Journal of Nuclear Medicine, published on March 30, 2017 as doi:10.2967/jnumed.116.184440

Simulation of tracer dose reduction in 18F-FDG-Positron emission tomography / magnetic resonance imaging (PET/MRI): Effects on oncologic reading, image quality and artifacts.

Ferdinand Seith[1], MD, Holger Schmidt[1], PhD, Julia Kunz[2], MD, Thomas Küstner[1,3], PhD, Sergios Gatidis[1], MD, Konstantin Nikolaou[1], MD, Christian la Fougère[2], MD, and Nina Schwenzer[1], MD.

[1] Diagnostic and Interventional Radiology, Department of Radiology, Eberhard Karls University, Hoppe-Seyler-Straße 3, 72076 Tuebingen, Germany

[2] Nuclear Medicine and Clinical Molecular Imaging, Department of Radiology, Eberhard Karls University, Otfried-Müller-Straße 14, 72076 Tuebingen, Germany

[3] Institute of Signal Processing and System Theory, University of Stuttgart, Pfaffenwaldring 47, 70550 Stuttgart, Germany

Corresponding author:

Holger Schmidt, Diagnostic and Interventional Radiology, Department of Radiology, Eberhard Karls University, Hoppe-Seyler-Straße 3, 72076 Tuebingen, Germany, phone: 0049-7071-29-80544, fax: 0049-7071-29-4958, e-mail: holger.schmidt@uni-tuebingen.de

Article word count: 6176

First author:

Ferdinand Seith, Diagnostic and Interventional Radiology, Department of Radiology, Eberhard Karls University, Hoppe-Seyler-Straße 3, 72076 Tuebingen, Germany, phone: 0049-7071-29-86607, fax: 0049-7071-29-5845, e-mail: ferdinand.seith@uni-tuebingen.de

ABSTRACT

Purpose: Aim of our study was to evaluate the effect of stepwise reduced doses on objective and subjective image parameters and on oncologic readings in whole-body 18F-FDG-PET/MRI.

Material and Methods: We retrospectively simulated the stepwise reduction of 18F-FDG doses of 19 patients (50.9 ± 11.7 y/o, Body-Mass-Index (BMI) 22.8 ± 3.2) who received a whole-body PET/MRI examination from 3 to 0.5 MBq/kgBW in intervals of 0.25. Objective imaging parameters were assessed by measuring the standardized uptake value (SUV) and coefficient of variation (CV) in different regions (aorta, liver, spleen, kidney, small bowel, lumbar vertebra, psoas muscle, urinary bladder) as well as the noise equivalent count rates (NECR) in each bed position. Subjective image quality was evaluated with a blinded reading of each simulated PET compared to the dose of 2 MBq/kgBW. Oncologic reading was performed first, according to PET response criteria in solid tumors (PERCIST) in each dose and second, by defining malignant lesions in doses of 2 MBq/kgBW and the maximum dose image (gold standard). Diagnostic confidence of each lesion was measured using a Likert scale.

Results: With decreasing doses, regions in the mid abdomen showed a stronger decrease of SUVmean and NECR than regions in the upper abdomen (SUVmean: -45% and -15% on average in the small bowel and the liver, respectively). CV showed a non-linear increase, pronounced below 1.5 MBq/kgBW. Subjective image quality was stable

over a range between 1.25 and 2.75 MBq/kgBW compared to 2 MBq/kgBW. However, large photopenic areas in the mid abdomen were observed in two patients. In the PERCIST reading, target lesions were above the liver threshold with a stable SUVpeak in all cases down to 2 MBq/kgBW. 86 of 90 lesions were identified correctly with a dose of 2 MBq/kgBW; Likert scores did not differ significantly.

Conclusion: A reduction of doses in 18F-FDG-PET/MRI might be possible down to 2 MBq/kgBW in oncologic whole-body examinations. The image quality in the mid abdomen seems to be more affected by lower doses than in the upper abdomen and in single cases, large photopenic areas can occur. Therefore, we do not recommend reducing doses below 3 MBq/kgBW in adults at this time.

RUNNING TITLE: Dose Reduction in 18F-FDG-PET/MRI

Key words:

PET

Oncology

Dose Reduction

PET/MRI

INTRODUCTION

PET provides information of tissue properties regarding, e.g., metabolism or surface receptors on a molecular level. The possibility of implementing PET into hybrid imaging by combination with computed tomography (PET/CT) or magnetic resonance imaging (PET/MRI) in one single scanner facilitates the anatomic correlation of PET data and has led to an increasing acceptance in daily clinical routine. 18F-labeled fluorodeoxyglucose (18F-FDG) is the most widely used tracer in whole-body oncology, as it is able to detect an increased consumption of glucose, which is a characteristic of many tumor types (*1*). Therefore, PET with 18F-FDG plays a pivotal role in several oncologic diseases like lung cancer or lymphoma (*2-5*). In oncology, PET is of increasing importance not only for the initial staging of patients but also in therapy monitoring or the evaluation of tumor recurrence (*6,7*). According to the response evaluation criteria in solid tumors for oncologic trials in radiology, Wahl et al. introduced a model in order to standardize oncologic reading of PET in clinical trials: PERCIST (*8*). With the increasing use of PET in oncologic imaging and therefore in repetitive follow-up examinations, radiation exposure increases for patients on the one hand, but also for technologists and physicians on the other hand (*9*). Today's guidelines for tracer doses in PET are mostly based on phantom studies, theoretical model calculations or retrospective evaluations of examinations of different patient populations (*10-15*). Due to the high sensitivity of PET detectors in current PET/MRI scanners (*16,17*) and commonly longer PET acquisition times compared to PET/CT (*18,19*), a reduction of PET tracer doses might be feasible for PET/MRI examinations. However, an intra-individual analysis of the influence of different tracer doses on image quality is ethically

not feasible in a prospective study setting. Recently, Gatidis et. al proposed a new approach to simulate PET tracer dose reduction of whole-body examinations by a randomized undersampling of PET list-mode data (*20*). Thus, it is possible to retrospectively simulate PET examinations of several lower doses of one patient by randomly deleting proportions of PET events over the acquisition time. In a recently published study, Schaefferkoetter et al. investigated the influence of a simulated reduction of 18F-FDG doses on image characteristics in the lung and found that even a low-dose PET of 18.5 MBq might be sufficient for lung cancer screening (*21*).

The aim of our study was to evaluate the effects of dose reduction in 18F-FDG-PET on objective and subjective imaging parameters and the influence on oncologic reading in whole-body examinations using stepwise simulated reduced tracer doses.

MATERIAL AND METHODS

Patients

From October 2012 to July 2014, we retrospectively evaluated all patients who were examined in a fully integrated simultaneous PET/MRI scanner (Biograph mMR, Siemens Healthineers) in a clinical setting. Inclusion criteria were a torso 18F-FDG-PET-scan of adults (>18y), an injected tracer dose of at least 3 MBq/kgBW, and an acquisition time per bed of at least 4 minutes. Written informed consent concerning the examination and scientific evaluation of their data was given by all patients.

