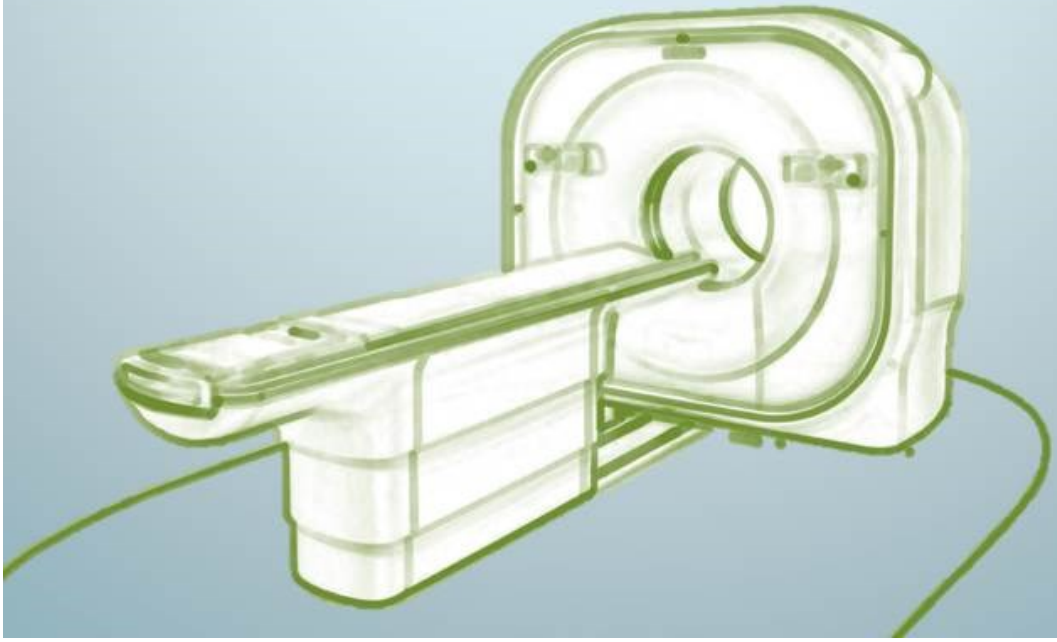
Ellis Winters

Digital PET-CT in oncology: image quality optimization and diagnostic comparison with conventional PET



Title page

E.G. Winters

Defense:	November 5th 2019, 14:00 Technohal TL 1133, University of Twente, Enschede
Master track:	Technical Medicine, Imaging and Intervention, University of Twente
Location:	Isala Hospital, Department of Nuclear Medicine, Zwolle
Title:	Digital PET-CT in oncology: image quality optimization and diagnostic comparison with conventional PET

Graduation committee

Chairman and Technological supervisor UT	Prof. dr. ir. C.H. Slump
Medical supervisor Isala	Dr. P.L. Jager
Technological supervisor Isala	Dr. J.A. van Dalen
Process supervisor	Drs. E.M. Walter
Daily supervisor and extra member	D. Koopman, MSc
Extern member	Dr. E. Groot Jebbink

Preface

Beste lezer,

Dit is de scriptie 'Digital PET-CT in oncology: image quality optimization and diagnostic comparison with conventional PET', geschreven in het kader van mijn klinische afstudeerstage van de opleiding Technical Medicine van Universiteit Twente. Ik heb dit onderzoek uitgevoerd op de afdeling Nucleaire Geneeskunde van Isala Zwolle, van Januari 2019 tot en met November 2019.

Deze stage ben ik erg goed begeleid door verschillende personen, en die wil ik bij deze hartelijk bedanken. Allereerst wil ik Daniëlle Koopman bedanken voor de intensieve begeleiding die ze heeft geboden. Je was altijd heel positief en je hebt me veel geleerd over alles wat met onderzoek doen te maken heeft, ook op persoonlijk vlak. Ik wil ook Jorn van Dalen bedanken voor de duidelijke feedback die je altijd gaf en de prettige manier waarop je dat deed. Ook wil ik Piet Jager bedanken voor je bereidheid om alle klinische vragen die ik had te beantwoorden en ook voor de positieve en enthousiaste begeleiding. Kees Slump wil ik graag bedanken voor de sturing tijdens het onderzoeksproces. Tijdens de stage heb ik meerdere intervisiebijeenkomsten gehad, waarbij ik begeleid ben door Elyse Walter. Ik wil je bedanken voor je luisterend oor en de heldere analyses die mij vaak beter inzicht gaven in mijn sterke punten en valkuilen, en hoe ik daar ook mijzelf in kan sturen. Daarnaast wil ik Erik Groot Jebbink hartelijk bedanken om de tijd te nemen om buitenlid te zijn van de afstudeercommissie. Ik wil ook graag Henk Stevens bedanken voor alle tijd die hij gestoken heeft in het visueel beoordelen van de scans.
Tijdens deze stage heb ik een goede tijd gehad op de afdeling, vooral met Tessa Gerritse, Maaike Dotinga en Sabine Koenders, bedankt daarvoor!

Ik hoop dat u mijn afstudeerthesis met veel plezier zult lezen.

Ellis Winters
Zwolle, 24 oktober 2019

Summary

^{18}F-FDG PET plays an important role in oncology for diagnosing, staging and therapy-response monitoring. Using ^{18}F-labeled glucose, tissues with high glucose-metabolism are visualized, a common characteristic for tumors. Recently introduced PET-scanners using Silicon photomultipliers with digital readout have a better timing resolution and photon detection accuracy than conventional PET (cPET).
The aims were to compare diagnostic outcome of digital PET (dPET) to cPET, secondly, to illustrate difficulties in the balance between sensitivity and specificity for cPET and dPET and lastly to determine an image reconstruction providing the best small lesion detectability without noise amplification for FDG-PET using dPET.

We scanned 80 patients with proven cancer on dPET (Vereos, Philips) and cPET (Ingenuity, Philips) using high-resolution reconstructions to compare diagnostic outcome. We evaluated SUV_{max}, SUV_{mean}, lesion-to-background ratio (LB_{ratio}) and metabolic tumor volume (MTV) in up to five lesions per patient and noise-levels in the liver. We determined ROC-curves, cut-off values and accuracy. TNM-staging was assessed by two NM-physicians. With dPET, uptake values and LB_{ratio} increased with 15% and 18% compared to cPET, MTV decreased with 7%. Malignant lesions showed higher SUV_{max} and LB_{ratio} than benign lesions ($p<0.05$). Moreover, we found higher AUCs, better characterization performance and TNM-upstaging in 4/26 patients with dPET. More research is necessary to assess the effect on different tumor types.

For illustration of difficulties in the balance between sensitivity and specificity for cPET and dPET, we performed a case-study. We described a case of a 75-year old male. On cPET, esophageal cancer was found and a lesion in Th6 was marked as degenerative. He was treated curative for esophageal cancer. Several months later he had bone metastasis. By using dPET, it could have been that the Th6 lesion was reported as malignant, leading to a different treatment plan. For other patients, it is recommended to use both dPET and cPET, to obtain best possible diagnostics.

For determination of an OSEM-based image reconstruction with the best small lesion detectability and acceptable noise level for FDG-PET using dPET, we performed a phantom study and a patient study. In the phantom study, highest noise was found for 51 updates without filtering, decreasing with 22% for 39 updates ($p<0.01$). In the patient study (n=24), noise-level decreased with 13% for 39 updates ($p<0.01$). For all lesions (n=80, size:0.1-8.1mL), SUV_{mean} and SUV_{max} were similar between updates. Visual analysis showed no significant preference. Adequate image quality for a dPET-system using OSEM requires a selected number of updates. Lowest noise-level without impairing small-lesion detectability with acceptable image quality can be obtained using 2x2x2mm^3 voxels with 3 iterations x 13 subsets (39 updates) without post-smoothing filter.

Abbreviations

BG	Background
COV	Coefficient Of Variation
cPET	Conventional Positron Emission Tomography
CT	Computed Tomography
dPET	Digital Positron Emission Tomography
EANM	European Association of Nuclear Medicine
EARL	EANM Research Ltd.
FDG	Fluordeoxyglucose
FOV	Field Of View
FWHM	Full Width At Half Maximum
IQ	Image Quality
LB_{ratio}	Lesion-to-Background ratio
LYSO	Lutetium-yttrium oxyorthosilicate
METC	Medical Ethical Committee
MRI	Magnetic Resonance Imaging
MTV	Metabolic Tumor Volume
NM	Nuclear Medicine
NEMA	National Electrical Manufacturers Association
OSEM	Ordered Subset Expectation-Maximization
PET	Positron Emission Tomography
PMT	Photomultiplier Tube
PSF	Point Spread Function
PVE	Partial Volume Effect
QC	Quality Control
RC	Recovery coefficient
RC_{max}	Maximum activity concentration Recovery Coefficient
RC_{mean}	Mean activity concentration Recovery Coefficient
ROC	Receiver Operating Characteristic
ROI	Region Of Interest
SD	Standard Deviation
SiPM	Silicon Photomultipliers
SPECT	Single-Photon Emission Computed Tomography
SUV_{max}	Maximum standardized uptake value
SUV_{mean}	Mean standardized uptake value
TNM	Tumor Node Metastasis staging method
ΔT	Time between FDG administration and start PET acquisition
Δt	Time between the start of the first scan and the start of the second scan
t_1	Time per bed position for the first scan
t_2	Time per bed position for the second scan
$T_{\frac{1}{2}}$	Half-life of a radioactive tracer
TOF	Time-Of-Flight
US	Ultrasound
VOI	Volume Of Interest

Thesis outline

This thesis is divided into three topics, related to diagnostic outcome and image improvement of digital PET technology.

First, a general introduction about PET in oncology, PET technology and the PETPET study, and a description of the three main goals is described in Chapter 1. Chapter 2 describes the results of a part of the patient population participating in the PETPET study, in which we analyzed the differences between conventional PET and digital PET regarding lesion detectability and TNM staging. Chapter 3 shows a case report of a patient with different lesion detectability between conventional PET and digital PET. Chapter 4 describes a method to improve image quality by reducing noise without impairing signal, by adjusting the reconstruction settings.

Table of contents

Chapter 1 – Introduction

PET imaging in oncology

Imaging technologies support a personalized approach for cancer care [1]. Early recognition of anatomical changes and physiologic behavior can lead to improved tumor diagnosis and management [1, 2]. Several imaging techniques are used for assessing tumor properties, including Magnetic Resonance Imaging (MRI), ultrasound (US), Computed Tomography (CT), Single-Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) [1]. PET plays an important role in oncology for diagnosing and staging of cancer as well as for therapy response, as it images whole body metabolism [1, 3, 4]. It is a non-invasive technique for quantitative assessment of tumor behavior, often combined with CT for anatomical information and attenuation correction. It is used for staging or restaging patients with several types of cancer, as it is considered as a safe, accurate and reproducible technique [2, 5]. By using a radioactive tracer coupled to glucose, tissues with high glucose metabolism can be visualized [5] as shown in Figure 1. High glucose metabolism is a common characteristic for tumors [6]. After imaging is performed, staging or restaging of the tumor is performed, based on the TNM-system. It evaluates the extent of tumor invasion (T), number and location of lymph nodes (N) and the presence and location of metastases (M) [7]. To further investigate the origin and extent of tumor tissue, pathology is often performed by means of surgery or biopsy, resulting in a pathological TNM (pTNM) staging. Based on TNM-staging, a patient-specific treatment plan is made.

PET technique

^{18}Fluorine-fluordeoxyglucose (^{18}F-FDG) is used as a radioactive tracer for detecting malignancies, based on increased glucose uptake in tumors [5]. The principle of PET technology is the detection of 511 keV photons originating from positron-electron annihilation events. Patients are intravenously injected with ^{18}F-FDG and have to rest for 60 minutes, in order to let the radiotracer accumulate in tissues with high glucose uptake. To prevent uptake in muscles, patients are not allowed to move. The ^{18}F nucleus is instable, resulting in radioactive decay with positron emission as shown in Figure 2. Those positrons will collide with electrons in surrounding tissue and an annihilation will occur, producing two 511 keV photons at approximately 180° to one another [6].

