Intra- and interobserver agreement of SUV SPECT® quantitative SPECT/CT processing software, applied in clinical settings for patients with solid renal tumours

Short running title: SUV SPECT® quantitative SPECT/CT

Authors:
Antonios Tzortzakakis¹,², Maria Holstensson³, Eva Hagel⁴, Mattias Karlsson³, Rimma Axelsson¹³

¹ Division of Radiology, Department for Clinical Science, Intervention and Technology (CLINTEC), 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
⁴ Learning, Informatics, Management and Ethics (LIME), Medical Statistics Unit, Karolinska Institutet

Corresponding author:
Antonios Tzortzakakis MD
C1-46 Karolinska University Hospital
SE-141 86 Huddinge, Stockholm, Sweden

E-mail: antonios.tzortzakakis@ki.se
Phone number: 0046702219919
Fax number: 004687114840
Aim: Quantification tools for Single Photon Emission Computed Tomography/Computed Tomography (SPECT/CT) is a field of ongoing research among hybrid imaging techniques. A recent multicentre study performed in phantoms evaluating Standardized Uptake Value (SUV) SPECT® showed that quantitative SPECT/CT is reproducible in HERMES Hybrid Viewer software. The aim of our study is to evaluate the intra- and interobserver agreement of quantitative SUV SPECT® measurements in clinical settings for patients with solid renal tumours.

Methods: The evaluation was part of a study which examines the role of $^{99m}$Tc-Sestamibi SPECT/CT in the characterization of solid kidney tumours and the differentiation of renal oncocytopas from renal cell carcinomas (RCC). Quantitative evaluation of SUV measurements was performed in HERMES Hybrid Viewer PDR v2.5. Forty-eight renal lesions were identified and examined twice by two independent readers. The agreement of the SUV measurements, concerning the renal tumours and the adjacent renal parenchyma, was evaluated using the intraclass correlation coefficient (ICC).

Results: ICC concerning the SUV$_{\text{max}}$ measurements of solid renal tumours in the same reader varies between 97%-99%. The ICC of SUV$_{\text{max}}$ measurements on the ipsilateral healthy renal parenchyma from the same reader varies between 92%-98%. ICC between readers concerning SUV$_{\text{max}}$ measurements of the renal tumours varies between 87%-89%. The ICC of SUV$_{\text{max}}$ measurements on the ipsilateral healthy renal parenchyma between readers was between 72%-73%. Estimated ICC concerning SUV$_{\text{mean}}$ values, for measurements on the solid renal tumour in the same reader varies between 95%-98%. ICC values of SUV$_{\text{mean}}$ measurements between the two readers concerning the renal tumours, varies between 86%-89%. Similar results were found for the SUV$_{\text{peak}}$ measurements.

Conclusion: The high ICC values indicate a strong agreement among SUV measurements for patients with solid renal lesions undergoing a $^{99m}$Tc-Sestamibi SPECT/CT examination not only in measurements of the same reader but also between two different readers.

Keywords: SUV SPECT®, SPECT/CT, $^{99m}$Tc-Sestamibi SPECT/CT, Quantitative SPECT/CT, SUV SPECT/CT
Introduction

For many years SPECT was undervalued in comparison to positron emission tomography (PET) (1). Not only because of the higher spatial resolution and sensitivity of PET but also due to the fact that PET/CT was enabling quantification of tissue radioactivity concentrations (2). Quantitative SPECT/CT is used routinely in some applications, however, for example in 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 has revived an 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 modelling of photon attenuation, scatter and collimator resolution degradation have 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, SUV SPECT® software (Hermes Medical Solutions, Stockholm Sweden) was developed and tested in a multicentre study performed in phantoms (8). The study showed that quantitative SPECT is reproducible in this setting. Based on that, 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) in patients (9). $^{99m}$Tc Sestamibi acts as a mitochondrial-imaging agent that can trace possible renal oncocytomas, since such benign solid kidney tumours contain an increased number of mitochondria. The visual evaluation was of high diagnostic yield, with 11 out 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 healthcare. Participation in this program gave to our group access to SUV SPECT® software designed for quantification of SPECT/CT-studies. Other published studies employing the SUV SPECT® software, demonstrate prominent results supporting its quantitative role for an accurate diagnostic
approach in evaluation of the myocardial viability, the absolute quantification of lumbar herniation and in the
evaluation of therapy response in prostate cancer patients with bone metastasis (10,11,12,13).