Hybrid Examinations

Prior to being injected with 18F-FDG, all patients fasted at least 6 h. Between the tracer injection and the examination, patients were resting on a patient couch. Acquisition time per bed position was 4 to 6 minutes. All PET-data were stored in list mode. A 3D T1-weighted spoiled gradient-echo sequence in end-expiratory breath-hold with Dixon-based fat-water separation (for vendor-provided PET attenuation correction) and a coronal T2-weighted short-time inversion-recovery sequence were acquired simultaneously with PET. Other MR sequences were chosen according to the disease and clinical question. If the acquisition time per bed position exceeded 4 minutes, the PET-data were shortened to 4 minutes per bed position by excluding the counts acquired in the excessive minutes. The simulation of doses was carried out by a retrospective randomized undersampling of list-mode data as described by Gatidis et al. (*20*). Here, a predefined proportion of PET events are randomly deleted in the PET list-mode data. These undersampled datasets can be reconstructed as simulated low-dose PET-images. Using this method, we calculated PET-images of the following doses of each patient (MBq/kgBW): 3.0, 2.75, 2.5, 2.25, 2.0, 1.75, 1.5, 1.25, 1.0, 0.75 and 0.5. All PET data were acquired in 3D mode and reconstructed using the vendor provided e7-tools software package (Siemens Healthineers, Erlangen, Germany) with following parameters: 3D ordered subset expectation maximization algorithm, 21 subsets, 2 iterations, 256x256 matrix size, voxel size 2.8x2.8x2mm^3 and 4mm Gaussian filter. The results were controlled visually.

Quantitative Assessment of PET Image Quality

In order to assess the reliability of the evaluation of tracer uptake and to give an impression of the image noise in different doses we acquired the SUVmean (corrected by bodyweight) and the CV in physiological organs. The estimated trues, prompts, NECR and the total scatter fraction (SF) demonstrate the effects of dose modulations on technical scan parameters in different body regions.

In the PET-images reconstructed with the original dose fused with the corresponding short-time inversion-recovery sequence, 2D circular regions-of-interest with a target diameter of 1cm were set in axial plane in the spleen, the descending aorta, the renal pelvis, the psoas muscle, the urinary bladder, the small bowel in the middle abdomen and in the fourth lumbar vertebra in coronal plane. Moreover, 3D spherical volume-of-interest of 3 cm diameter was set in the right liver lobe as proposed by Wahl et al. for PERCIST evaluations (*8*). Care was taken to avoid organ borders, and only organs without PET-positive lesions were included. Those regions or volumes-of-interest were copied to all PET-images at various dose simulations of the respective patient to measure the SUVmean and the standard deviation (SD). These steps were performed using PMod (PMOD Technologies Ltd, Zurich, Switzerland) and MATLAB (The MathWorks Inc, Natick, MA). The CV within a region-of-interest was defined as SD / SUVmean.

Depending on the covered body region, the bed positions of an examination were classified in the following categories: I) head/neck, II) thorax/neck, III) upper abdomen/lower thorax, IV) abdomen/pelvis and V) pelvis/upper thighs (Figure 1). The

estimated trues, prompts, NECR and SF of each bed position were extracted from the PET data header.

Oncologic Readings

All examinations were evaluated in a clinical setting in consensus reading by a radiologist and a nuclear medicine physician, each with more than eight years of experience in hybrid imaging. In this reading, all malignant lesions in various oncologic diseases were defined. All examinations with PET-positive malignant findings were included in the oncologic readings.

We performed two different oncologic readings. First, according to PERCIST, to demonstrate the effects of dose reduction on clinical oncologic trials (PERCIST). Second, in a nuclear medicine physician reading procedure to represent the effects of a dose reduction down to 2 MBq/kgBW on oncologic readings in daily clinical routine compared to the full-dose PET-images (lesion detection).

PERCIST. We calculated a liver threshold according to PERCIST with the 3D spherical volume-of-interest in the right liver lobe in all simulated PET-images of a patient as follows: 1.5*SUVmean + 2*SD. According to PERCIST, a maximum of five malignant lesions with the highest SUVpeak were defined as target lesions (max. two per organ) in PET-images with doses of 3 MBq/kgBW. A 3D spherical volume-of-interest was drawn covering the whole lesion to assess the SUVmax, SUVpeak and the metabolic volume (MTV) of each lesion in each PET-image. The SUVpeak was defined

by Wahl et al. (*8*) as the largest possible mean of value of a 1cm^3 spherical volume-of-interest positioned within a tumor. The MTV was defined as a volume within a 40% threshold of the SUVmax. These steps were performed with the software MMOncology (syngo.via, Siemens Healthineers, Erlangen, Germany).

Lesion Detection. In a second oncologic reading, a nuclear medicine physician with five years of experience in PET imaging (J.K.) evaluated the examinations with a simulated dose of 2 MBq/kgBW and the full-dose PET-image (gold standard) of each oncologic patient to define malignant lesions. As a first benchmark, 2 MBq/kgBW was chosen, based on the observations of the image parameters (relative stable behaviour of SUV, CV and NECR, see Figs. 1 and 2) and the results from the first oncologic reading (no lesion with SUVpeak below this threshold). The reader was blinded for the respective dose and waited at least six weeks between the readings of the different datasets. A lesion was defined as a focal excess of tracer uptake compared to the surrounding regional uptake. A maximum of ten lesions per organ were evaluated. For each lesion, the diagnostic confidence was rated using a Likert scale: 1=highly suspicious for a malignant lesion, 2=suspicious for a malignant lesion, 3=probably a malignant lesionand 4=possibly a malignant lesion.

Subjective Image Evaluation

We assessed the subjective image impression in order to demonstrate the effects of dose reduction on the overall confidence of a nuclear medicine physician reading PET-

images of lower doses. For this purpose, the nuclear medicine physician compared all reconstructed doses of a patient with the simulated PET of 2 MBq/kgBW in a blinded randomized setting with an in-house written MATLAB reading software (https://sites.google.com/site/kspaceastronauts/blindfold) in the following manner: Two PET-images of a randomly picked patient were simultaneously presented on a screen, side by side. One PET-image was reconstructed from the 2 MBq/kgBW dose and was always presented randomly on the right or left side. We chose 2 MBq/kgBW instead of 3 MBq/kgBW to prevent a selection bias for the higher dose. The other PET-image was of a randomly picked dose (higher or lower). The reader always rated the image on the left.hand side in comparison to the image on the right hand side as of (+) better image quality, (=) comparable image quality or (-) worse image quality regarding the subjective overall diagnostic validity focussed on image artifacts and overall homogeneity.

Statistical Analysis

The results of the second oncologic reading were compared with a McNemar's test with the continuity correction. The scores of the Likert scale were statistically compared using analysis of variance with Bonferroni post hoc correction. A p-value of <0.05 was considered as significant. The correlation of the trend in % of MTV and SUVmax in % was calculated with a linear regression using the ordinary least-squares algorithmic technique. Analyses were performed with the software JMP (Version 11, SAS institute, Cary, NY).