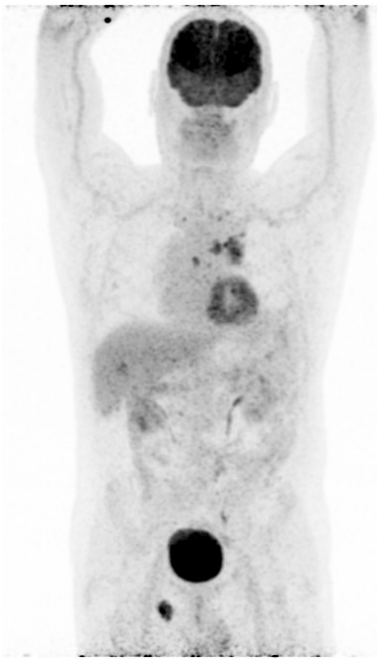


Figure 1 Example of a ^{18}F-FDG PET scan (Ingenuity TF, Philips) of a male patient with metastasized lung cancer. Primary tumor is located in the left lung, with mediastinal lymph node metastasis. Increased uptake in right ischium. Physiological glucose uptake in brain, myocardium and bladder. Data are from Isala Hospital, Zwolle, The Netherlands.

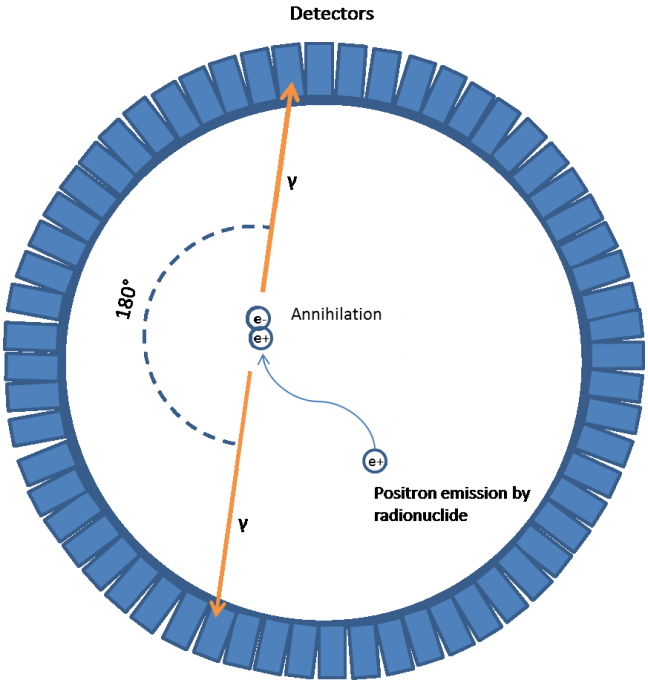


Figure 2 Annihilation of a positron-electron pair, resulting in two photons, detected by the detector ring.

Photons detected simultaneously on the detector ring surrounding the patient will be taken into account to calculate the location of the annihilation. After entering the detector, the signal will first be translated into visible light by scintillator crystals and after that into an electrical signal using photo-multiplier tubes (PMT) [8]. A CT scan is often combined with the PET scan to assess anatomy. Besides, the CT is necessary for photon-attenuation correction.

Accurate detection of the primary tumor, lymph nodes or metastasis is essential for determining therapy management. Therefore, FDG uptake should be as representative as possible for lesion metabolism in those lesions [9]. For accurate quantitative measurements of FDG uptake in small lesions, sufficient spatial resolution of the PET scanner is required. Over the past years, many improvements have been made in PET technology to enhance lesion detectability [3]. One of those improvements is Time-Of-Flight (TOF) technology. This comprises that the determination of the location of the positron-electron annihilation event is based on the detection times of the two emitted photons and the velocity of light [10]. The implementation of TOF technology results in a higher signal-to-noise-ratio (SNR) and therefore better small lesion detection [10–12]. However, the advantage of the better timing resolution that TOF can offer is limited when using current conventional PET/CT systems. This is induced by the PMTs. Those tubes are larger in size than the detectors, leading to uncertainty in determining the location of the annihilation [10]. Besides, PET systems using PMTs have a low timing resolution of 400-540 ps, while PET systems with SiPMs have a timing resolution of 345-385 ps [13]. Recently introduced PET scanners using SiPMs with digital readout have a better timing resolution and photon detection accuracy as compared to conventional PET [8]. By using SiPMs, spatial resolution can be improved, resulting in better small lesion detectability [3]. The diagnostic outcome could improve, and therefore higher sensitivity and specificity and more reliable staging.

For both conventional (cPET) and digital PET (dPET), the electrical signal produced by the PMTs or the SiPMs has to be converted into images using reconstruction techniques. The reconstruction method that is commonly used is ordered subset expectation-maximization (OSEM) [14]. The main principle of the OSEM reconstruction method, is the partition of the data into subsets and using several iterations. With each iteration, an image estimation is projected and adjusted to reach an image that satisfies regarding quantitative accuracy and noise level. By using subsets of the whole dataset, image reconstruction accelerates [14, 15]. With increasing number of subsets and iterations,

quantitative accuracy increases. However, the noise level will increase as well [13]. For optimal image quality, the iterative process is often terminated early and smoothing filters are added to reduce the noise [14].

Other improving technologies are based on new reconstruction methods. Image voxel sizes of 4x4x4 mm^3 are most common in current practice, but this is sensitive for the Partial Volume Effect (PVE) [16]. It has been shown that with smaller voxels (2x2x2 mm^3), the SNR increases and small lesion detection can be improved [16]. Nonetheless, a disadvantage of the use of smaller voxels is the increased amount of noise due to the lower amount of counts per voxel [11]. The noise level can be decreased by adjusting the reconstruction method or by adding post-smoothing filters [16].

PETPET study

In the department of Nuclear Medicine, Isala hospital, Zwolle, a conventional PET/CT-scanner (Ingenuity TF) is available, as well as a digital PET/CT system (Vereos), both developed by Philips Healthcare. In order to investigate the clinical added value of dPET, a head-to-head comparison study of both systems is performed. The primary aim is to investigate how the diagnostic outcome of dPET is compared to the diagnostic outcome of cPET. The second aim was to evaluate the image quality of both scanners [17]. Patients included in this study are referred to the department for (re)staging of lung-, breast- or esophageal cancer. Patients with other types of cancer are only included when they are referred for primary disease staging. Informed consent was obtained from all participating patients and the study was approved by the Medical Ethical Committee of the Isala hospital (METC Isala, Zwolle). A patient population of 225 patients is scanned twice, both on the conventional and the digital PET. The scan order is randomized per scanning day to prevent possible bias. A dedicated dose protocol depending quadratically on patient's body weight is used to determine the dose per patient. The equation to determine this patient specific dose is

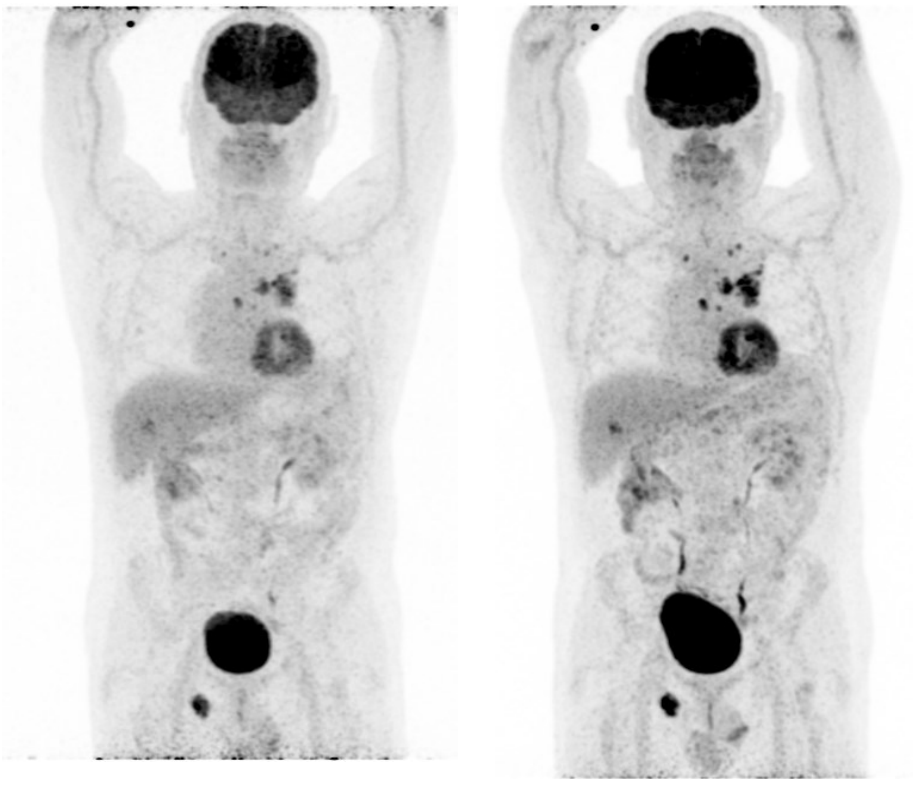


Figure 3 Example of a ^{18}F-FDG PET scan from a patient with metastasized lung cancer who participated in the PETPET study. Left scan was made on cPET (ΔT = 62 min), the right scan was made on dPET (ΔT = 83 min).

$$A \cdot t_1 = 6.22\, w^2, \quad (1)$$

with *A* the FDG activity administered in MBq, t_1 the time per bed position for the first scan (s) and *w* the body weight in kg. For patients ≤80 kg, t_1 is 72 seconds and for patients >80 kg, t_1 is 144 seconds. The first PET scan is made approximately 60 minutes after FDG-injection of the radioactive tracer (*ΔT*). To compensate for the delay between the injection and the second PET scan, the second scan has prolonged scanning time, according to:

$$t_2 = \left(0.5^{\frac{\Delta t}{T_{1/2}}}\right)^{-1} \cdot t_1\,, \quad (2)$$

with t_2 the time per bed position for the second scan (s), Δt the time between the start of the first scan and the start of the second scan (min), $T_{\frac{1}{2}}$ the half-life of the tracer (min) and t_1 the time per bed position for the first scan (s). Figure 3 demonstrates the PET scans of a patient who participated in the PETPET study.

Aims of the study

The first aim of this study was to investigate how the diagnostic outcome of digital PET compares to the outcome of conventional PET. The second aim of our study was to illustrate the difficulties in the balance between sensitivity and specificity in FDG-PET between cPET and dPET. The third aim was to determine an image reconstruction that provides the best small lesion detectability without noise amplification for FDG-PET using a dPET system.

Chapter 2 – Diagnostic outcome dPET compared to cPET

Abstract

Introduction With the introduction of PET using SiPMs with digital read-out (dPET), spatial resolution increases compared to conventional PET, with potentially improved lesion detectability. The aim of this study was to compare the diagnostic outcome of d PET to conventional PET (cPET) in patients with cancer regarding lesion detectability, TNM staging, sensitivity and specificity.

Methods We prospectively included 80 patients with proven cancer who were scanned both on the digital (Vereos, Philips Healthcare) and the conventional PET scanner (Ingenuity, Philips Healthcare), with a 2x2x2 mm^3 OSEM reconstruction. Participants were scanned in randomized order. We evaluated SUV_{max}, SUV_{mean}, lesion-to-background ratio (LB_{ratio}) and metabolic tumor volume (MTV) in the smallest lesions of each patient, with a maximum of five lesions per patient. For the uptake values, LB_{ratio} and MTV we determined the average difference ΔP from cPET to dPET, in which P can be replaced for each of the parameters. Noise levels were measured in the liver. We determined ROC curves using the final diagnosis based on histological, imaging or clinical follow-up information for each lesion, together with optimal cut-off values, sensitivity and specificity. TNM staging was performed on patients who were referred for initial staging (n=26/80) by two nuclear medicine physicians.

Results We analyzed 281 lesions, with 271 lesions measurable on dPET (96%) and 244 on cPET (87%). 86% of the lesions (242/281) were measurable on both PET scans. Of those lesions, 59% (143/242) were measured in the dPET second group and the other 41% (99/242) were detected in the dPET-first group. When we averaged over the dPET-first and dPET-second group, we found an increase from dPET to cPET of 15% for both SUVs, 18% for LB_{ratio} and a decrease of 7% for MTV. On cPET we found higher SUVs and LB_{ratios} for the dPET-second group as compared to the dPET-first group ($p = 0.04$ and $p = 0.01$, respectively). On dPET we found similar SUVs for both groups ($p=0.17$) and higher LB_{ratios} in the dPET-second group ($p<0.001$). Overall for all lesions, SUV_{max} was 61% (cPET) and 64% (dPET) and) higher for malignant lesions than for benign lesions, as well as for LB_{ratio} with 62% (cPET) and 65% (dPET) ($p < 0.05$). Furthermore, we calculated ROC curves of SUV_{max} and LB_{ratio} for both cPET and dPET, showing a significant difference between cPET and dPET regarding AUCs for SUV_{max} ($p = 0.02$), but not for LB_{ratio} ($p = 0.49$). Using optimal cut-off values to distinguish benign from malignant lesions, we found better characterization performance for dPET compared to cPET for SUV_{max}, but not for LB_{ratio}. In 4/26 patients (15%), we found TNM upstaging from cPET to dPET regarding lymph nodes or metastasis, with three cases from the dPET-second group. In 1/26 patients, we found TNM upstaging from dPET to cPET.