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

Material and Methods
The Regional Ethical Review Board and the local Radiation Safety Committee approved the MIDOR study.
Forty-one patients with 48 solid renal tumours were included in this evaluation between September 2015 and May
2017. Prior to any invasive procedures (renal biopsy or surgical excision) all study participants underwent a $^{99m}$Tc-
Sestamibi SPECT/CT examination. Patients were injected intravenously with 925 ± 25 MBq $^{99m}$Tc-Sestamibi
(produced by the National Centre for Nuclear Research, Poland; distributed by S. Ahlén Medical Nordic AB,
Stockholm, Sweden). SPECT/CT imaging was performed 60-90 minutes after the injection using a dual headed
SPECT/CT (Symbia T16, Siemens, Erlangen, Germany) equipped with low-energy high-resolution collimators.

SPECT imaging was performed using a 128×128-pixel matrix size (zoom factor 1) and 32 projections per detector
head. Each projection was acquired for 40 seconds. Directly following the SPECT acquisition, a CT was performed
using 130 kV tube voltage, 5 mm slice width and the automatic exposure control CARE Dose4D activated to provide
proper tube current modulation, with a quality reference mAs setting of 10. The projections were reconstructed using
the Hermes SUV SPECT Hybrid Recon™ Oncology software (v.1.2). The iterative OSEM algorithm (6 iterations,
16 subsets) was used, applying corrections for scatter, attenuation and collimator resolution followed by post-
filtration using a Gaussian 3D-filter (8 mm full-width at half-maximum).

In order to enable quantitative evaluation, information about syringe activity, residual activity in syringe, patient
weight, time points of injection and scan start were specified in the reconstruction software. A calibration factor for
the SPECT system was determined beforehand with a known amount of $^{99m}$Tc in a cylindrical phantom, which
enables conversion from pixel count rate values to activity concentration and standardized uptake value normalized to patient body weight (SUVw). SUVw was defined as the decay-corrected tissue concentration (in kilobecquerels/millilitre) divided by the dose injected per body weight (in kilobecquerels/gram) (14).

Forty-eight solid renal lesions (6 chromophobe RCC, 9 papillary RCC, 12 clear RCC, 12 oncocyomas, 3 hybrid oncocyomic chromophobe renal tumour HOCT, 1 lymphoma, 1 adult Wilms’ tumour, 1 angiomyolipoma, 1 chromophobe-papillary RCC and 2 clear-papillary RCC) were identified on the multiphase CT or magnetic resonance imaging (MRI) scans that patients have previously undergone for diagnostic purposes and detailed anatomical mapping. All images (diagnostic CT or MRI scans and the ⁹⁹mTc-Sestamibi SPECT/CT examinations) were evaluated twice by two independent readers and measurements performed within a one-week interval in HERMES Hybrid Viewer PDR v2.5 (fig.1).

The two readers had different professional backgrounds. One of the readers is a specialist in nuclear medicine dealing with scintigraphy studies in everyday clinical praxis. The other reader is a consultant radiologist with experience in oncological imaging. SUV measurements were performed in the solid renal tumours and the adjacent healthy renal parenchyma. The two readers drew regions of interest (ROI) free hand at the site of the renal tumour on the axial fused SPECT/CT images often guided by the diagnostic CT or MRI examination that each patient underwent previously for tumour staging. Free hand 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 (VOIs) based on the manually drawn ROIs. The two readers drew the ROIs for the tumours and the healthy tissues independently. No minimum or maximum number of ROIs were proposed. Additionally, the readers freely selected the regions of the healthy renal parenchyma to be measured. At least 4 ROIs obtained from the examined renal tumours and the ipsilateral healthy kidney parenchyma. The software then extracted \( \text{SUV}_{\text{max}} \), \( \text{SUV}_{\text{mean}} \) and \( \text{SUV}_{\text{peak}} \) for the renal tumours and the 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). In order to estimate ICC a two-way random single measurement, absolute agreement approach was employed. 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 0-100%, 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 cut-off values of ICC or even what is considered strong agreement.