RESULTS

In total, 19 patients (14 female, 50.9±11.7 y/o, BMI 22.8±3.2) met the inclusion criteria and were included in the evaluation. Patients' diseases were distributed as follows: teratocarcinoma (n=1), squamous cell carcinoma (n=1), bronchial carcinoma (n=3), breast cancer (n=7), esophageal cancer (n=1), rectal carcinoma (n=2), lymphoma (n=1), gastric cancer (n=1), neurofibromatosis (n=1) and fever of unknown origin (n=1). The mean injected dose was 229.6±67.2 MBq; the mean uptake time was 64.8±9.3 minutes. The simulation of tracer dose reductions in dependence on patients' bodyweight led to the following doses of 18F-FDG in MBq (± standard deviations): 203.0±34.9, 186.1±32.0, 169.2±29.1, 152.2±26.2, 135.3±23.3, 118.4±20.4, 101.5±17.4, 84.6±14.5, 67.7±11.6, 50.7±8.7 and 33.8±5.8. An example of all reconstructed simulated doses of a patient is given in figure 3.

Quantitative Assessment of PET Image Quality

The results of the measurement of SUVmean and CV in physiological tissue are given in figure 2. The SUVmean in the small bowel, the psoas muscle and the lumbar vertebra showed a stronger dependence on doses with a pronounced decrease for reduced doses (up to -45% on average in the psoas muscle) compared to other investigated regions (up to -25% in the spleen). At doses of 3 MBq/kgBW, the CV in the renal pelvis was highest (CV=0.18). The CV in all organs showed a pronounced increase for decreasing doses, especially for doses below 1.5 MBq/kgBW. Fitting the CV

with a power law (R^2 > 0.91 for all organs) indicates that the noise characteristics follow the Poisson relationship (respecting the fact that dead time, normalization, attenuation correction and the reconstruction algorithm might alter this relationship (*11,22*)). In all examined body regions, the NECR tended to decrease while the estimated SF tended to increase with decreasing doses (Figure 1). In the abdominal/pelvis region, this dependence of NECR and SF on doses was most pronounced (NECR at 2 MBq/kgBW: abdominal/pelvis regions: 88.2%, head/neck region: 99.9% of the NECR at 3 MBq/kgBW). Moreover, the highest estimated SF of PET-images was found in this region (48.7% at 3 MBq/kgBW), and the estimated SF reached the system maximum of 0.75 in low doses in seven patients: In three cases at 0.5 MBq/kgBW, in two cases below 0.75 MBq/kgBW, in one case below 1.5 MBq/kgBW and in one case below 1.75 MBq/kgBW. This is a predefined limit in the vendor's software to assure stable data correction and image reconstruction. In one of those patients, the system maximum of the estimated SF was also reached in the pelvis/upper thighs region at 0.5 MBq/kgBW.

Oncologic Readings

Thirteen patients out of 19 had PET-positive malignant lesions and were therefore included in the oncologic reading.

PERCIST. In one of those patients, the SUVpeak of the lesions was below the liver threshold. Therefore, this patient was excluded from the PERCIST oncologic reading. Overall, 20 lesions in 12 patients were examined in the PERCIST oncologic reading. Of

these, three were located in the skeleton, two in the mammary gland, one in the lung, one in the adrenal gland and one in the rectum. Moreover, twelve metastatic lymph nodes were evaluated. The results of the PERCIST reading including SUVmax, SUVpeak, the metabolic volume, and the PERCIST threshold which apply to the liver are presented in figure 4. The PERCIST threshold was revealed to be of relatively stable behavior, with an increase for doses below 1.25 MBq/kgBW. The SUVpeak showed a slight decrease for decreasing doses, less pronounced however, compared to the SUVmean in organs. (The mean value of SUVpeak at a dose of 0.75 MBq/kgBW was 96% of 3 MBq/kgBW compared to 90% at the same dose of SUVmean in the liver.) The SUVmax and the metabolic volume showed the strongest dependence on dose reduction (MTV: 77% of 3 MBq/kgBW on average at a dose of 0.75 MBq/kgBW). The trend of the MTV showed a moderate inverse correlation with the trend of SUVmax (R^2=0.27). With decreasing doses, the SUVpeak of a target lesion was below the liver threshold in three lesions: in one lesion at 1.75 MBq/kgBW, in one lesion below 1.0 MBq/kgBW and in one lesion at 0.5 MBq/kgBW.

Lesion Detection. In the second oncologic reading, 90 lesions were defined in the original PET-images of the maximum doses (gold standard): 51 metastatic lymph nodes, 15 osseous lesions, 9 pleural lesions, 6 lung metastases, 3 lesions in the mammary glands, 2 in the adrenal glands, 2 in the subcutaneous fat, 1 in the thyroid gland and 1 in the rectum.

Of those, 86 were identified correctly in the 2 MBq/kgBW PET-images. Overall, the difference to full dose PET data was not significant (p=0.95). No false-positive lesions

were found in the 2 MBq/kgBW PET-images. Four metastatic lymph nodes were missed in two patients in the 2 MBq/kgBW PET, located in the mediastinum (n=2) and the neck (n=2). One of them is demonstrated in Supplemental Material 1. Three of those lymph nodes were rated with a Likert score of 4 (possibly a malignant lesion) and one with a Likert score of 3 (probably a malignant lesion) in the original PET data with full dose. The Likert score on the 2 MBq/kgBW PET-images was rated lower than on the full-dose PET-images in five lesions of two patients. However, the average Likert score was 1.8 in both PET-images and did not differ significantly (p=0.84).

Subjective Image Evaluation

The results of the evaluation of the subjective impression of image quality are given in figure 5. Over a range between doses of 1.25 and 2.75 MBq/kgBW most PET-images were rated as equally good compared to the 2 MBq/kgBW. One PET-image with 2.75 MBq/kgBW and nine PET-images with 3 MBq/kgBW were rated as substantially better. One PET-image was rated worse at doses of 1.75, 1.5 and 1.25 MBq/kgBW, respectively. However, five PET-images at 1.75 Mq/kgBW were rated as better than 2 MBq/kgBW.

With decreasing doses increasing photopenic areas were observed. Typical photopenic artifacts in PET-images were located around the urinary bladder in three patients. Moreover, in two different patients, we observed large photopenic areas in the middle abdomen, not surrounding the kidneys or the urinary bladder; one of them is demonstrated in figure 6. BMI of those patients was 24 and 26, respectively.