Conclusion
With the use of digital PET, uptake values and LB_{ratio} increased for dPET compared to cPET. With the use of digital PET, uptake values and LB_{ratio} increased and MTV decreased for digital PET compared to conventional PET. We found higher cut-off values for distinguishing benign and malignant lesions for dPET and significant higher AUCs and better characterization performance for dPET compared to cPET for SUV_{max}, but not for LB_{ratio}. Digital PET resulted in upstaging of 4 patients compared to conventional PET. More research is necessary to assess the effect on different tumor types.

Introduction

In oncology, FDG-PET/CT plays an important role in tumor staging and treatment response monitoring [4, 10]. Accurate staging is important, as treatment plans highly depend on the presence of any metastases [18]. However, FDG-PET/CT technology is associated with poor small-lesion detectability [18–21]. This is due to its low spatial resolution, leading to the PVE [21, 22]. Especially small lesions are prone to blurring by the PVE, resulting in inaccurate visualization of FDG uptake [22]. With the use of high-resolution reconstructions, in particular using small voxels of 2x2x2 mm^3, lesion detectability improves [16, 23].

With the introduction of digital PET, spatial resolution increases even more, because of implementation of SiPM [24–27]. Comparison studies of digital PET and conventional PET in patient data were performed previously [3, 4, 26, 28]. However, those comparisons were made between a 4x4x4 mm^3 voxel reconstruction on cPET and a 2x2x2 mm^3 on dPET [26]. As stated by Koopman et al. [29], this in an invalid comparison, as the difference between those two voxel sizes will already result in different uptake values [9, 16]. Moreover, no analysis was performed regarding sensitivity, specificity and TNM staging.

Therefore, the aim of this study was to compare the diagnostic outcome of dPET to cPET using a 2x2x2 mm^3 reconstruction in both systems, regarding lesion detectability and TNM staging together with sensitivity and specificity.

Methods

Study population

We prospectively included 80 patients with proven cancer who participated in the PETPET-study. Furthermore, only patients with at least one FDG positive lesion and a homogeneous liver on at least three slices were included. A sub-analysis of 26 patients who were referred for initial staging was performed to assess differences in TNM staging between digital and conventional PET. Patients with melanoma or thymoma were excluded.

All patients agreed on the use of their data for this study. Patients fasted for at least 6 hours before scanning and prior to intravenous injection of FDG, blood glucose levels were assessed to ensure a value below 15 mmol/L. FDG was administered according the PETPET dose protocol, depending quadratically on a patient's body weight. This protocol is described by equation 1 [30]. Scan duration for the first scan was 72 seconds or 144 seconds per bed position, for patients with body weight ≤80 kg and >80 kg respectively. For the second scan, scan duration was calculated according equation 2. Scan order was randomized for all patients, with 36 patients scanned first on dPET (dPET-first group) and 44 patients scanned first on cPET (dPET-second group). All scans were acquired with patients in supine position with the arms along their head.

Data acquisition

Data were acquired on a digital PET system (Vereos PET/CT, Philips Healthcare) and a conventional PET system (Ingenuity PET/CT, Philips Healthcare), with the use of Time of Flight (TOF). The PET detector ring of the dPET system consists of 18 detector modules, each with a 40x32 array of 4x4x19 mm^3 LYSO crystals combined with SiPMs. The dPET system is combined with a 128-channel CT scanner. Scan parameters for the CT scan were 120 kV tube voltage, dose modulation with an average tube current of 64 mAs (35-136 mAs range), slice collimation 64 x 0.625 mm, pitch 0.83 and rotation time of 0.5 s. PET data were reconstructed using 2x2x2 mm^3 voxels, with an OSEM reconstruction with 3 iterations and 17 subsets without Point Spread Function (PSF). The PET detector ring of the cPET system consists of 28 detector modules, each with 23x44 matrices of 4x4x22 mm^3 LYSO crystals combined with conventional PMTs. The cPET system is combined with a comparable CT system as the digital PET [23]. cPET data were reconstructed using 2x2x2 mm^3 voxels with a relaxation parameter of 0.6 and 3 iterations and 43 subsets without PSF.

Data analysis

Lesion detectability

On each PET scan, we obtained noise levels, Standard Uptake Values (SUV_{max} and SUV_{mean}) and Metabolic Tumor Volume (MTV) in mm^3. The SUV_{max}, SUV_{mean} and MTV were determined in a 3D ROI at 70% of the maximum pixel value, with the MTV the volume of a lesion with increased FDG uptake. All lesions were determined as a lesion if they were visible in a SUV-window from 0 to 5. A maximum of five lesions per patient was chosen to prevent bias from patients with many lesions. Lesions were included if they were smaller than or equal to 10 mL, corresponding with a diameter of 26 mm [31]. Noise levels were calculated using:

$$Noise = \frac{SD_{liver}}{SUV_{mean_{liver}}} \quad (3)$$

SD_{liver} and $SUV_{meanliver}$ were calculated by averaging nine regions of 900 mm^2 each in homogeneous areas in the liver, measured in three axial slices with three ROIs per slice.

If it was not possible to delineate a lesion with the 3D VOI at 70% of SUV_{max} without including background, we marked that lesion as 'unmeasurable'. If a lesion was not visible at all at one of both scans, we marked it as 'invisible'.

Semi-quantitative analysis for dPET compared to cPET

For equal comparison, in cases where lesions were visible on one of the two scans, we only used SUV_{max} of a ROI instead of the SUV_{max} of the VOI. The SUV_{mean} in the ROI of unmeasurable or invisible lesions was not used.

Next, we measured the SUV_{mean} of the lesion determined in a 2D ROI (ROI_1) which fitted the lesion at the slice with the highest FDG uptake in the lesion, $SUV_{meanlesion}$, and the SUV_{mean} in the surrounding background of the lesion, $SUV_{meanbackground}$. The $SUV_{meanbackground}$ was measured in a 2D ROI (ROI_2) with an area of 800 mm^2 larger than the area of ROI_1 and including the lesion, and corrected according to [9]:

$$SUV_{mean_{background}} = \frac{(SUV_{meanROI1} \cdot ROI_{1_{area}}) - (SUV_{meanROI2} \cdot ROI_{2_{area}})}{ROI_{1_{area}} - ROI_{2_{area}}} \quad (4)$$

After that, we calculated the lesion-to-background ratio, LB_{ratio}, defined as following:

$$LB_{ratio} = \frac{SUV_{max}}{SUV_{mean_{background}}} \quad (5)$$

Comparison between cPET and dPET was performed using the following equation, in which *P* can be replaced by SUVs, MTV and LB_{ratio}:

$$\Delta P = \frac{P_{dPET} - P_{cPET}}{P_{cPET}} \cdot 100\% \quad (6)$$

To correct for unequal sizes between groups, we used the following correction method, in which c_a is the correction factor for group a (dPET-first group), c_b for group b (dPET-second group, n_a the number of lesions in group a and n_b the number of lesions in group b, assuming a 50/50 distribution between both groups would be ideal:

$$c_a = \left(\frac{0.5}{\frac{n_a}{n_a+n_b}}\right) \; and \; c_b = \left(\frac{0.5}{\frac{n_b}{n_a+n_b}}\right) \quad (7)$$

The ΔP values for each lesion in both groups were multiplied by the corresponding correction factors. The average change between the dPET-first and the dPET-second group was calculated by averaging the corrected ΔP values.

All delineations were performed using IntelliSpace Portal (Philips Healthcare, 2015).

To assess differences regarding SUVs, LB_{ratio} and MTV for cPET compared to dPET, we used lesions measurable on both scans. We performed a sub-analysis for the dPET-first group and the dPET second group and we averaged the results of those groups using the correction factors according to Equation 7.

Performance analysis

Additionally, we collected follow-up information for the final diagnosis of each lesion. This was based on histological information, follow-up imaging (FDG-PET/CT, CT, MRI, ultrasound or X-ray) or clinical follow-up information. For the lesions with only clinical follow-up information, we assumed a lesion was benign when it was reported as 'probably benign' or when the lesion was not mentioned at all. It was assumed that the largest lesions in the lesion area were analyzed for histological assessment. We calculated optimal cut-off values based on ROC curves of SUV_{max}, SUV_{mean} and LB_{ratio}. Based on the ROC curves, we determined sensitivity, specificity and accuracy. The optimal cut-off value was the value corresponding to the sensitivity and specificity of the left upper corner of the ROC-curve, calculated using Pythagoras' algorithm [32]:

$$Optimal\ cut-off\ value = (1 - sensitivity)^2 + (1 - specificity)^2 \quad (8)$$

We used SPSS 24.0 for calculating ROCs and cut-off values. A sub-analysis of 26 patients was performed to assess differences in TNM staging between dPET and cPET. Two nuclear medicine specialists determined the TNM score based on both cPET and dPET. They were blinded for scanner type. The differences between cPET and dPET were analyzed. We used the newest version of the TNM system for each tumor type.

Statistical analysis

The Shapiro-Wilk test was used to assess normality of the data. We used the Wilcoxon signed-rank test to compare noise levels, MTV, SUVs and LB_{ratio} of dPET compared to cPET. The Mann-Whitney U test was used to assess significant differences between the dPET-first group and dPET second group and for analysis of differences between benign and malignant lesions. For comparisons of AUCs we used a chi-square test. We used the McNemar test for paired samples to test characterization performances for SUV_{max} and LB_{ratio} based on optimal cut-off values. A p-value below 0.05 was considered to indicate statistical significance. Data were displayed as mean ± standard deviation, and if applicable including the range (minimum and maximum value). We used SPSS 24.0 for Windows for those analyses.

Results

The characteristics of all patients and all lesions are shown in Table 1. We included 80 patients with a total of 281 lesions and an average of 4 ± 1 lesions per patient. Average lesion size was 13 ± 7 mm. Based on histological information, follow-up imaging or clinical follow-up information, 119 lesions were classified as benign and 162 lesions were classified as malignant. Overall, the noise level in dPET was slightly higher with 15.0%± 1.9% compared to 14.0%± 1.8% in cPET ($p < 0.001$).

Data analysis

Lesion detectability

Lesion delineation was possible in 271 lesions (96%) on dPET and in 244 lesions (87%) on cPET. The remaining lesions were visible but unmeasurable (n = 10 on dPET and n = 31 on cPET, with 23 lesions measurable on dPET only) or not visible at all (n = 0 on dPET and n = 6 on cPET). 86% of the lesions (242/281) were measurable on both PET scans.

