Patients were injected intravenously with 925 ± 25 MBq of ^{99m}Tc-sestamibi (produced by the National Centre for Nuclear Research, Poland; distributed by S. Ahlén Medical Nordic AB). SPECT/CT imaging was performed 60–90 min after the injection using a dual-head SPECT/CT device (Symbia T16; Siemens) equipped with low-energy high-resolution collimators.

SPECT imaging was performed using a 128 × 128 pixel matrix (zoom factor, 1) and 32 projections per detector head. Each projection was acquired for 40 s. Directly after the SPECT acquisition, CT was performed using 130-kV tube voltage, a 5-mm slice width, and the automatic exposure control CARE Dose4D activated to provide proper tube current modulation, with a quality reference milliampere-second setting of 10. The projections were reconstructed using the SUV SPECT Hybrid Recon Oncology software (version 1.2; Hermes). The iterative ordered-subsets expectation-maximization algorithm (6 iterations, 16 subsets) was used, applying corrections for scatter, attenuation, and collimator resolution followed by postfiltration using a gaussian 3-dimensional filter (8 mm in full width at half maximum).

To enable quantitative evaluation, information about syringe activity, residual activity in the syringe, patient weight, time points of injection, and scan start time was specified in the reconstruction software. A calibration factor for the SPECT system was determined beforehand with a known amount of ^{99m}Tc in a cylindric phantom, which enabled conversion from pixel count rates to activity concentration and SUV normalized to patient body weight (decay-corrected tissue concentration [kBq/mL] divided by the dose injected per body weight [kBq/g]) (14).

Forty-eight solid renal lesions (6 chromophobe RCCs, 9 papillary RCCs, 12 clear RCCs, 12 oncocytomas, 3 hybrid oncocytic chromophobe renal tumors, 1 lymphoma, 1 metanephric adenoma, 1 angiomyolipoma, 1 chromophobe-papillary RCC, and 2 clear-papillary RCCs) were identified on the multiphase CT or MRI scans that patients had previously undergone for diagnostic purposes and detailed anatomic mapping. All images (diagnostic CT or MRI and the ^{99m}Tc-sestamibi SPECT/CT) were evaluated twice by 2 independent readers, and measurements were performed within a 1-wk interval in Hybrid Viewer PDR (version 2.5; Hermes) (Fig. 1).

The 2 readers had different professional backgrounds. One of the readers was a specialist in nuclear medicine dealing with scintigraphy studies in everyday clinical practice. The other reader was a consultant radiologist with experience in oncologic imaging. SUV measurements were performed on the solid renal tumors and the adjacent healthy renal parenchyma. The 2 readers drew regions of interest (ROIs) freehand at the site of the renal tumor on the axial fused SPECT/CT images, often guided by the diagnostic CT or MRI that each patient underwent previously for tumor staging. Freehand ROIs were also obtained from the ipsilateral healthy kidney parenchyma on the axial fused SPECT/CT images. The software could then generate volumes of interest based on the manually drawn ROIs. The 2 readers drew the ROIs for the tumors and the healthy tissues independently. No minimum or maximum number of ROIs was proposed. Additionally, the readers freely selected the regions of healthy renal parenchyma to be measured. At least 4 ROIs were obtained from the examined renal tumors and ipsilateral healthy kidney parenchyma. The software then extracted SUV$_{max}$, SUV$_{mean}$, and SUV$_{peak}$ for renal tumors and healthy renal parenchyma, according to the definitions given by Wahl et al. (15).

The level of agreement among the measurements was evaluated using the intraclass correlation coefficient (ICC) (16). To estimate ICC, a 2-way random single-measurement absolute-agreement approach was used. The purpose of the ICC was to estimate the amount of variability due to measurement uncertainty relative to the total variability in the samples. The estimate was given with a 95% confidence interval. ICC ranges from 0% to 100%, where absolute agreement between readers is 100%. The values of ICC depend heavily on the heterogeneity of the study population and the field of research; hence, there is little consensus in the statistical community on definite cutoffs for ICC or even on what is considered strong agreement.