DISCUSSION

In positron emission tomography, the annihilation of positrons and electrons leads to the emission of 511 keV gamma photons. This high-energy radiation might potentially harm patients and personnel in the long term (*9,23,24*). The increasing use of 18F-FDG-PET in oncologic imaging not only in initial staging but also in follow-up examinations, also results in an increase of radiation exposure for both patients and the medical personnel. Thus, a reduction of PET tracer doses is desirable. The definition of today's doses in 18F-FDG-PET is mostly based on theoretical models, phantom studies and the retrospective comparison of different groups of patients (*10-15*). As PET is a quantitative statistical imaging technique which is based on the Poisson-distributed random process of positron emission, theoretical calculations are legitimate. Those models use theoretical benchmarks like image noise (which is calculated as the standard deviation of the counts in a region-of-interest) and NECR as surrogates for the assessment of image quality. They are aimed at optimizing objectified parameters like the NECR peak and do not consider individual characteristics of the patient, physiological tracer distributions or the subjective requirements of the physician for an adequate reading of the examination. It is therefore not clear, how representative, e.g., a phantom study is for an individual patient (*25*). Moreover, Chang et al. (*14*) reported in a retrospective study with a large number of patients that a scanner with a higher NECR does not necessarily translate to lower image noise and therefore better image quality, depending on the technique used for image reconstruction. Therefore, the conjunction between NECR and the real image impression or subjective image quality might be more complex. Moreover, to evaluate image quality, many studies measure the image noise in

homogeneous liver tissue (*12*); however, this might not be sufficient to represent the PET-image quality of a whole-body examination, as PET-image quality of individual regions might be influenced by regional patient characteristics. This uncertainty is also represented in the heterogeneous recommendations for doses in 18F-FDG-PET, e.g., comparing the respective European and American societies (*13,26*). By applying the proposed method of retrospectively modifying PET list-mode data, we could perform a retrospective intra-individual comparison of PET-images with different doses.

The first part of our study revealed the effects of dose reduction on objective imaging parameters. We observed variable influences of doses on the system NECR and estimated SF in different body regions, pronounced in the abdominal/pelvis region. Therefore, the influence of doses on image quality differs in different body regions. With decreasing doses, an overall non-linear increase of the CV in organs can be seen. Additionally, an enhanced effect on the average SUVmean of the small bowel, the psoas muscle and the lumbar vertebra (mid abdominal region) with a pronounced decrease compared to the other investigated regions (upper abdomen or urinary bladder) was observed with decreasing doses. However, it has to be noted that those regions provide low SUV and thus small changes cause increased relative deviations. The finding of increasing estimated SF and decreasing SUVmean with decreasing doses is likely to be connected, as an artificially growing SF results in an underestimation of measured tracer uptake.

In two patients, we observed large photopenic areas in the middle abdomen with decreasing dose; one of them is demonstrated in figure 6. With BMIs of 24 and 26, those patients were not obese - a common potential reason for impaired PET-image quality. In

both cases, the estimated SF reached the system limit of 0.75 in those bed positions at doses of 1.75 MBq/kgBW and below, respectively. The system NECR in these cases showed a break in the decreasing curve at the corresponding doses (Figure 7). Therefore, the high amount of scatter resulting from an erroneous scatter estimation (or more precisely from erroneous scatter scaling) hampers the image reconstruction or the scatter correction in particular. A reconstruction of PET-images without scatter correction or scatter scaling did not reveal these photopenic areas (see Supplemental Material 2). It has to be noted that the NECR and SF reported were not measured but estimated (calculated) by the scanner. Thus, the computation of the SF and consequently the SUV is incorrect in those regions at low tracer doses. However, since no PET-positive lesions were found in those regions for either patient, the clinical impact on oncologic examinations remains unclear.

Scatter estimation is usually carried out by single scatter simulation (*27,28*) while the correction for multiple scatter is done via scatter scaling (scatter fitting outside attenuating regions in the emission sinogram (*29*)). Thus, scatter scaling can be problematic for low tracer concentrations or specific tracers (low count rate outside attenuating regions; (*30*)) as well as for high tracer concentrations outside the PET field-of-view (*31*). This is especially true for current PET/MRI systems which have higher estimated scatter fractions than PET/CT systems (*32*). Thus, to allow PET imaging with low tracer activities, an optimization of the scatter correction is needed (e.g., via fast Monte Carlo simulations; (*33,34*)). However, since the attenuation correction map in an essential input for the scatter estimation, an optimization of the MR-based attenuation correction might also be an important point of action to improve the reconstruction

results (e.g., (*35,36*)). A workaround to reduce photopenic areas in the abdominal region might be an optimized positioning of PET fields-of-view (PET field-of-view not between kidney and bladder). Also, prolonged acquisition times in this region might compensate for low-dose exams. However, this has to be further evaluated in future studies. Of course, ensuring the voiding of the patients' bladder is very important.

Other investigated regions showed a more stable behavior with a pronounced decrease of SUVmean from below 1.25 MBq/kgBW. A typical photopenic artifact in PET-images was located around the urinary bladder. This artifact was severely increasing with decreasing doses in three patients (different patients than the previously described cases). As a consequence, in one of those cases, a PET-positive carcinoma of the rectum could not be distinguished from the background at a dose of 0.5 MBq/kgBW (Supplemental Material 3).

The oncologic evaluation according to PERCIST revealed that the calculated threshold in the liver shows a relative stable behavior compared to the other examined parameters even with a slight increase of the threshold at very low doses. This behavior can be explained by the increase of standard deviation in lower dose images (Threshold = 1.5*SUVmean + 2*SD), which therefore countervails the trend of decreasing SUVmean (Figs. 2 and 4). In some target lesions, the SUVpeak in the highest dose was just close above the liver threshold. In three of those lesions, the SUVpeak was below the threshold in lower doses (max. 1.75 MBq/kgBW). This might have severe influence on oncologic readings, as those lesions would not have been rated as target lesions. Moreover, considering the Deauville criteria for lymphoma (*7*) and the overall trend

towards quantitative assessment in PET, a reliable measurement of SUV is of utmost importance, especially in clinical trials.

The SUVmax and the metabolic volume were revealed to be unstable parameters with a strong correlation to doses. The metabolic tumor volume can influence the oncologic reading in follow-up examinations, in texture analysis, as well as on therapy planning in radiation therapy (*8,37,38*).

With a simulated dose of 2 MBq/kgBW, the subjective image impression as well as the diagnostic performance of oncologic reading did not differ substantially from higher doses in most cases: The evaluation of the subjective image impression revealed that over a wide range of doses, no substantial differences compared to 2 MBq/kgBW were found. The reading of the image impression represents an individual subjective evaluation by an experienced nuclear medicine physician at a certain time; however, five PET-images with a simulated dose of 1.75 MBq/kgBW were rated better than the 2 MBq/kgBW PET-images. With a dose of 2 MBq/kgBW, most lesions could be identified correctly and no false negative or false positive findings in an organ or patient based analysis were observed. Moreover, Likert scores regarding diagnostic confidence of lesions did not differ significantly from those of the maximum doses.