Semi-quantitative analysis for dPET compared to cPET

In the group of measurable lesions, 59% (143/242) were detected in the dPET second group and the other 41% (99/242) of the measurable lesions were detected in the dPET-first group. The average

Table 1 Patient- and lesion characteristics

Patient characteristics (n = 80)			
Gender	Female	41	
	Male	39	
Age (years)		65 ± 10	
Body weight (kg)		79.4 ± 13.1	
Glucose (mmol/L)		6.7 ± 5.5	
Administered FDG dose (MBq)		405 ± 93	
Cancer type	Lung cancer	30	
	Breast cancer	20	
	Esophageal cancer	10	
	Others	20	
Lesion characteristics (n = 281)			
Type	Primary lesion	34	(12%)
	Lymph node	145	(52%)
	Distant lesion	102	(36%)
Final diagnosis	Benign	119	(42%)
	Malignant	162	(58%)
		Benign	Malignant
Final diagnosis based on	Histological information	32 (11%)	75 (27%)
	Additional imaging	43 (15%)	45(16%)
	Clinical follow-up information	44 (16%)	42 (15%)
Scan order group	dPET-first	60 (21%)	54 (19%)
	dPET-second	59 (21%)	108 (38%)
Measurable with 3D VOI on	dPET	110	161
	cPET	87	157
Visible, but unmeasurable with 3DVOI on	dPET	9	1
	cPET	26	5
Invisible on	dPET	0	0
	cPET	6	0

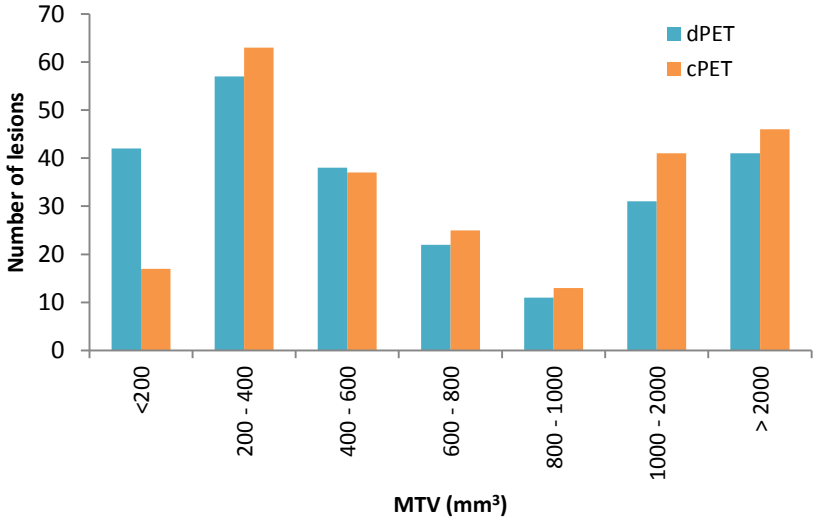


Figure 4 Histograms of all measurable lesions (n = 242), for cPET and dPET.

SUV_{max}, SUV_{mean}, LB_{ratio} and MTV for cPET and dPET in groups are shown in Table 2.
Overall, higher SUV_{max} and LB_{ratio} were found in the dPET second group compared to the dPET-first group for cPET results ($p = 0.04$ and $p = 0.01$, respectively) but for dPET results this was only the case for the LB_{ratio} ($p = 0.17$ and $p < 0.001$).
In the dPET-first group, for dPET we found averaged increases of 2% for both SUVs, a decrease of 1% for LB_{ratio} and an increase of 1% for MTV, compared to cPET ($p > 0.05$). In the dPET second group this increase was 33% for both SUVs, 44% for LB_{ratio} and a decrease of 19% for MTV ($p < 0.001$). When we average over both groups, with correction factors $c_a = 1.22$ and $c_b = 0.85$, we found an increase from dPET to cPET of 15% for both SUVs, 18% for LB_{ratio} and a decrease of 7% for MTV. MTV of measurable lesions was smaller on dPET ($p < 0.001$), as illustrated in Figure 4, with the lesion volume distribution per scan.

Performance analysis
Overall for all lesions, SUV_{max} was 61% (cPET) and 64% (dPET) and) higher for malignant lesions than for benign lesions, as well as for LB_{ratio} with 62% (cPET) and 65% (dPET) ($p < 0.05$), see Table 3. In benign lesions we found average increases from cPET to dPET of 16% and 21% for SUV_{max} and LB_{ratio} respectively ($p < 0.001$). Average increases in malignant lesions were 18% and 23% respectively ($p < 0.001$).
Furthermore, we calculated ROC curves of SUV_{max} and LB_{ratio} for both cPET and dPET, as illustrated in Figure 5 and Figure 6. For SUV_{max}, we found a significant higher AUC for dPET compared to cPET ($p = 0.02$), while for LB_{ratio} the AUCs were similar ($p = 0.49$). Corresponding optimal cut-off values for dPET showed increased values compared to cPET with 5.1 to 6.0 for SUV_{max} and 4.0 to 4.6 for LB_{ratio}, as summarized in Table 4. We found a significant better characterization performance for dPET compared to cPET for SUV_{max} ($p = 0.01$), but not for LB_{ratio} ($p = 0.67$).

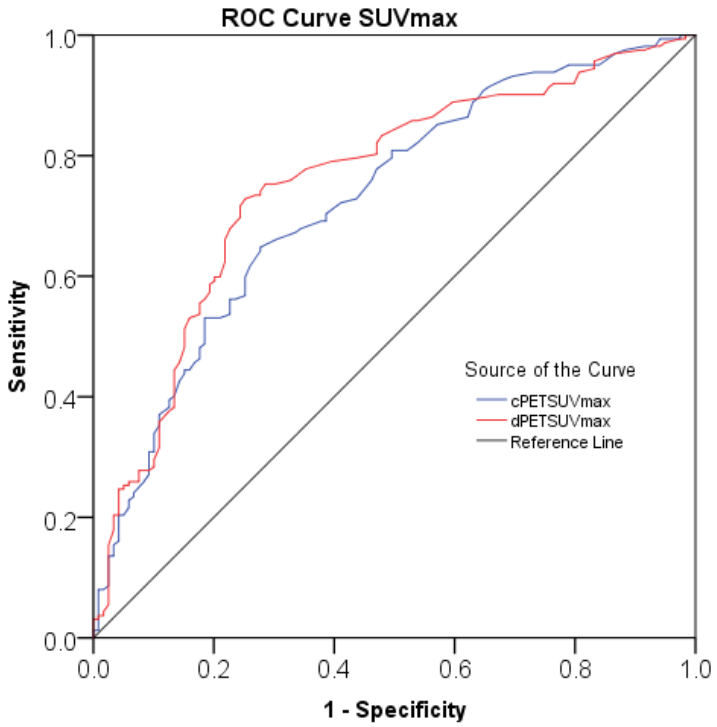


Figure 5 ROC curves using SUV_{max} for all lesions. AUC is 0.73 (95% CI 67% - 79%) for cPET and 0.76 for dPET (95% CI 70% - 81%). The AUC of dPET was higher compared to cPET (p = 0.02).

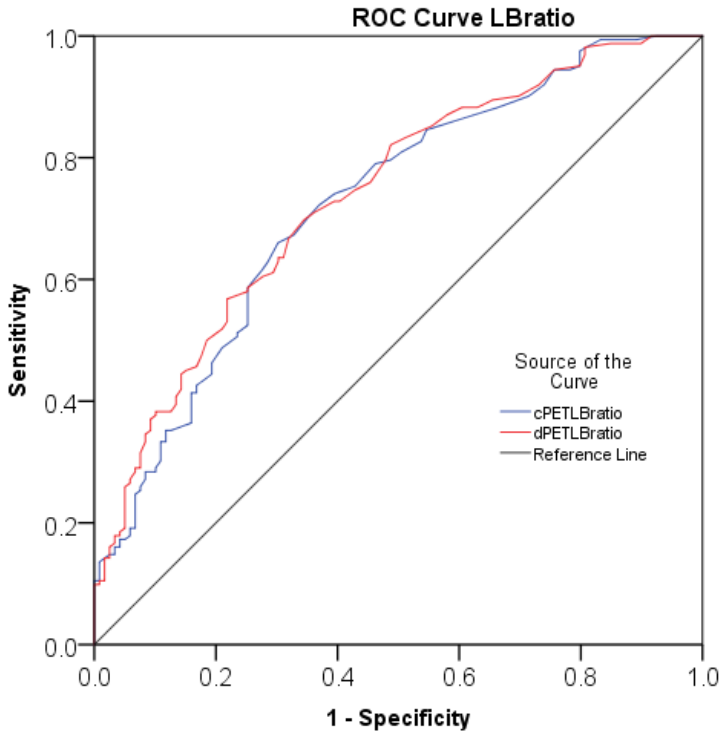


Figure 6 ROC curves using LB_{ratio} for all lesions. AUC is 0.73 (95% CI 67% - 78%) for cPET and 0.74 for dPET (95% CI 68% - 79%). The AUC of dPET was similar to cPET (p = 0.49).

Table 2 Lesion characteristics for cPET and dPET (only lesions measurable on both scans) *corrected for unequal number of patients per group

	SUV_{max}	SUV_{mean}	LB_{ratio}	MTV (mm^3)
dPET-first (n=99)				
cPET	7.8 ± 4.6 (1.8 – 25.5)	6.0 ± 3.6 (1.4 – 18.6)	6.2 ± 3.7 (2.2 – 19.1)	1513 ± 2147 (112 – 9928)
dPET	7.6 ± 4.0 (2.3 – 18.5)	5.8 ± 3.1 (1.8 – 14.0)	5.9 ± 3.5 (2.1 – 20.8)	1288 ± 1745 (112 – 9352)
ΔP	2% ± 18%	2% ± 16%	-1% ± 21%	1% ± 46%
dPET-second (n=143)				
cPET	6.5 ± 4.3 (1.5 – 30.6)	5.0 ± 3.5 (1.1 – 24.4)	5.2 ± 3.3 (1.9 – 18.5)	1261 ± 1590 (152 – 9056)
dPET	8.5 ± 5.5 (1.9 – 39.0)	6.6 ± 4.4 (1.5 – 31.2)	7.3 ± 4.7 (2.3 – 30.9)	1058 ± 1535 (88 – 8704)
ΔP	33% ± 20%	33% ± 21%	44% ± 37%	-19% ± 36%
Averaged change between dPET-first and dPET-second*	**15%**	**15%**	**18%**	**-7%**

Table 3 Lesion characteristics for benign and malignant lesions (for all lesions)

	SUV_{max}	LB_{ratio}
Benign (n = 119)		
cPET	4.7 ± 3.3 (1.2 – 25.5)	3.8 ± 2.2 (1.3 – 11.2)
dPET	5.5 ± 3.2 (1.9 – 19.3)	4.5 ± 2.4 (1.5 – 14.7)
ΔP	16% ± 3%	21% ± 12%
Malignant (n = 162)		
cPET	7.6 ± 4.8 (1.5 – 30.6)	6.1 ± 3.9 (1.7 – 19.1)
dPET	9.0 ± 5.3 (2.0 – 39.0)	7.5 ± 4.8 (2.4 – 30.9)
ΔP	18% ± 11%	23% ± 22%
Benign vs Malignant for cPET	**61% ± 45%**	**62% ± 81%**
Benign vs Malignant for dPET	**64% ± 68%**	**65% ± 97%**

Table 4 Optimal cut-off values for SUV_{max} and LB_{ratio} dPET first and second together (for all lesions)

		Optimal cut-off	Sensitivity	Specificity	Accuracy
SUV_{max}	cPET	5.1	65%	72%	68%
	dPET	6.0	73%	75%	73%
LB_{ratio}	cPET	4.0	66%	70%	67%
	dPET	4.6	70%	66%	68%

Visual analysis

In 4/26 patients (15%), we found TNM upstaging from cPET to dPET regarding lymph nodes or metastasis, with three cases scanned dPET second and one case scanned dPET-first, as summarized in Table 5. An example of one patient with mamma carcinoma is shown in Figure 7. With cPET, a T2N2aM1 tumor was found in the left mamma, but with dPET the staging was T3N3M1. This was due to the higher number of regional lymph nodes that was found on the dPET. This patient was scanned in the dPET-second group. In 1/26 patients, we found TNM upstaging from dPET to cPET. This patient was scanned on dPET first.

Table 5 Overview of five patients with different TNM staging between cPET and dPET.

Patient	Scan order	TNM score cPET	TNM score dPET	Comments
1	dPET-first	T2bN1M0	T2aN1M1b	Suspected lung cancer. pT2bN1PL0 in left upper lobe with one metastasis in N10. The lesion in the shoulder was indicated at dPET as metastasis in the sub cutis but at cPET as tendinitis, but there was no validation information available for that lesion.
2	dPET-second	T-N0M0	T2N1cM0	Mamma carcinoma in the right mamma with histological proven pT1cN0 (i+) (sn), so with isolated tumour cells in the sentinel node. Scan was made after a puncture in the right axilla.
3	dPET-second	T2N2aM1	T3N3M1	Adenocarcinoma mamma left. Axilla left metastasis. T4N1M1 with bone metastasis and lymph nodes, based on pathology before the PET-scan was made. N staging depended on number of regional lymph nodes, with more lymph nodes visible on dPET compared to cPET.
4	dPET-second	T1cN0M0	T3N2M0	Primary tumour located in pancreas, but the indication was abnormality in left upper lobe. It was uncertain whether the mass in the lung was a primary tumour or a metastasis. No histological validation, only CT proven growth between 2017 and 2019 and metastasis to liver, lung and peritoneum. TNM staging was performed under the assumption of primary lung cancer.
5	dPET-first	T3N3M0	T3N2M0	Barret oesophagus with histological proven adeno- carcinoma and lymph nodes behind aorta, cT3N1M1. N staging depended on number of regional lymph nodes, with more lymph nodes visible on cPET compared to dPET.