The appearance, location, size, ^{99m}Tc-sestamibi uptake intensity, and nature (benign or malignant) of the kidney tumors were highly variable. These tumor features

FIGURE 1. Female patient with 15-mm solid neoplasm (arrows) on lateral aspect of left kidney. Scintigraphy coronal view (left) demonstrates absence of ^{99m}Tc-sestamibi in place of renal tumor. Directly after SPECT acquisition, CT scan (right) was performed for anatomic correlation. Middle panel is fused coronal SPECT/CT image. Renal tumor was diagnosed as clear cell carcinoma on histopathologic grounds.
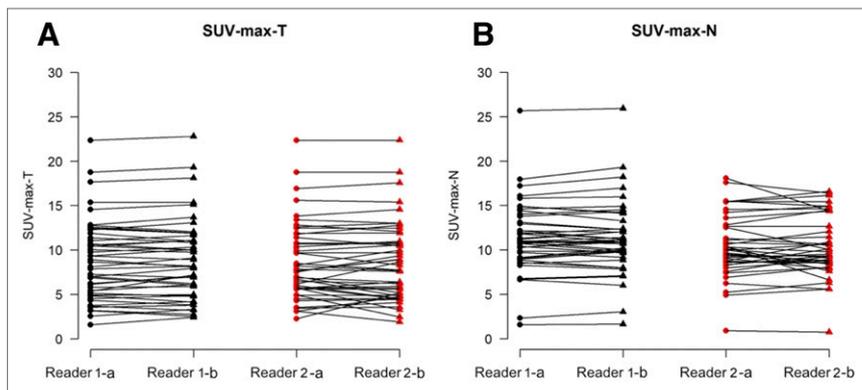
FIGURE 2. (A) All individual SUV$_{max}$ measurements in solid renal tumors for both readers. (B) All individual SUV$_{max}$ measurements in ipsilateral renal parenchyma. T = renal tumor; N = healthy renal parenchyma.

were not examined or correlated with the SUV measurements since this study examined only the level of agreement (high or low) among the different SUV SPECT measurements.

RESULTS

Figure 2 demonstrates in a purely descriptive way the individual values of SUV$_{max}$ for the solid renal tumors and the healthy renal parenchyma measured twice by the 2 readers.

Subsequent statistical analysis showed the same tendencies of agreement among different measurements as are presented descriptively in Figure 2. Tables 1 and 2 summarize the ICCs and their upper and lower margins derived from SUV SPECT measurements performed by the 2 readers on solid renal tumors and adjacent healthy renal parenchyma.

As Table 1 shows, ICC for SUV$_{mean}$ measurements of solid renal tumors by the same reader varied between 95% and 98%. ICC for SUV$_{mean}$ measurements of ipsilateral healthy renal parenchyma by the same reader varied between 93% and 98%. Table 2 shows that ICC between readers for SUV$_{mean}$ measurements of solid renal tumors varied between 86% and 89%. ICC for SUV$_{mean}$ measurements of ipsilateral healthy renal parenchyma between readers was 73%. In the same tables, ICCs for SUV$_{peak}$ and SUV$_{max}$ by the same reader and between readers are listed.

DISCUSSION

Regarding SUV SPECT measurements of renal tumors between readers, the estimated ICCs of 85%–90% indicate

that 85%–90% of the observed variability was due to actual variability among the solid renal tumors and only 10%–15% was due to variability between readers' measurements. This is considered high in many fields of research, although more information within this specific field would be needed for comparison.

SPECT/CT with ^{99m}Tc-sestamibi has been used for diagnosis of renal oncocytomas by another research group, from John Hopkins Hospital, under the guidance of Steven P. Rowe and Michael A. Gorin. This group developed its own quantitative SPECT/CT (QSPECT) reconstruction models for quantification. In their evaluations, any ICC above 80% was considered a good correlation (*17*). Our group also agrees that a 10%–15% measurement error is acceptable in everyday clinical practice (*18*). The ICCs for the same reader concerning SUV SPECT measurements of the kidney lesions were even higher, with values around 97%.

Apart from the well-known intrinsic limitations of SPECT/CT imaging (*1–4*), further limitations were observed in this study. The different backgrounds of the readers can affect how easily the renal tumor is detected on the CT or the MRI scan and subsequently how the ROIs are drawn on the fused SPECT/CT images. The reader participating in this study with the radiologic profile generated more consistent measurements.

Since some of the examined renal tumors were very small, difficulties in obtaining SUV$_{peak}$ were often encountered. By definition, SUV$_{peak}$ represents the maximum activity concentration in a 1-cm^3 volume of a larger volume of interest drawn. This explains the difficulties in measuring SUV$_{peak}$ in tumors smaller than 15 mm. The readers had to be precise when drawing ROIs on those small lesions, to include the whole tumor mass in the measurement. The last-mentioned difficulty did not seem to affect the ICCs compared with those of SUV$_{mean}$ or SUV$_{max}$ measurements. ICCs for SUV$_{peak}$ between readers were actually higher than ICCs for SUV$_{mean}$ or SUV$_{max}$ measurements.