Previously conducted studies investigated the effects of simulated dose reduction in dependence on predefined true count levels on image characteristics of 18F-FDG- PET-images of the lung (*21,39*). We chose a more clinical approach by calculating whole-body PET-images in dependence on patients' characteristics, since count levels differ for different body regions and cannot be predicted. De Groot et al. (*12*) suggest that the signal-to-noise-ratio in PET is more strongly correlated to the body mass than to the lean

body mass, the BMI or the fat mass; thus we decided to calculate our doses in relation to kilogram bodyweight. Yan et al. (*39*) determined a lower limit of 5 million true counts for the Biograph mMR, which would roughly correspond to a dose below 0.5 MBq/kgBW in our setting. However, the dose-determining factor is the lower abdomen, which was not assessed by their work. Schaefferkoetter et al. (*21*) found similar results for PET/CT. Interestingly, they also found an increase in the estimated SF below 1 million true counts in the lung. Unfortunately, the abdominal region was not assessed here as well.

Our study has several limitations. The low number of patients with a relative homogeneous BMI and with different types of tumors does not allow for general recommendations. The investigated regions were mostly located in the abdomen; other regions like the thorax or the brain were not evaluated and are beyond the scope of our work. In the oncologic reading according to PERCIST, only target lesions were investigated; however, non-target lesions can also have a significant influence on the evaluation of PET examinations in oncology. The initial doses in our study are already relatively low compared to guidelines, as it is generally accepted that prolonged acquisition times have the potential to compensate for lower doses (*19*). Bed positions of a patient were classified in different categories; as our study was performed in a clinical setting, patients show different body heights and body proportions. Therefore, the different beds do not always cover exactly the same body region or organs. Also, the random deletion of PET events from listmode data assumes a linear dependence of true and random events to the single counts (similar to Schaefferkoetter et al. (*21*)). However, it is known that randoms show a quadratic dependence. Even though results from Gatidis et al. (*20*) indicated that there is no significant difference between

measured and simulated PET-images for the activity concentration range used here, our results represent a worst-case scenario; image quality might be improved for measured low-dose PET-images as randoms and dead-time should be lower at lower doses. Finally, all PET-images in all simulated doses were reconstructed with the same parameters. However, for PET-images of lower doses an adjustment of reconstruction parameters like Gaussian filtering or the number of subsets might lead to improved image impressions or lesion detections.

CONCLUSION

It could be shown that a reduction of 18F-FDGdoses down to 2 MBq/kgBW might be possible in whole-body PET/MRI examinations of oncologic patients with no distinct effect on lesion detection, diagnostic confidence or a study evaluation according to PERCIST criteria. However, the impact of reducing doses on image quality and SUV measurements is highly dependent on body regions, mostly affecting the middle abdomen. Moreover, with decreasing doses we observed the occurrence of large photopenic artifacts in the middle abdomen in single patients. The occurrence of this artifact cannot currently be predicted and the clinical impact is unclear. Therefore, we do not recommend reducing doses below 3 MBq/kgBW in oncologic PET/MRI examinations in adults at this time.

REFERENCES

1. Larson SM, Schwartz LH. 18F-FDG PET as a candidate for "qualified biomarker": functional assessment of treatment response in oncology. *J Nucl Med.* 2006;47:901-903.

2. Hoppe RT, Advani RH, Ai WZ, et al. Hodgkin lymphoma, version 2.2012 featured updates to the NCCN guidelines. *J Natl Compr Canc Netw.* 2012;10:589-597.

3. Barrington SF, Mikhaeel NG, Kostakoglu L, et al. Role of imaging in the staging and response assessment of lymphoma: consensus of the international conference on malignant lymphomas imaging working group. *J Clin Oncol.* 2014;32:3048-3058.

4. Ioannidis JP, Lau J. 18F-FDG PET for the diagnosis and grading of soft-tissue sarcoma: a meta-analysis. *J Nucl Med.* 2003;44:717-724.

5. Berghmans T, Dusart M, Paesmans M, et al. Primary tumor standardized uptake value (SUVmax) measured on fluorodeoxyglucose positron emission tomography (FDG-PET) is of prognostic value for survival in non-small cell lung cancer (NSCLC): a systematic review and meta-analysis (MA) by the european lung cancer working party for the IASLC lung cancer staging project. *J Thorac Oncol.* 2008;3:6-12.

6. Murakami R, Kumita S, Yoshida T, et al. FDG-PET/CT in the diagnosis of recurrent breast cancer. *Acta Radiol.* 2012;53:12-16.

7. Meignan M, Gallamini A, Haioun C. Report on the first international workshop on interim-PET-Scan in lymphoma. *Leuk Lymphoma.* 2009;50:1257-1260.

8. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med.* 2009;50 Suppl 1:122S-150S.

9. Leide-Svegborn S. Radiation exposure of patients and personnel from a PET/CT procedure with 18F-FDG. *Radiat Prot Dosimetry.* 2010;139:208-213.

10. Daube-Witherspoon ME, Karp JS, Casey ME, et al. PET performance measurements using the NEMA NU 2-2001 standard. *J Nucl Med.* 2002;43:1398-1409.

11. Watson CC, Casey ME, Bendriem B, et al. Optimizing injected dose in clinical PET by accurately modeling the counting-rate response functions specific to individual patient scans. *J Nucl Med.* 2005;46:1825-1834.

12. de Groot EH, Post N, Boellaard R, Wagenaar NR, Willemsen AT, van Dalen JA. Optimized dose regimen for whole-body FDG-PET imaging. *EJNMMI Res.* 2013;3:63.

13. Boellaard R, Delgado-Bolton R, Oyen WJ, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging.* 2015;42:328-354.

14. Chang T, Chang G, Kohlmyer S, Clark JW, Rohren E, Mawlawi OR. Effects of injected dose, BMI and scanner type on NECR and image noise in PET imaging. *Phys Med Biol.* 2011;56:5275-5285.

15. Makris NE, Huisman MC, Kinahan PE, Lammertsma AA, Boellaard R. Evaluation of strategies towards harmonization of FDG PET/CT studies in multicentre trials: comparison of scanner validation phantoms and data analysis procedures. *Eur J Nucl Med Mol Imaging.* 2013;40:1507-1515.

16. Delso G, Furst S, Jakoby B, et al. Performance measurements of the siemens mMR integrated whole-body PET/MR scanner. *J Nucl Med.* 2011;52:1914-1922.

17. Grant AM, Deller TW, Khalighi MM, Maramraju SH, Delso G, Levin CS. NEMA NU 2-2012 performance studies for the SiPM-based ToF-PET component of the GE SIGNA PET/MR system. *Med Phys.* 2016;43:2334.

18. Schafer JF, Gatidis S, Schmidt H, et al. Simultaneous whole-body PET/MR imaging in comparison to PET/CT in pediatric oncology: Initial Results. *Radiology.* 2014:131732.

19. Oehmigen M, Ziegler S, Jakoby BW, Georgi JC, Paulus DH, Quick HH. Radiotracer dose reduction in integrated PET/MR: implications from national electrical manufacturers association phantom studies. *J Nucl Med.* 2014;55:1361-1367.

20. Gatidis S, Wurslin C, Seith F, et al. Towards tracer dose reduction in PET studies: simulation of dose reduction by retrospective randomized undersampling of list-mode data. *Hell J Nucl Med.* 2016;19:15-18.