The appearance, location, size, $^{99}$Tc-Sestamibi uptake intensity as well as the nature (benign or malignant) of the kidney tumours were highly variable. These tumour features 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 $\text{SUV}_{\max}$ of the solid renal tumours and the healthy renal parenchyma respectively measured twice by the two readers.

Subsequent statistical analysis shows the same tendencies of agreement among different measurements presented descriptively in figure 2. Tables 1-2 summarize the ICC values as well as their upper and lower margins derived from SUV SPECT® measurements performed by the 2 readers on the solid renal tumours and the adjacent healthy renal parenchyma.

As Table 1 shows, ICC for the $\text{SUV}_{\text{mean}}$ measurements of solid renal tumour (T) in the same reader varied between 95%-98%. The ICC of $\text{SUV}_{\text{mean}}$ measurements on the ipsilateral healthy renal parenchyma (N) in the same reader varied between 93%-98%. Table 2 shows that ICC between readers for $\text{SUV}_{\text{mean}}$ measurements of the renal tumours varied between 86%-89%. The ICC of $\text{SUV}_{\text{mean}}$ measurements on the ipsilateral healthy renal parenchyma between readers was found to be 73%. At the same tables ICC values for $\text{SUV}_{\text{peak}}$ and $\text{SUV}_{\max}$ by the same reader and between readers are listed.

**Discussion**
The estimated ICC values for SUV SPECT® measurements of the renal tumour between readers that are around 85%-90%, mean that 85%-90% of the observed variability is due to actual variability between individuals and only 10%-15% is due to variability between readers. 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 oncocyotomas by another research group from John Hopkins Hospital, USA under the guidance of S.P Powe and M.A Gorin. The above-mentioned group developed its own quantitative SPECT/CT (QSPECT) reconstruction models for quantification. In their evaluations any ICC value above 80% was considered as good correlation (17). Our group also agrees that a 10%-15% measurement error is acceptable in everyday clinical practice (18). The ICC values on the same reader concerning SUV SPECT® measurements of the kidney lesion is even higher with values around 97%.

Apart from the well-known intrinsic limitations of SPECT/CT imaging (1, 2, 3, 4) further limitations were observed in this study. The different backgrounds of the readers can affect how easily the renal tumour 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 radiological profile generated more consistent measurements.

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

Bigger renal tumours often exhibit regions of cystic degeneration or necrosis that were also difficult to assess. The readers performed measurements in the solid components of these tumours that exhibited higher $^{99m}$Tc-Sestamibi uptake than the necrotic parts. A similar strategy can also be adopted for the SUV measurements of the healthy renal
parenchyma. It is easier for readers to detect and perform their SUV measurements in the regions of the healthy parenchyma with the highest $^{99m}$Tc-Sestamibi uptake, in those cases that uneven uptake is observed (fig. 3). In many cases though the $^{99m}$Tc-Sestamibi uptake was equally distributed in the healthy renal parenchyma.

Our measurements were done unsystematic in the healthy renal parenchyma of the ipsilateral kidney and that partly explains the lower ICC values around 70% obtained between the two readers, compared to the very high ICC values from the renal tumours which are by definition more focal and often well-defined lesions. The ICC values for the healthy renal parenchyma in the same reader were very high, however, at around 95%.

The different nature of the solid renal tumours, namely the higher mitochondrial context of renal oncocytomas, HOCT and some variants of chromophobes 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:

ICC values obtained between the two readers concerning measurements on the healthy renal parenchyma around 70% suggest a 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 the 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 examined in this study with quantitative SUV SPECT® showed high ICC values around 90% not only in measurements of the same reader but also between two different readers. Those high ICC values indicate a strong agreement among SUV measurements for patients with solid renal lesions undergoing a $^{99m}$Tc-Sestamibi SPECT/CT examination.
Competing interests
The authors declare that they have no competing interests.

Ethics approval and consent to participate
Our pilot study is approved by the Regional Ethical Review Board and the local Radiation Safety Committee, Drn 2015/923-31/4. A written informed consent was acquired from all patients who participated in our study.