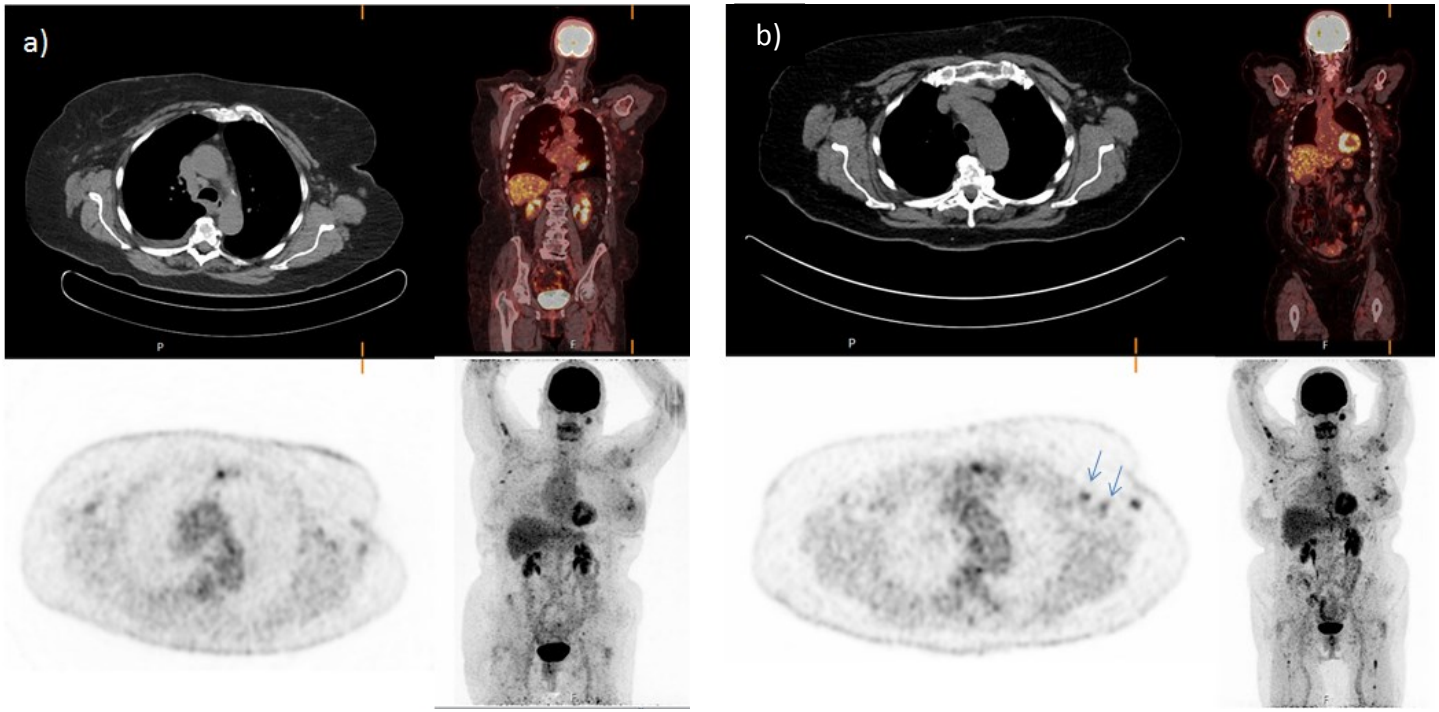


Figure 7 Example of a patient with mamma carcinoma in the left mamma with T2N2aM1 on cPET (a) and T3N3M1 on dPET (b). On dPET, more regional lymph nodes were found (blue arrows).

Discussion

In this study, we have demonstrated that the $SUV_{max,}$ SUV_{mean} and LB_{ratio} are higher with dPET compared to cPET and that the MTV is smaller for dPET compared to cPET. Furthermore, we detected 6 more lesions on dPET and 23 lesions were measurable on dPET only. Besides, with dPET we found significant higher AUCs and better characterization performance compared to cPET for SUV_{max}, but not for LB_{ratio}. Moreover, dPET resulted in TNM upstaging in 4 patients and TNM downstaging in 1 patient.

This study demonstrated improved small lesion detectability for dPET compared to cPET, with 29 lesions visible or measurable on dPET only. When we separated the data to scan order, 22 of those lesions were found in the dPET second group and 7 in the dPET-first group. We found 2 lesions measurable on cPET only, both in the dPET-first group. We found increased uptake values (15%) and LB_{ratio} (18%) for dPET compared to cPET in 242 FDG positive lesions when averaging the results of the dPET first and dPET-second group. On the other hand, the MTV showed a decrease of 7% for dPET compared to cPET, which is likely to be caused by the decrease of the PVE in dPET due to higher spatial resolution of this system. Baratto et al. found 37 more lesions on dPET with an increase in SUV of 53%, Nguyen et al. 8 extra lesions with a SUV_{max} increase of 36%, and the study of López-Mora found more lesions in 22 patients [3, 4, 26]. The study of Fuentes-Ocampo et al. showed 35% higher SUV_{max} values for dPET compared to cPET [26]. However, those studies compared the 2x2x2 mm^3 voxels of dPET with 4x4x4 mm^3 of cPET, and in Baratto et al. and Nguyen et al, they performed the dPET only after the cPET [3, 4]. Both aspects make the comparison between cPET and dPET invalid.

It is important to notice the difference between the dPET-first and dPET second groups. We noticed significant higher SUV_{max} and LB_{ratio} in the dPET second group regarding cPET measurements, but for dPET measurements only for the LB_{ratio}. Moreover, the increase from cPET to dPET was only significant in the dPET second group. Overall, this indicates that the delayed uptake between the two scans resulted in increased FDG uptake. Several groups investigated the effect of longer uptake time, revealing that this depends on tumor type [33, 34]. As in our study several tumor types were analyzed, it was difficult to quantitatively correct for the difference in uptake time. Therefore, we corrected for the effect of time delay by averaging the differences from cPET to dPET of both groups. This was allowed to do as the time delay between both scans was similar for both the dPET-first and dPET second group ($p = 0.6$).

When considering the results of benign and malignant lesions separately, we found averaged increases between cPET and dPET of 16% and 21% for SUV_{max} and LB_{ratio} respectively in benign lesions, and 18% and 23% in malignant lesions. Additionally, malignant lesions showed higher uptake values than benign lesions. This indicates that, despite scan order, both benign and malignant lesions show higher FDG uptake values with dPET. This is comparable to earlier studies on cPET systems [9, 35].

For SUV_{max} and LB_{ratio}, the optimal cut-off values were slightly higher for dPET (SUV_{max}: 6.0 and LB_{ratio}: 4.6) compared to cPET (SUV_{max}: 5.1 and LB_{ratio}: 4.0). The higher cut-off values for dPET are assumable, as SUVs and LB_{ratio} were larger in dPET. The optimal cut-off values are higher than in studies with conventional PET scanners [9, 36]. Moreover, the ROC curves for SUV_{max} were statistically significant different for cPET compared to dPET ($p = 0.02$). For LB_{ratio}, there was no difference between cPET and dPET regarding the ROC curves ($p = 0.49$). Furthermore, we found a significant better characterization performance for dPET compared to cPET for SUV_{max} ($p = 0.01$), but not for LB_{ratio} ($p = 0.67$). This indicates that LB_{ratio} was not that good at distinguishing benign from malignant lesions.

The sensitivity and specificity are not as high as mentioned in some other studies [9, 35, 37]. It is very likely that this is caused by the heterogeneity of the tumor type and the different validation methods that were used. As the ROC curves are based on all lesions, it could have been that the more

unreliable validation methods influenced the results. Most lesions (107/281) were proven benign or malignant by histology. However, it was not certain to what extent the biopsy was taken in the exact the same lesion as was seen on the scan. Other lesions (89/281) were proven by follow-up imaging. The remaining lesions were estimated by clinical follow-up, but this is less reliable as it depends mainly on reader's interpretation whether a lesion is not mentioned because they missed it or because they did not found it malignant.

In 4 out of 26 patients, we found TNM upstaging with dPET by visual analysis of two NM physicians. For the first patient, the extra lesion resulting in a difference from M0 to M1b but this was not validated. In the second patient, a difference from N0 to N1c was noticed, which was validated by pathology, as well as in the third patient with a difference from N2a to N3. For the fourth patient, the NM physicians performed the TNM staging under the assumption of primary lung cancer, as the indication was 'abnormality in left upper lobe'. However, this patient had suffered pancreatic cancer before, and it could have been that this abnormality in the lung was a metastasis. On the other hand, as both NM physicians made this assumption for both cPET and dPET, we can assume that the comparison was equal. In this case, they noticed a difference from N0 (cPET) to N2 (dPET).
In 1 out of 26 patients, we found TNM upstaging with cPET. This difference was from N2 on dPET to N3 on cPET, which depends on the number of regional lymph nodes. However, it could be that the lymph nodes visible on cPET in the abdominal region are counted for this as well, as they did not notice at as metastasis. Moreover, this patient was scanned on the dPET system first, so the time delay between dPET and cPET could have resulted in more lesions visible on cPET.

Limitations

Our study also has some limitations. The patient population was heterogeneous making it difficult to assess which effects can be related to dPET specifically, and which effects to the behavior of different tumor types, related to the prolonged FDG uptake. Therefore, for calculations between cPET and dPET for the dPET-first and dPET-second group, we corrected for the unequal group sizes by weighing. Another limitation could be the learning curve for delineating lesions, as the PET reader was not that experienced with lesion delineation in FDG PET scans. This could have resulted in delineations of increased uptake in areas with for example brown fat or muscle activity. However, all lesions without validation with pathology or imaging were validated by an experienced NM physician. For unmeasurable lesions, we only used SUV_{max} measured with an ROI instead of the SUV_{max} of the VOI. However, the ROI measurement was performed at the slice showing highest intensity in the lesion, so they should be comparable. Furthermore, the dPET reconstruction settings used in this study were not optimized yet. It is recommended to improve reconstruction settings to optimize noise level without impairing signal. Moreover, in this study we only measured lesion SUV_{max} and SUV_{mean}. It has been suggested that with the use of SUV_{peak} , the outcome could further improve [38]. Furthermore, we did not examine the differences between tumor types regarding cut-off values, so that could be an valuable addition to this study. Moreover, a study with separation of validation methods to histology and other validation methods could result in higher sensitivity and specificity for specific validation methods. At last, the improved lesion detectability of dPET could make that there will be more suggestive lesions, because more lesions are visible. This is important to take into account when using dPET. The results of this study can be used to improve visual evaluation about lesion detectability with dPET.

Conclusion

The aim of this study was to compare the diagnostic outcome of digital PET to conventional PET in patients with cancer regarding lesion detectability and TNM staging, sensitivity and specificity. With the use of digital PET, uptake values and LB_{ratio} increased and MTV decreased for digital PET compared to conventional PET. Moreover, we found higher cut-off values for distinguishing benign and malignant lesions for dPET. Furthermore, we found significant higher AUCs and better characterization performance for dPET compared to cPET for SUV_{max}, but not for LB_{ratio}. Digital PET resulted in upstaging of 4/26 patients compared to conventional PET. More research is necessary to assess the effect on different tumor types.

Chapter 3 - Impact of digital PET imaging on the detection of bone metastases in esophageal cancer staging: a case report

Introduction

The incidence of esophageal cancer in The Netherlands is 2500 patients per year with a five-year survival rate of 24% [39]. In esophageal cancer metastasis are most often present in lymph nodes, lung, liver, bones, adrenal glands and brain [40–42]. Bone metastases are found in 9% of all esophageal cancer patients [43]. The five-year survival rate for non-metastasized esophageal cancer is 24%, but it declines rapidly to 4% if distant metastasis are found [44, 45]. For whole-body staging of the disease FDG PET/CT is an important tool because of adequate sensitivity and specificity [46]. Neo-adjuvant chemotherapy followed by surgical resection is the standard of care for non-metastasized tumors [47, 48]. However, after curative surgery many patients still show recurrence of their disease [49].