Bigger renal tumors often exhibit regions of cystic degeneration or necrosis that are difficult to assess. The readers

TABLE 1
ICCs for SUV$_{mean}$, SUV$_{max}$, and SUV$_{peak}$ for Repeated Measurements by Same Reader

Same reader	SUV$_{mean}$	SUV$_{peak}$	SUV$_{max}$
Reader 1: T	0.984 (0.972–0.991)	0.989 (0.980–0.994)	0.989 (0.981–0.994)
Reader 2: T	0.954 (0.920–0.974)	0.969 (0.946–0.982)	0.965 (0.939–0.980)
Reader 1: N	0.975 (0.956–0.986)	0.982 (0.968–0.989)	0.983 (0.970–0.990)
Reader 2: N	0.932 (0.883–0.961)	0.932 (0.884–0.961)	0.921 (0.865–0.954)

T = renal tumor; N = healthy renal parenchyma.
Data are ICC followed by 95% confidence interval in parentheses.

TABLE 2
ICCs for SUV_{mean}, SUV_{max}, and SUV_{peak} for Repeated Measurements Between Readers

Between readers	SUV_{mean}	SUV_{peak}	SUV_{max}
First measurement - T	0.890 (0.814–0.936)	0.887 (0.809–0.934)	0.866 (0.775–0.922)
First measurement - N	0.732 (0.530–0.848)	0.705 (0.526–0.823)	0.715 (0.518–0.835)
Second measurement - T	0.858 (0.762–0.917)	0.915 (0.856–0.951)	0.890 (0.812–0.936)
Second measurement - N	0.734 (0.560–0.843)	0.656 (0.447–0.794)	0.729 (0.524–0.847)

T = renal tumor; N = healthy renal parenchyma.
Data are ICC followed by 95% confidence interval in parentheses.

measured the solid components of these tumors that exhibited higher ^{99m}Tc-sestamibi uptake than the necrotic parts. A similar strategy can be adopted for SUV measurements of healthy renal parenchyma. It is easier for readers to perform SUV measurements on the regions of healthy parenchyma with the highest ^{99m}Tc-sestamibi uptake; in those cases, that uneven uptake was observed (Fig. 3). In many cases, though, the ^{99m}Tc-sestamibi uptake was equally distributed in the healthy renal parenchyma.

Our measurements were done unsystematically on the healthy renal parenchyma of the ipsilateral kidney, partly explaining the low ICCs (~70%) obtained between the 2 readers, compared with the very high ICCs from the renal tumors, which are by definition more focal and often well-defined lesions. The ICCs for the healthy renal parenchyma in the same reader were very high, however, at around 95%.

The different nature of the solid renal tumors, namely the higher mitochondrial content of renal oncocytomas, hybrid oncocytic chromophobe renal tumor, and some variants of chromophobe RCC, clearly affects the uptake and distribution of the ^{99m}Tc-sestamibi, indicating the regions of maximal ^{99m}Tc-sestamibi uptake [19]. Future clinical applications and SUV SPECT measurements should be performed on those regions exhibiting the highest ^{99m}Tc-sestamibi uptake since they are easily identified visually by any reader. No further analysis based on the nature (benign or malignant) of the examined solid renal lesions or correlation with the visual evaluation of ^{99m}Tc-sestamibi uptake was performed in the present study.

CONCLUSION

The ICCs of around 70% obtained between the 2 readers concerning measurements on healthy renal parenchyma suggest moderate agreement, a finding that is partly explained by the absence of predefined criteria about the location of SUV measurements. Regions of high radioisotope uptake on healthy renal parenchyma should be preferred for future measurements since they are easily detected by any reader. At the same time, the examination of solid renal lesions in this study with quantitative SUV SPECT showed high ICCs of around 90% not only for measurements by the same reader but also for measurements between 2 different readers. Those high ICCs indicate strong agreement among SUV measurements for patients with solid renal lesions undergoing ^{99m}Tc-sestamibi SPECT/CT.

FIGURE 3. Coronal ^{99m}Tc-sestamibi SPECT/CT fused image of bilaterally healthy kidney parenchyma, with regions of uneven uptake in left kidney and more even uptake in right kidney.