21. Schaefferkoetter JD, Yan J, Soderlund TA, et al. Quantitative accuracy and lesion detectability of low-dose FDG-PET for lung cancer screening. *J Nucl Med.* September 29, 2016 [Epub ahead of print].

22. Teymurazyan A, Riauka T, Jans HS, Robinson D. Properties of noise in positron emission tomography images reconstructed with filtered-backprojection and row-action maximum likelihood algorithm. *J Digit Imaging.* 2013;26:447-456.

23. Manning G, Taylor K, Finnon P, Lemon JA, Boreham DR, Badie C. Quantifying murine bone marrow and blood radiation dose response following (18)F-FDG PET with DNA damage biomarkers. *Mutat Res.* 2014;770:29-36.

24. Harrison J, Day P. Radiation doses and risks from internal emitters. *J Radiol Prot.* 2008;28:137-159.

25. Badawi RD, Adam LE, Zimmerman RE. A simulation-based assessment of the revised NEMA NU-2 70 cm long test phantom for PET. Paper presented at: Nuclear Science Symposium Conference Record, 2001 IEEE; 4-10 Nov. 2001, 2001.

26. American College of Radiology. ACR–SPR practice parameter for performing FDG-PET/CT in oncology. https://www.acr.org/~/media/ACR/Documents/PGTS/guidelines/FDG_PET_CT.pdf. Accessed December 15, 2016.

27. Watson CC, Newport D, Casey ME. A single scatter simulation technique for scatter correction in 3D PET. In: Grangeat P, Amans J-L, eds. *Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*. Dordrecht: Springer Netherlands; 1996:255-268.

28. Watson CC, Newport D, Casey ME, deKemp RA, Beanlands RS, Schmand M. Evaluation of simulation-based scatter correction for 3-D PET cardiac imaging. *IEEE Trans Nucl Sci.* 1997;44:90-97.

29. Watson CC. New, faster, image-based scatter correction for 3D PET. *IEEE Trans Nucl Sci.* 2000;47:1587-1594.

30. Afshar-Oromieh A, Haberkorn U, Hadaschik B, et al. PET/MRI with a 68Ga-PSMA ligand for the detection of prostate cancer. *Eur J Nucl Med Mol Imaging.* 2013;40:1629-1630.

31. Schmidt H, Schwenzer NF, Bezrukov I, et al. On the quantification accuracy, homogeneity, and stability of simultaneous positron emission tomography/magnetic resonance imaging systems. *Invest Radiol.* 2014;49:373-381.

32. Delso G, Martinez MJ, Torres I, et al. Monte carlo simulations of the count rate performance of a clinical whole-body MR/PET scanner. *Med Phys.* 2009;36:4126-4135.

33. Gaens M, Bert J, Pietrzyk U, Autret A, Shah NJ, Visvikis D. GPU-accelerated monte carlo based scatter correction in brain PET/MR. *EJNMMI Phys.* 2014;1:A32.

34. Kim KS, Son YD, Cho ZH, Ra JB, Ye JC. Ultra-fast hybrid CPU-GPU multiple scatter simulation for 3-D PET. *IEEE J Biomed Health Inform.* 2014;18:148-156.

35. Blumhagen JO, Ladebeck R, Fenchel M, Scheffler K. MR-based field-of-view extension in MR/PET: B0 homogenization using gradient enhancement (HUGE). *Magn Reson Med.* 2013;70:1047-1057.

36. Paulus DH, Quick HH, Geppert C, et al. Whole-body PET/MR imaging: quantitative evaluation of a novel model-based MR attenuation correction method Including Bone. *J Nucl Med.* 2015;56:1061-1066.

37. Wang K, Heron DE, Clump DA, et al. Target delineation in stereotactic body radiation therapy for recurrent head and neck cancer: a retrospective analysis of the impact of margins and automated PET-CT segmentation. *Radiother Oncol.* 2013;106:90-95.

38. Yip SS, Coroller TP, Sanford NN, Mamon H, Aerts HJ, Berbeco RI. Relationship between the temporal changes in tositron-emission-tomography-imaging-based textural features and pathologic response and survival in esophageal cancer patients. *Front Oncol.* 2016;6:72.

39. Yan J, Schaefferkoette J, Conti M, Townsend D. A method to assess image quality for low-dose PET: analysis of SNR, CNR, bias and image noise. *Cancer Imaging.* 2016;16:26.

Figures

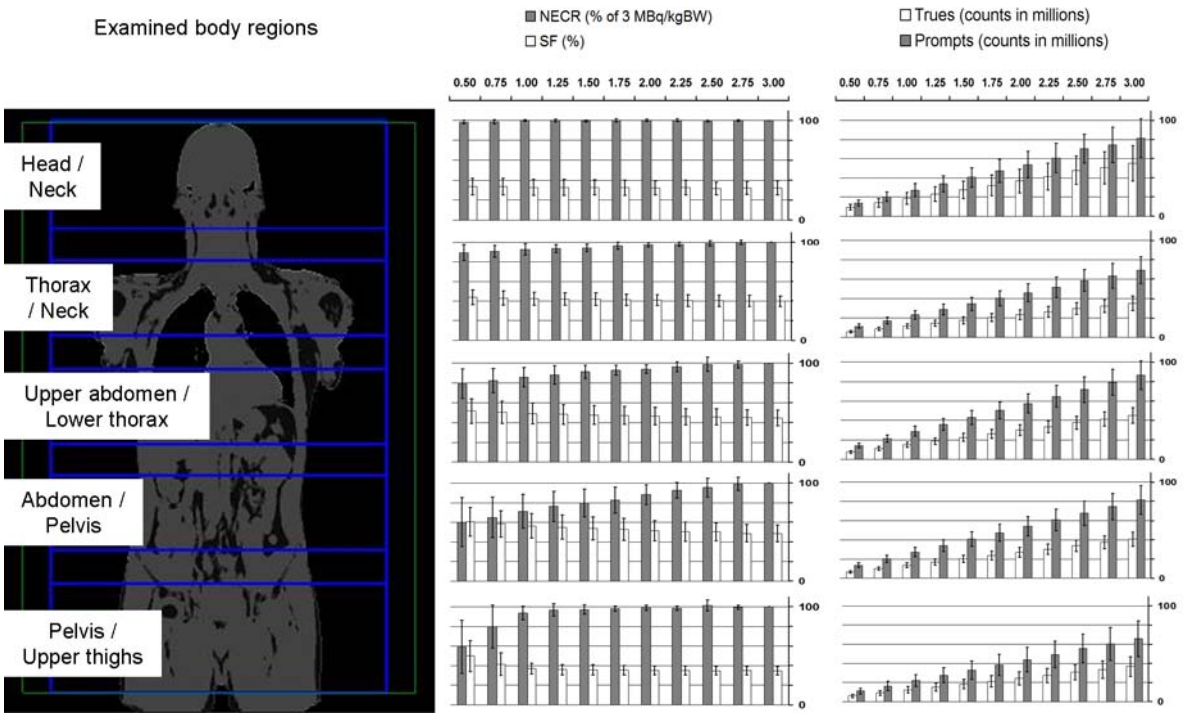

Figure 1: Overview of the NECR in % compared to 3 MBq/kgBW, the estimated SF in %, the trues and the prompts in each examined body region. Error bars represent the standard deviations. Note the different levels of SF and the strong dependence of the NECR and the SF on the reduction of doses in the abdomen/pelvis region.