Despite its importance FDG PET is not an perfect technique. False positive lesions on PET/CT can be of great impact. An important cause of false positive findings in FDG PET/CT is the accumulation of FDG in inflammatory tissue or fractures [50, 51]. On the other hand, false negative findings may lead to under diagnosis and errors in treatment choice. Both false positive and false negative findings are caused by interpretation by the readers. A gray zone is present between normal and increased uptake, leading to different conclusions. With the introduction of digital PET scanners, the performance of PET technique may improve. This case report describes the diagnostic evaluation and treatment of a patient with esophageal cancer and illustrates the difficulties in the balance between sensitivity and specificity in FDG-PET between conventional PET (cPET) and modern digital PET (dPET) scanners.

Case

A 75-year old man with a cardiac history presented in May 2018 at his general practitioner with difficulties in eating and the loss of 4-5 kilograms body weight in one month time. Besides, he had pain behind his sternum. He underwent gastroscopy and an esophageal tumor was found. Pathology analysis of a biopsy revealed a well-differentiated tubular adeno-carcinoma, pT2-3N0M0. Afterwards, a FDG-PET/CT scan was acquired for further staging. The scan was performed on a cPET system (Ingenuity, Philips Healthcare). For research purposes, a second scan was performed on a dPET system (Vereos, Philips Healthcare). Based on the conventional PET scan, the nuclear medicine physician noted focal activity in the esophagus (SUV_{max} = 6.37, SUV_{mean} = 3.54), clearly suspect for malignancy, as well as a suspect hotspot in the descending colon (SUV_{max} = 10.31, SUV_{mean} = 8.10), which could be a polyp. Furthermore, in the Th6 vertebra a small lesion with increased FDG-uptake was found (SUV_{max} = 3.69, SUV_{mean} = 2.90), but this was marked as degenerative. No other spots were noted. The diagnostic CT performed on the same day showed a space-occupying lesion at the transition from esophagus to stomach, without invasion in surrounding tissue and without lymphadenopathy. After colonoscopy, the intestinal uptake could be attributed to several benign polyps that were removed. Based on imaging and endoscopy, the diagnosis was T2-3N0M0 again.

Treatment

Therefore, the patient was accepted for a curative approach with chemo-radiation and surgery. At the end of May 2018, neo adjuvant chemo-radiation was started with 23 x 1.8 Gy on the tumor and lymph node areas combined with Taxol/Carboplatin. In August 2018, another CT scan was acquired. The tumor showed regression and no new lesions were seen. Afterwards, in September 2018, a radical resection was performed. In the resection specimen one malignant lymph node was found, resulting in a ypT3N1M0 staging.

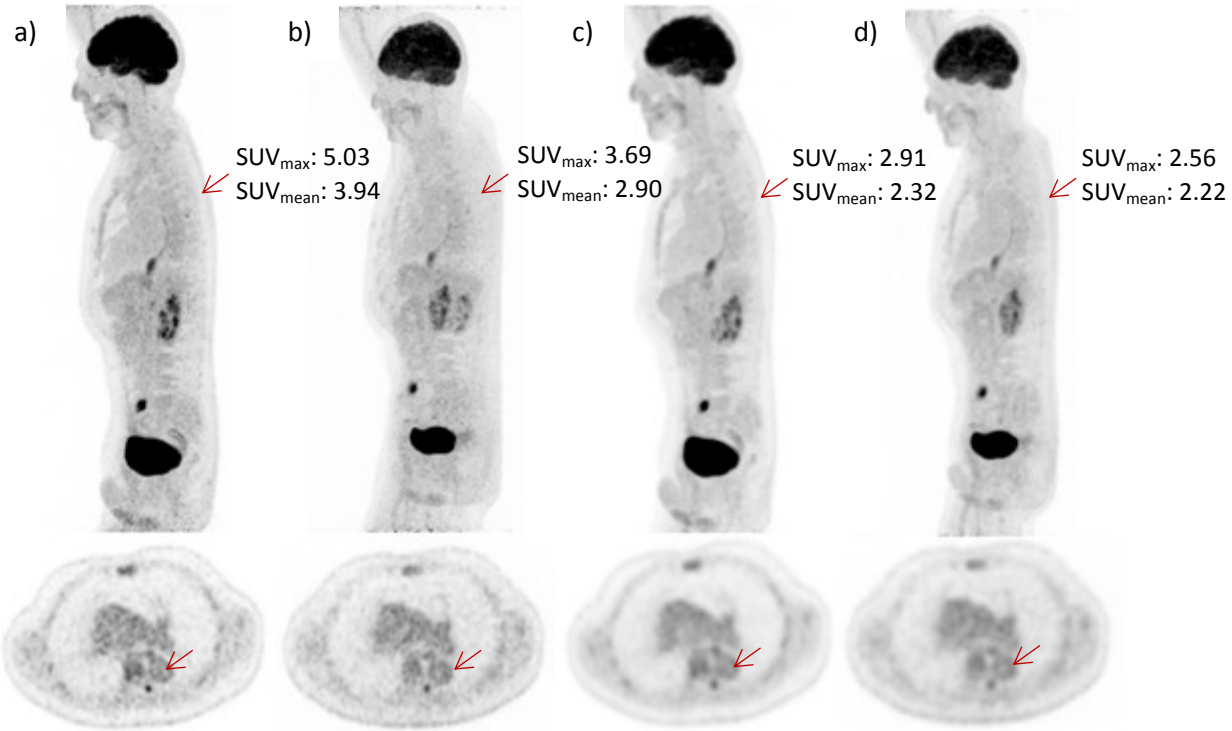


Figure 8 Lesion in Th6 in May. a) dPET 2mm^3, b) cPET 2mm^3, c) dPET 4mm^3, d) cPET 4mm^3. SUVs clearly increase on dPET with small voxels compared to dPET with large voxels and cPET.

Three months later the patient complained about a swelling in his left shoulder. An ultrasound suggested an intramuscular lipoma. A few days later he presented at the emergency department, because his wife found him unconscious and sweating. It appeared to be a hypoglycemia due to unknown origin. He also suffered palpitations and tachycardia. In January 2019, he underwent a PET scan to verify the lipoma lesion in the left shoulder and his ongoing weight loss. On this second PET-scan, which was a dPET as well, several bone metastases were found in the vertebrae, the ribs and the pelvic bones, Figure 9. He refused further therapy.

Discussion

The most unfortunate point in this case is the curative treatment that was given but within half a year metastatic disease was detected. On the first cPET-scan, no other malignancies than in the esophagus were reported. Looking in further detail at the initial cPET scan, a lesion in the vertebra at Th6 was marked as suspect but was not noted in the conclusion of the cPET report. At the second PET-scan six months later, several FDG-positive lesions were found, as summarized in Table 6. This demonstrates a substantial increase in number of lesions. Would it have been possible to detect these lesions or some of them on the first PET scan? In May 2018, the patient had actually been scanned twice, both on a cPET and dPET system. On both scans, there was a minor increased uptake in Th6. This was therefore interpreted as benign degenerative uptake, which is assumable as no other lesions were found and an uptake which is also suspect for degenerative was seen in the C2 vertebra. However, looking retrospectively at the dPET, the lesion in Th6 clearly had increased uptake (SUV_{max} = 5.03, SUV_{mean} = 3.94), and was clearly more sharp, especially on the scan with 2x2x2 mm^3 voxel size, as shown in Figure 8, a and b. On the scan with 4x4x4 mm^3 voxels , the lesion again had a low intensity, Figure 8, c and d. However, for the 2x2x2 mm^3 voxels , any nuclear medicine physician would call such a lesion at least to be suspect, requiring further evaluation. It is worth mentioning that de dPET in May 2018 was performed as second scan, so the time between FDG injection and scan was prolonged with 30 minutes. However, we expect that the effect of time delay is only partly the cause of increased uptake and that the other part is due to the better image quality of dPET. So, if the dPET would have been the basis for the clinical evaluation, the lesion in Th6 would have been evaluated further, e.g. with MRI. In retrospect, this might have changed the treatment

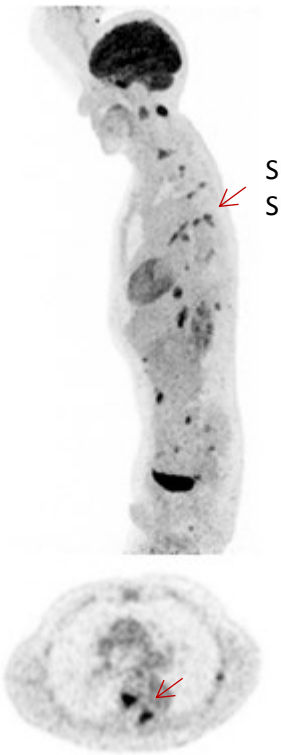


Figure 9 FDG-PET scan of the same patient in January with multiple metastasis as summarized in Table 6. Lesion in Th6 (arrow) shows increase in SUVmax and SUVmean compared to the scan in May. Scan was performed on dPET with 2x2x2 mm^3 voxels.

plan to palliative treatment only, rather than the very demanding chemo-radiation-surgery approach. It is easy to retrospectively consider, but in this case we think the dPET hotspot was significant. In addition, both benign and malignant lesions have increased contrast and sharpness on dPET, but the increase in malignant lesions may push the contrast over the visual limit of readers and make the lesions suspect, which could have led to further evaluation.

Conclusion

It is likely that if dPET rather than cPET would have been the base for the initial interpretation of this patients scan, the uptake in Th6 would have been interpreted as metastatic rather than degenerative. This could have led to a different treatment plan. For other patients, it may be valuable taking into account both the dPET and the cPET, if available, in order to obtain best possible diagnostics.

Table 6 Lesions detected per scan (based on the scan reports made by nuclear physicians)

May 2018	January 2019
Head and neck:	
Normal uptake brain, physiological uptake neck.	Normal uptake brain, physiological uptake neck.
Thorax:	
	Focal uptake mediastinum, below left bronchus
	Lymph node between tube stomach and vertebra
	Focal uptake in crux diaphragm
No lymph nodes	
Focal activity distal esophagus	
Abdomen:	
Left adrenal gland enlarged, right adrenal gland not suspect	Enlarged adrenal glands
	Focal uptake abdominal wall left
	Hot spot musculus psoas left
	Higher uptake in small intestines and sigmoid
Hot spot distal part colon	
Skeleton and soft tissues:	
	Intense high uptake musculus deltoideus left
	Vertebra C2
	Vertebra C7 (including processus transversus left)
Small hot spot processus transversus Th6/Th7 left, degenerative	Vertebra Th6 left dorsal (including processus spinosis Th5)
	Some ribs
	Processus spinosis vertebra L1 right
	Os ileum left
	Os ileum right
	Acetabulum right
	Muscles posterior to femur diaphysis right
	Slightly higher uptake was found at the left hip, degeneration or bursitis?

Chapter 4 – Optimizing Image Reconstruction on Digital PET

Abstract

Aim/Introduction: PET imaging with digital photon counting technology is associated with better spatial resolution, however, at the expense of increased noise-levels. We aimed to determine an image reconstruction that provides the best small lesion detectability without noise amplification for FDG-PET using a digital PET system.

Materials and Methods: We have performed a phantom study with the NEMA image quality phantom and micro-phantom (sphere diameters: 4-37 mm) using a digital PET system (Vereos PET/CT, Philips Healthcare). We used 2x2x2 mm^3 voxels and compared six different OSEM reconstruction settings (51, 45, 39, 34, 26 and 21 updates (=iterations x subsets)). For every setting, the effect of applying no filter and post-smoothing Gaussian filters of 2 and 4 mm was determined as well. Noise-levels were calculated by dividing the standard deviation to mean pixel value in a background area of the NEMA-phantom. Furthermore, we determined contrast recovery coefficients based on the mean activity (RC_{mean}) and maximum activity (RC_{max}) in all spheres in both phantoms. In the patient study, the reconstruction settings providing the lowest noise-levels without impairing RCs, were tested on FDG-PET scans of patients with breast cancer (n=5) and lung cancer (n=19). We measured SUV_{max} and SUV_{mean} of 80 FDG-avid lesions and noise-levels in the liver. Two NM physicians also visually evaluated these reconstructions. They indicated for each reconstruction whether the image quality was acceptable and they rated which reconstruction was preferred per patient.