Funding
The above pilot project is financially supported from VINNOVA. MIDOR study, Drn 2015-0180.
**References:**


*EJNMMI Res* (2017) 7:80
Figure 1: Female patient with a 15mm solid neoplasm on the lateral aspect of the left kidney. Scintigraphy coronal view on the left demonstrates the absence of $^{99m}$Tc-Sestamibi (arrow) in the place of the renal tumour. Directly following the SPECT acquisition, a CT scan was performed (on the right) for anatomical correlation. Fused coronal SPECT/CT image in the middle. The above renal tumour was diagnosed as a clear cell carcinoma on histopathological grounds.
Figure 2: Graph (A) demonstrates all individual SUV\textsubscript{max} measurements in the solid renal tumours (reader 1: a and b measurements, reader 2: a and b measurements) for both readers. Graph (B) shows all individual SUV\textsubscript{max} measurements in the ipsilateral renal parenchyma (reader 1: a and b measurements, reader 2: a and b measurements). T: renal tumour, N: healthy renal parenchyma
Figure 3: Coronal $^{99m}$Tc-Sestamibi SPECT/CT fused image of healthy kidney parenchyma bilateral with regions of uneven $^{99m}$Tc-Sestamibi uptake on the left kidney. Right kidney though shows a more even $^{99m}$Tc-Sestamibi uptake.
### Tables:

<table>
<thead>
<tr>
<th>Test-retest same reader</th>
<th>ICC SUVmean</th>
<th>ICC SUVpeak</th>
<th>ICC SUVmax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reader 1 - T</td>
<td>0.984 (0.972-0.991)</td>
<td>0.989 (0.980-0.994)</td>
<td>0.989 (0.981-0.994)</td>
</tr>
<tr>
<td>Reader 2 - T</td>
<td>0.954 (0.920-0.974)</td>
<td>0.969 (0.946-0.982)</td>
<td>0.965 (0.939-0.980)</td>
</tr>
<tr>
<td>Reader 1 - N</td>
<td>0.975 (0.956-0.986)</td>
<td>0.982 (0.968-0.989)</td>
<td>0.983 (0.970-0.990)</td>
</tr>
<tr>
<td>Reader 2 - N</td>
<td>0.932 (0.883-0.961)</td>
<td>0.932 (0.884-0.961)</td>
<td>0.921 (0.865-0.954)</td>
</tr>
</tbody>
</table>

**Table 1**) Different ICC values for SUV\(_{\text{mean}}\), SUV\(_{\text{max}}\) and SUV\(_{\text{peak}}\) for repeated measurements from the same reader. In brackets the lower and the upper ICC values are presented with 95% confidence interval.

T: renal tumour, N: healthy renal parenchyma

<table>
<thead>
<tr>
<th>Between readers</th>
<th>ICC SUVmean</th>
<th>ICC SUVpeak</th>
<th>ICC SUVmax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reader 1 - T</td>
<td>0.890 (0.814-0.936)</td>
<td>0.887 (0.809-0.934)</td>
<td>0.866 (0.775-0.922)</td>
</tr>
<tr>
<td>Reader 2 - T</td>
<td>0.732 (0.530-0.848)</td>
<td>0.705 (0.526-0.823)</td>
<td>0.715 (0.518-0.835)</td>
</tr>
<tr>
<td>Reader 1 - N</td>
<td>0.858 (0.762-0.917)</td>
<td>0.915 (0.856-0.951)</td>
<td>0.890 (0.812-0.936)</td>
</tr>
<tr>
<td>Reader 2 - N</td>
<td>0.734 (0.560-0.843)</td>
<td>0.656 (0.447-0.794)</td>
<td>0.729 (0.524-0.847)</td>
</tr>
</tbody>
</table>

**Table 2**) Different ICC values for SUV\(_{\text{mean}}\), SUV\(_{\text{max}}\) and SUV\(_{\text{peak}}\) for repeated measurements between readers (reader 1 and reader 2). In brackets the lower and the upper ICC values are presented with 95% confidence interval.

T: renal tumour, N: healthy renal parenchyma