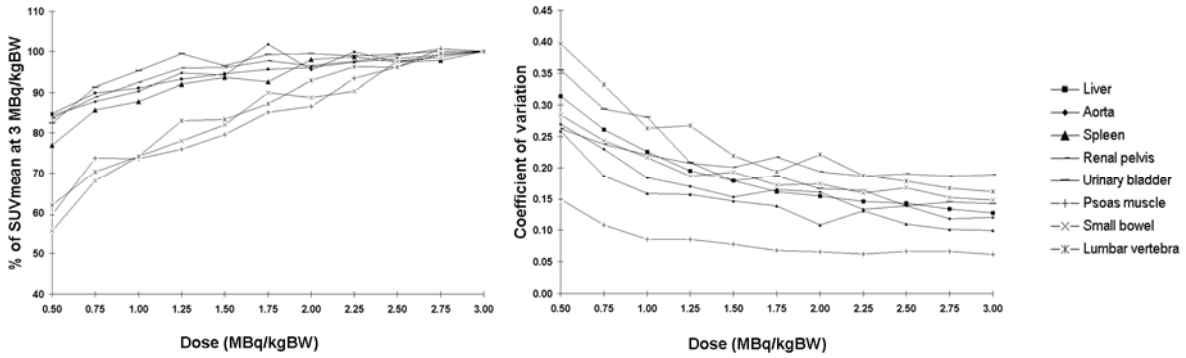

Figure 2: Means of SUV and CV in organs compared to 3 MBq/kgBW. Regions-of-interest located in the abdominal region with a low tracer uptake like the psoas muscle, the small bowel or the lumbar vertebra showed a stronger relative decrease of SUVmean compared to the other regions. A pronounced increase of CV in doses below 1.5 MBq/kgBW was observed. A power-law-fit of the CV revealed $R^2 > 0.96$ except for aorta, spleen, pelvis and muscle ($R^2 > 0.91$).

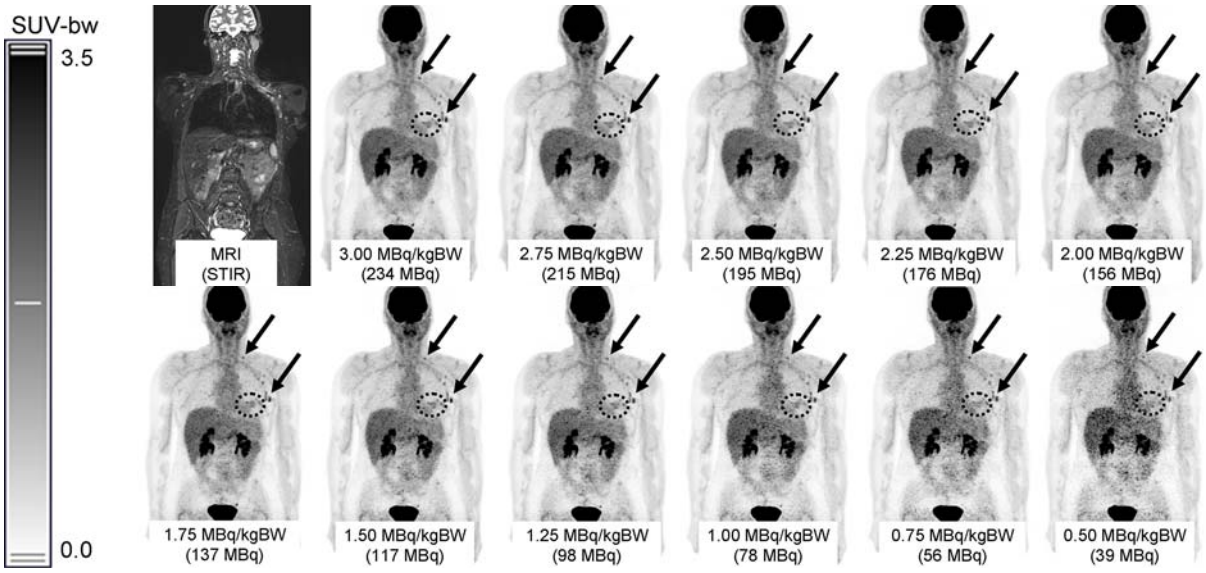

Figure 3: Example of a 57 y/o female patient with metastasized breast cancer. STIR = short-time inversion-recovery. Lymph node metastases in the left axilla and supraclavicular region are marked with black arrows, the cancer in the patient's left breast is highlighted with a dotted circle. In all simulated doses, the tumor and the lymph node metastases can be delineated, although the image quality suffers from an increasing noise level. In this patient, no severe photopenic artifacts were observed.

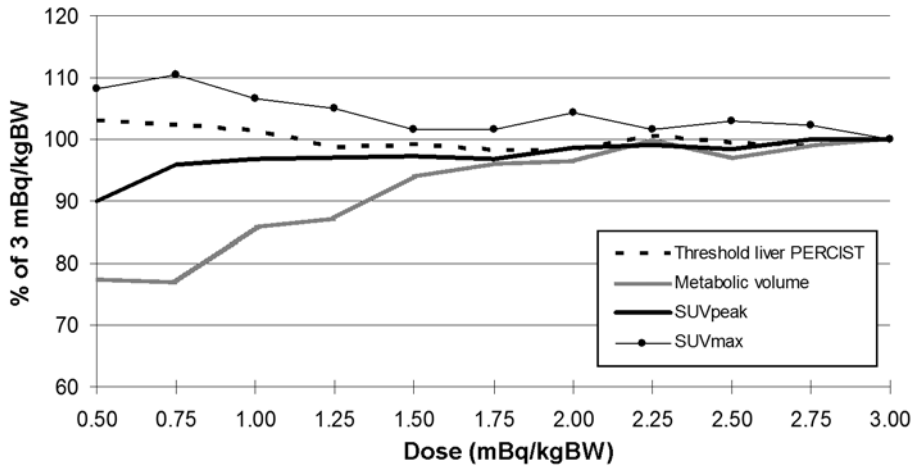

Figure 4: Overview of the SUVpeak, the SUVmax, the metabolic volume of the evaluated lesions and the liver threshold according to PERCIST. The liver threshold also includes the standard deviation; therefore, a decreasing SUV can be compensated by increasing standard deviations which leads to a relative stable behaviour. The trend of the metabolic volume shows an inverse linear correlation to the trend of SUVmax with R^2=0.27.