Results: In the phantom study, the highest noise was found for 51 updates without filtering. Noise decreased with 14% for 45 updates and 22% for 39 updates ($p<0.01$). RC_{mean} and RC_{max} were similar for those reconstructions ($p>0.12$). By further reducing the number of updates, both noise and RCs decreased, especially for the smaller spheres. For all updates, applying a 2 mm filter resulted in decreasing noise levels (9-19%, $p\leq0.03$) and decreasing RCs (3-4%, $p<0.01$). A 4 mm filter led to a noise reduction of 14-39% ($p\leq0.03$) and a RC decrease of 10-12% ($p<0.01$). Consequently, on patient data we applied 51, 45 and 39 updates without filtering. Noise-level decreased with 6% for 45 updates and with 13% for 39 updates when compared to 51 updates ($p<0.01$). For all lesions (size: 0.1-8.1 mL), SUV_{mean} and SUV_{max} were similar between updates. Visual analysis showed no significant preference.

Conclusion: Adequate image quality for a digital PET system using the OSEM reconstruction algorithm requires to select a number of updates. The lowest noise-level without impairing small lesion detectability and with acceptable image quality can be obtained using 2x2x2 mm^3 voxels with 3 iterations x 13 subsets (39 updates) without post-smoothing filter.

Introduction

Positron Emission Tomography (PET) technology is increasingly used for diagnosing cancer [2, 5, 52]. For adequate small lesion detection, sufficient spatial resolution is necessary. PET imaging with digital photon counting technology is associated with better spatial resolution, due to the implementation of silicon photomultipliers instead of conventional photomultiplier tubes [13]. This results in increasing spatial resolution and effective sensitivity compared to analogous counting techniques [4, 8, 10, 13, 25, 27, 53].

Further optimization of scan quality can be obtained by adjusting the reconstruction settings. The main principle of the reconstruction method used for the current PET scanners, called ordered subset expectation-maximization (OSEM), is the partition of the data into subsets and using several iterations to estimate the best image [14, 15]. With an increasing number of subsets, the computation accelerates [54]. The product of subsets and iterations is referred to as the number of updates. However, with increasing the number of updates, the noise level is amplified as well [13]. For best lesion detectability, the iterative process is often terminated early and smoothing filters are added to reduce the noise [14].

The European Association of Nuclear Medicine guideline for FDG-PET contains recommendations for adequate repeatable and reproducible reconstruction settings [5]. However, those guidelines recommend to use a minimum voxel size of 3-4 mm, while it is proven to have better small-lesion detectability when using small voxels [16]. Small-lesion detectability is described as the highest SUV for the smallest lesions with a low noise level, because a high noise level is known to influence lesion detectability as well, making measurements less reliable [15]. For digital PET, no guidelines are available yet. Therefore, our aim was to determine an image reconstruction that provides the best small lesion detectability without noise amplification for FDG-PET using a digital PET system.

Methods

In order to optimize the current reconstruction settings for a digital PET system, we have performed a phantom study and a patient study, both with the use of a ^{18}F-FDG tracer.

Phantom study

Materials

For the phantom study we used the National Electrical Manufacturers Association (NEMA) NU2-2001 IQ phantom with sphere diameters of 10 to 37 mm, and a Micro Hollow Sphere phantom with sphere diameters of 4 to 8 mm. Both phantoms were prepared according to the IQRC QC procedure by EARL [5]. The NEMA phantom had an interior length of 18 cm, with a background compartment of 9,300 mL. The background compartment was filled with 3.41 kBq/mL ^{18}F-FDG and all spheres with 43.15 kBq/mL, resulting in a sphere-to-background ratio of 10:1. The background compartment of the Micro phantom had a volume of 103 mL, which was filled with 2.20 kBq/mL and the spheres with 22.30 kBq/mL, resulting in a sphere-to-background ratio of 10:1.

Data acquisition

All data was acquired with our department's digital PET system (Vereos PET/CT, Philips Healthcare), with the use of Time of Flight (TOF). The PET detector ring consists of 18 detector modules, each with a 40x32 array of 4x4x19 mm^3 LYSO crystals. Scan duration per bed position was 180 seconds for the NEMA phantom and 247 seconds for the micro phantom. The PET system is combined with a 128-channel CT scanner. Scan parameters were 120 kV tube voltage, dose modulation with an average tube current of 53 mA (37-94 mA range), slice collimation 64 x 0.625 mm, pitch 0.83 and rotation time of 0.5 s.

Data reconstruction

Data was reconstructed using 2x2x2 mm^3 voxels and six reconstruction settings. All reconstruction options were analyzed first on the NEMA phantom, without post-reconstruction Gaussian filtering and with a 2 mm and a 4 mm filter in full width at half maximum (FWHM) as well. For the MICRO phantom, only the reconstructions with acceptable RCs and noise levels based on the NEMA phantom were tested.

Data analysis

Quantitative measurements were performed on a dedicated workstation (IntelliSpace Portal, Philips Healthcare). We determined the Coefficient of Variation (COV) in the background compartment (BG) of the NEMA phantom according to:

$$COV = \frac{SD_{BG}}{Mean_{BG}} \tag{9}$$

with SD_{BG} the standard deviation of pixel values measured in three regions of interest (ROI) of approximately 900 mm^2 each in three subsequent slices the BG, and $Mean_{BG}$ the mean pixel value (*PV*) of the those ROI, in kBq/mL. Because of the small background compartment of the MICRO phantom, no noise level measurements were performed.

In all spheres of the NEMA phantom and the micro phantom, RC_{mean} and RC_{max} were calculated, according to the following equation:

$$RC = \frac{PV}{A_{true}} \tag{10}$$

PV is the mean or maximum pixel value in kBq/mL within a sphere, determined in a 3D VOI within 70% of the maximum pixel value, to obtain the best small lesion delineation. A_{true} is the injected activity in kBq/mL, corrected for scan time.

Patient study

Study population

For the patient study, we retrospectively included 24 patients who participated in the PETPET-study, a study to compare diagnostic outcome and image quality of a conventional PET system and a digital PET system. Patients included in our study underwent the first scan on the digital Vereos PET system (Philips Healthcare). Furthermore, patients should have at least one FDG positive lesion and a homogeneous liver on at least three slices. Lesions should be smaller than 10 mL to obtain only small lesions. All patients agreed on the use of their data for research. Patients fasted for at least 6 hours before scanning and prior to intravenous injection of FDG, blood glucose levels were assessed to ensure a value below 15 mmol/L. FDG was administered according a dose protocol depending quadratically on a patient's body weight. This protocol is described by:

$$A \cdot t = 5.0 \cdot w^2, \tag{11}$$

with A the FDG activity to administer (in MBq), t the time per bed position (in seconds) and w the patient's body weight (in kilograms) [30]. Scan duration was 72 seconds or 144 seconds per bed position, for patients with body weight ≤80 kg and >80 kg respectively. All scans were acquired with patients in supine position, using the same PET-CT scanner settings as for the phantom study.

Data reconstruction and analysis

We selected the reconstructions with the lowest COV and highest RCs, based on phantom data, to obtain a selection of reconstructions with in patient scans the highest achievable signal to background ratio. We obtained noise levels, Standard Uptake Values (SUV_{max} and SUV_{mean})and Signal-to-Noise Ratios (SNR_{max} and SNR_{mean}) in 24 FDG-PET scans of breast- and lung cancer patients. A maximum of five lesions per patient was chosen to prevent bias from patients with many lesions. Lesions were included if they were smaller than or equal to 10 mL. A sub analysis for lesions with a volume up to 4 mL was performed to investigate the effects particularly on small lesions. The 4 mL was chosen according to literature, corresponding with a lesion diameter of 2 cm [13]. Noise levels were determined using:

$$Noise = \frac{SD_{liver}}{SUV_{mean_{liver}}} \tag{12}$$

SD_{liver} and $SUV_{meanliver}$ were calculated by averaging nine regions of 900 mm^2 in homogeneous areas in the liver, measured in three axial slices with three ROIs per slice.

Signal intensity in the lesions was assessed by SUV_{max} and SUV_{mean} measurements, as well as mean and maximum SNR, defined by:

$$SNR_{mean} = \frac{SUV_{mean}}{SD_{liver}} \quad \text{and} \quad SNR_{max} = \frac{SUV_{max}}{SD_{liver}} \tag{13}$$

Each update combined with all three filter options were compared to a reconstruction setting with such a number of updates to ensure maximal SUVs. Comparison was performed using the following equation, in which *P* can be replaced by noise levels, SUVs and SNRs:

$$\Delta P = \frac{P_{test}}{P_{reference\ reconstruction}} \cdot 100\% \tag{14}$$

Visual assessment was performed by two nuclear medicine physicians. The three updates performed on patient data were displayed side-by-side on a screen, in random order. The physicians were asked to answer the following questions:

1. Which reconstruction do you prefer?
2. Do you think the image quality of that reconstruction is acceptable?
3. Do you see differences between the three reconstructions?
4. If yes, what are those differences?

Statistical analysis

For both the phantom study and patient study, we used the Wilcoxon signed-rank test to compare noise level, RCs, SUVs and SNRs of the reconstruction settings and filtering options with the current reconstruction setting. A McNemar test was used to assess significant differences between both observers. The Chi-Square Goodness of Fit test was used for assessing significant differences between preferences of reconstructions. A p-value below 0.05 was considered to indicate statistical significance. Data is displayed as mean ± standard deviation, and if applicable, with the minimum and maximum value of the range. We used SPSS 24.0 for Windows for those analyses.

Results

An overview of the used reconstructions for both phantom data and patient data is shown in Figure 10. We started with the reconstruction settings recommended by the vendor of our digital PET system and after that we tested lower number of updates [55].

Phantom study

In the phantom study, a noise level of 15% was found for 51 updates without filtering. A relative decrease in noise level of 14% for 45 updates ($p < 0.01$) and 22% for 39 updates ($p < 0.01$) was found, as can be seen in Figure 11, Figure 12 and Figure 13. By lower number of updates, noise level decreased more, with a relative decrease up to 52% for 21 updates.

RC_{mean} and RC_{max} were similar for 51, 45 and 39 updates ($p > 0.12$). By further reducing the number of updates, both noise and RCs decreased, especially for the smaller spheres, as shown in Table 7 and Figure 14.

For all updates, applying a 2 mm filter resulted in decreasing noise levels (9-19%, $p \leq 0.03$) and decreasing RCs (3-4%, $p < 0.01$), Figure 12 and Figure 13. A 4 mm filter led to an averaged noise reduction of 14-39% ($p \leq 0.03$) and an averaged RC decrease of 10-12% ($p < 0.01$), Figure 12. For smaller spheres, the effect of filtering was substantial, with a decrease in RCs up to 27% for the 10 mm sphere, Table 8.

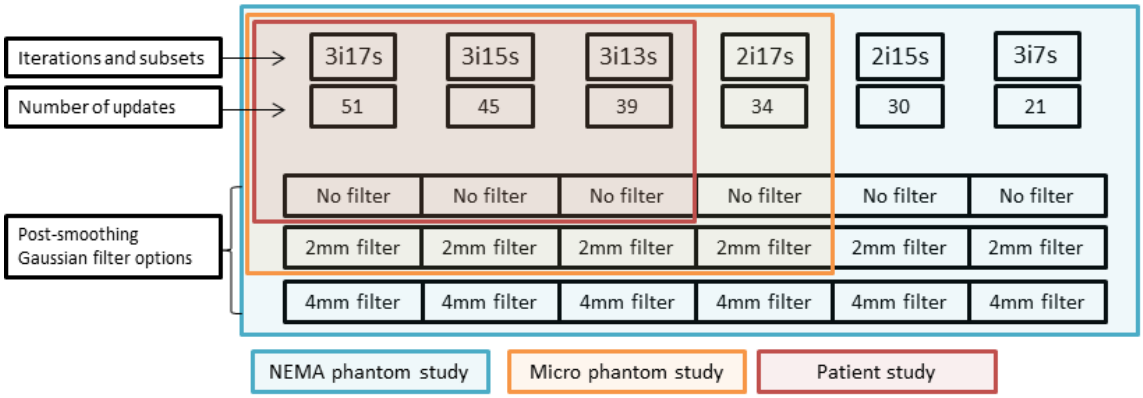


Figure 10 Overview of the used reconstructions. In blue the reconstructions used for the NEMA phantom study. In orange the reconstructions used for the Micro phantom study, we excluded 30 and 21 updates, as well as the 4 mm filter. In the patient study, in red, we included only 51, 45 and 39 updates, without post-smoothing Gaussian filtering.