REFERENCES

1. Beyer T, Veit-Haibach P. State-of-the-art SPECT/CT: technology, methodology and applications—defining a new role for an undervalued multimodality imaging technique. *Eur J Nucl Med Mol Imaging.* 2014;41(suppl 1):S1–S2.
2. Pfeifer A, Knigge U, Mortensen J, et al. Clinical PET of neuroendocrine tumors using ^{64}Cu- DOTATATE: first-in-humans study. *J Nucl Med.* 2012;53:1207–1215.
3. D'Asseler Y. Advances in SPECT imaging with respect to radionuclide therapy. *Q J Nucl Med Mol Imaging.* 2009;53:343–347.
4. Cachovan M, Vija AH, Hornegger J, et al. Quantification of ^{99m}Tc-DPD concentration in the lumbar spine with SPECT/CT. *EJNMMI Res.* 2013;3:45.
5. Bailey DL, Willowson KP. Quantitative SPECT/CT: SPECT joins PET as a quantitative imaging modality. *Eur J Nucl Med Mol Imaging.* 2014;41(suppl 1): S17–S25.
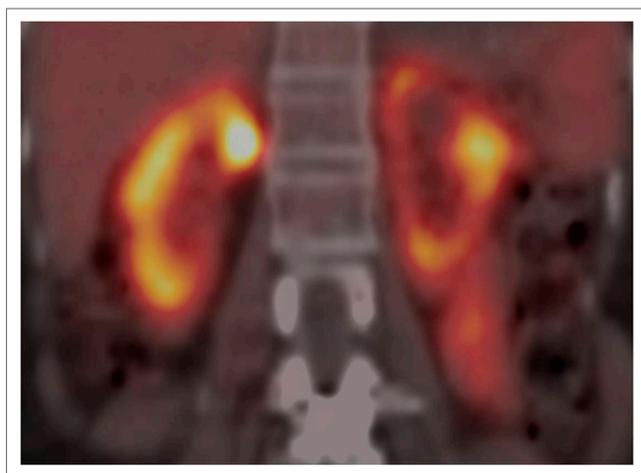
6. Ahmadzadehfar H, Biersack HJ. *Clinical Applications of SPECT-CT*. Berlin, Germany: Springer-Verlag. 2014.

7. Ljungberg M, Pretorius PH. SPECT/CT: an update on technological developments and clinical applications. *Br J Radiol.* 2018;91:20160402.

8. Kangasmaa TS, Constable C, Hippeläinen E, et al. Multicenter evaluation of single-photon emission computed tomography quantification with third-party reconstruction software. *Nucl Med Commun.* 2016;37:983–987.

9. Tzortzakakis A, Gustafsson O, Karlsson M, et al. Visual evaluation and differentiation of renal oncocytomas from renal cell carcinomas by means of [99mTc]-sestamibi SPECT/CT. *EJNMMI Res.* 2017;7:29.

10. He W, Yu W, Hacker M, Sohlberg A, Bertling J, Bertling S. The clinical application of absolute quantification SUV SPECT [99mTc]-MIBI myocardial perfusion imaging to translate myocardial viability—a pilot study [abstract]. *J Nucl Med.* 2015;56(suppl 1):1538.

11. He W, Zhai W, Yu W, et al. A novel SUV-based quantification of [99mTc]-MDP SPECT/CT uptake and [18F]-FDG PET/CT metabolism in patients with lumbar disc herniation [abstract]. *J Nucl Med.* 2015;56(suppl 3):1774.

12. Mao JL, Zhai W, Lewis R, et al. The clinical value of SUV measurements in [99mTC]-MDP SPECT/CT bone scans in the evaluation of therapy response in prostate cancer patients with bone metastases [abstract]. *J Nucl Med.* 2015;56(suppl 3):1789

13. Bailey DL, Willowson KP. An evidence-based review of quantitative SPECT imaging and potential clinical applications. *J Nucl Med.* 2013;54:83–89.

14. Sugawara Y, Zasadny KR, Neuhoff AW, et al. Reevaluation of the standardized uptake value for FDG: variations with body weight and methods for correction. *Radiology.* 1999;213:521–525.

15. Wahl RL, Jacene H, Kasamon Y, et al. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med.* 2009;50(suppl 1):122S–150S.

16. Hallgren KA. Computing inter-rater reliability for observational data: An overview and tutorial university of New Mexico, Department of Psychology. *Tutor Quant Methods Psychol.* 2012;8:23–34.

17. Jones KM, Solnes LB, Rowe SP, et al. Use of quantitative SPECT/CT reconstruction in [99mTc]-sestamibi imaging of patients with renal masses. *Ann Nucl Med.* 2018;32:87–93.

18. Lodge MA. Repeatability of SUV in oncologic [18F]-FDG PET. *J Nucl Med.* 2017;58:523–532.

19. Rowe SP, Gorin MA, Solnes LB, et al. Correlation of [99mTc]-sestamibi uptake in renal masses with mitochondrial content and multi-drug resistance pump expression. *EJNMMI Res.* 2017;7:80.