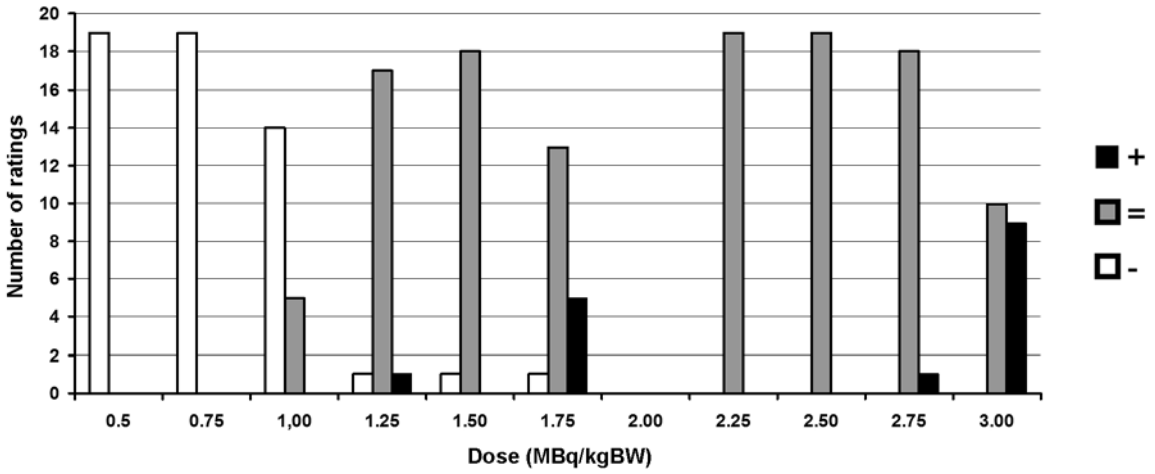

Figure 5: Overview of the evaluation of the subjective image impression compared to a dose of 2 MBq/kgBW. (+) means better image quality, (=) comparable image quality and (-) worse image quality.

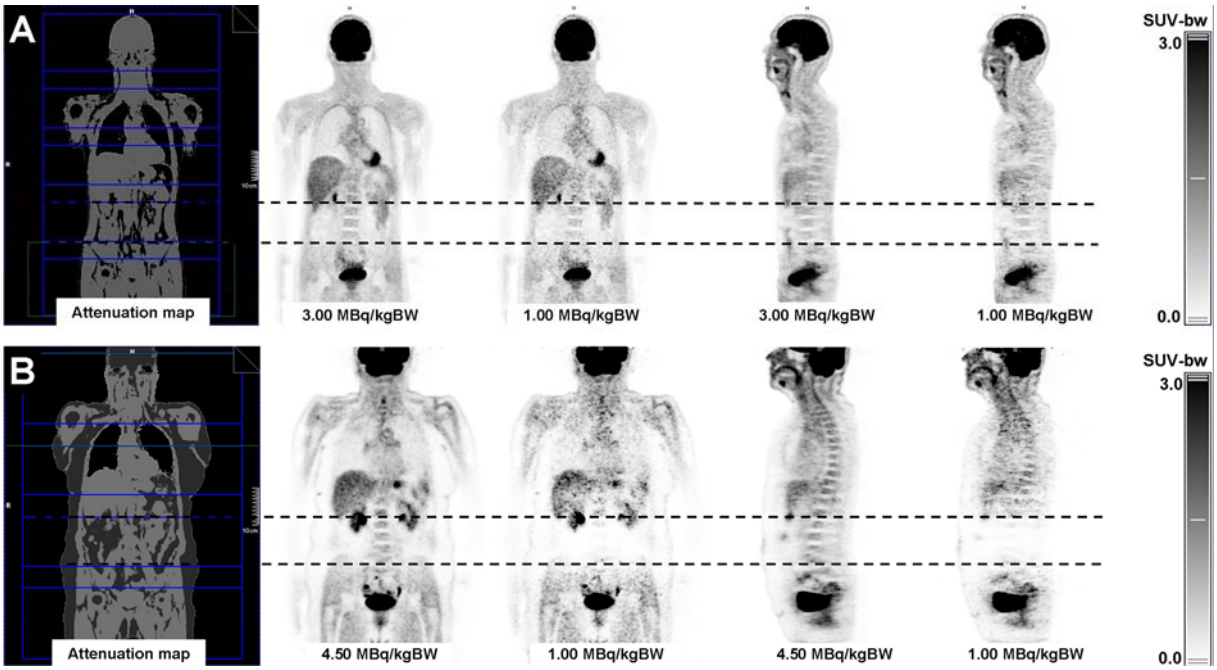

Figure 6: Two examples of a simulated dose reduction in 18F-FDG-PET. The upper line (A) shows an example of a 43 y/o male patient; even with a dose of 1 MBq/kgBW the image impression is good. In the lower line (B), a pronounced photopenic area was observed in the abdominal region at lower doses. However, this region was located completely within one bed position (dotted line).

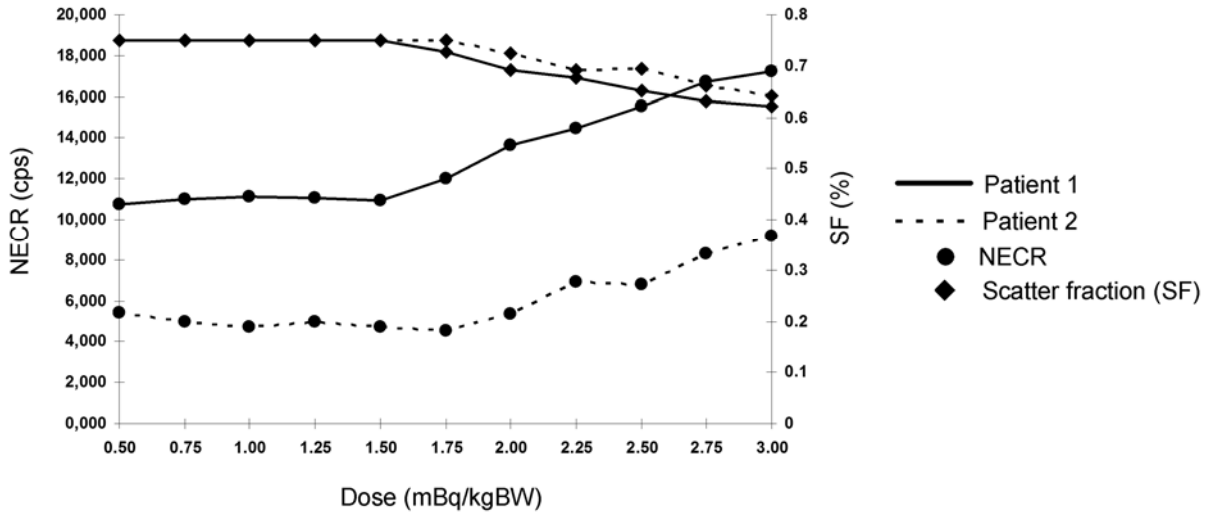

Figure 7: NECR and estimated SF of the bed position covering the abdomen/pelvis region of two patients with large photopenic artifacts in the abdominal region. Note the break in the curves of NECR when the estimated SF reaches the internal maximum of 0.75, limited by the scanner software.