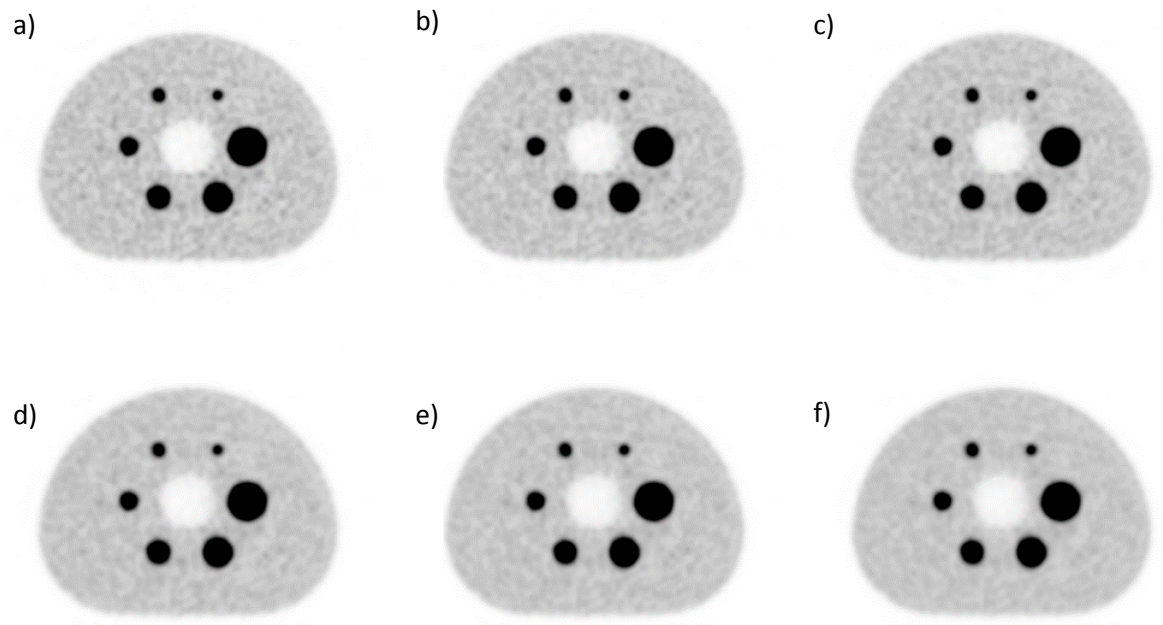


Figure 11 NEMA phantom results for different number of updates. From a to f: 51, 45, 39, 34, 30 and 21 updates. Those results are without post-smoothing Gaussian filtering. A reduction in noise can be seen at lower number of updates.

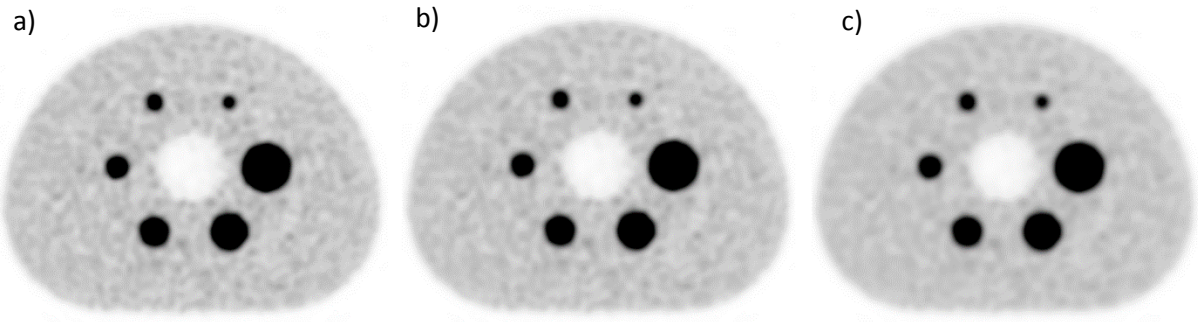


Figure 12 NEMA phantom results for three options for post-smoothing Gaussian filtering. From a to c: no filter, a 2 mm filter and a 4 mm filter. A reduction in noise can be seen when adding filtering, especially for the 4 mm filter.

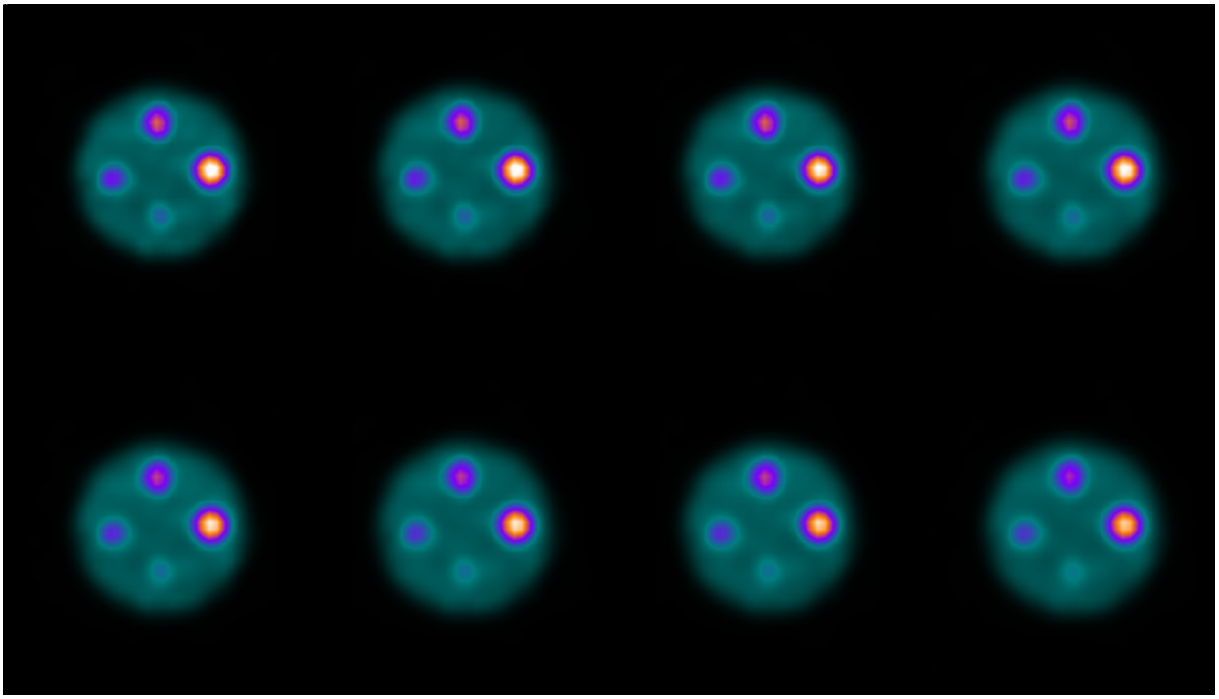


Figure 13 Results of the MICRO phantom. From left to right, the number of updates is 51, 45, 39 and 34. First row shows images without post-smoothing. Second row shows images with a 2 mm Gaussian post-smoothing filter. A slightly lower intensity is seen for the reconstructions with filtering, especially for the smallest sphere.

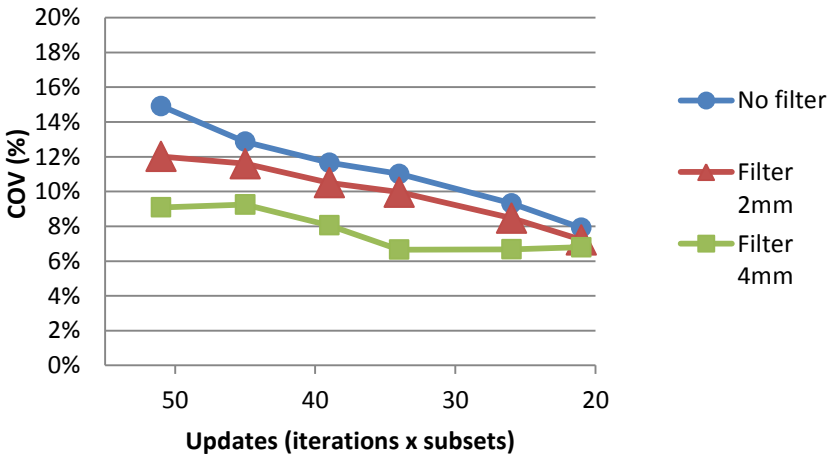


Figure 14 Averaged COVs for all spheres of the NEMA phantom, with averaged standard deviations (SD) for the highest three updates. Lower updates and adding filtering results in lower noise levels.

Table 7 RC_{mean} and RC_{max} phantom studies, per update, no filter applied.

		Updates					
		51	45	39	34	26	21
RC_{mean}							
	Sphere diameter						
NEMA phantom	37	0.87	0.85	0.85	0.85	0.84	0.84
	28	0.81	0.82	0.83	0.81	0.83	0.81
	22	0.78	0.78	0.76	0.77	0.76	0.77
	17	0.72	0.75	0.72	0.71	0.73	0.74
	13	0.62	0.62	0.59	0.63	0.59	0.59
	10	0.56	0.54	0.55	0.53	0.52	0.50
MICRO phantom	8	0.49	0.50	0.47	0.47	-	-
	6	0.30	0.30	0.30	0.29	-	-
	5	0.22	0.21	0.21	0.20	-	-
	4	0.14	0.14	0.14	0.13	-	-
RC_{max}							
	Sphere diameter						
NEMA phantom	37	1.10	1.12	1.11	1.10	1.10	1.09
	28	1.09	1.10	1.11	1.11	1.09	1.04
	22	1.07	1.07	1.04	1.05	1.05	1.04
	17	1.00	1.02	1.01	1.02	1.03	1.06
	13	0.91	0.90	0.88	0.94	0.88	0.88
	10	0.85	0.81	0.84	0.82	0.80	0.76
MICRO phantom	8	0.61	0.62	0.60	0.60	-	-
	6	0.40	0.40	0.40	0.39	-	-
	5	0.28	0.27	0.27	0.26	-	-
	4	0.18	0.18	0.18	0.17	-	-

Table 8 RC_{mean} and RC_{max} for each filter option, for 51 updates

		Filter diameter		
		No filter	2 mm	4 mm
RC_{mean}				
	Sphere diameter			
NEMA phantom	37	0.87	0.84	0.82
	28	0.81	0.79	0.76
	22	0.78	0.75	0.71
	17	0.72	0.69	0.64
	13	0.62	0.60	0.53
	10	0.56	0.50	0.41
MICRO phantom	8	0.49	0.44	-
	6	0.30	0.27	-
	5	0.22	0.19	-
	4	0.14	0.13	-
RC_{max}				
	Sphere diameter			
NEMA phantom	37	1.10	1.08	1.04
	28	1.09	1.08	1.00
	22	1.07	1.03	0.97
	17	1.00	0.97	0.92
	13	0.91	0.89	0.79
	10	0.85	0.78	0.62
MICRO phantom	8	0.61	0.55	-
	6	0.40	0.36	-
	5	0.28	0.25	-
	4	0.18	0.17	-

Patient study

Patient characteristics

Characteristics from 24 patients are shown in Table 9. A total of 80 lesions was included (size: 0.1-8.1 mL, mean: 0.8 mL).

Quantitative analysis

On patient data we applied 51, 45 and 39 updates without filtering. The average noise level decreased with 6% for 45 updates and 13% for 39 updates when compared to 51 updates ($p < 0.01$), as can be seen in Figure 11. For all lesions (size: 0.1-8.1 mL). SUV_{mean} and SUV_{max} were similar between those updates, see Table 10**.** SNR increases for lower updates up to 12.6% for 39 updates for SNR_{max} and 12.4% for SNR_{mean}.

Analysis of the sub selection of small lesions (<0.5 mL = 1 cm diameter) showed similar results as for the overall analysis, as can be found in Table 11. SUVs were slightly lower for each update, but relative changes between the three updates still were small. SNR increased for lower updates as well, up to 12% for 39 updates.

Table 9 Patient- and lesion characteristics

Patient characteristics (n = 24)		
Gender	Female	11
	Male	13
Age (years)		64 ± 11
Body weight (kg)		79.9 ± 14.4
Glucose (mmol/L)		5.5 ± 0.7
Proven malignancy	Yes	21
	No	2
	Unknown	1
Lesion characteristics (n = 80)		
Type	Primary tumor	15
	Lymph node	29
	Metastasis	34
	Unknown	2
Lesion size (mL) per reconstruction	3i 17s	0.8 ± 1.2 (0.1-7.0)
	3i 15s	0.8 ± 1.3 (0.1-8.1)
	3i 13s	0.8 ± 1.3 (0.1-7.4)

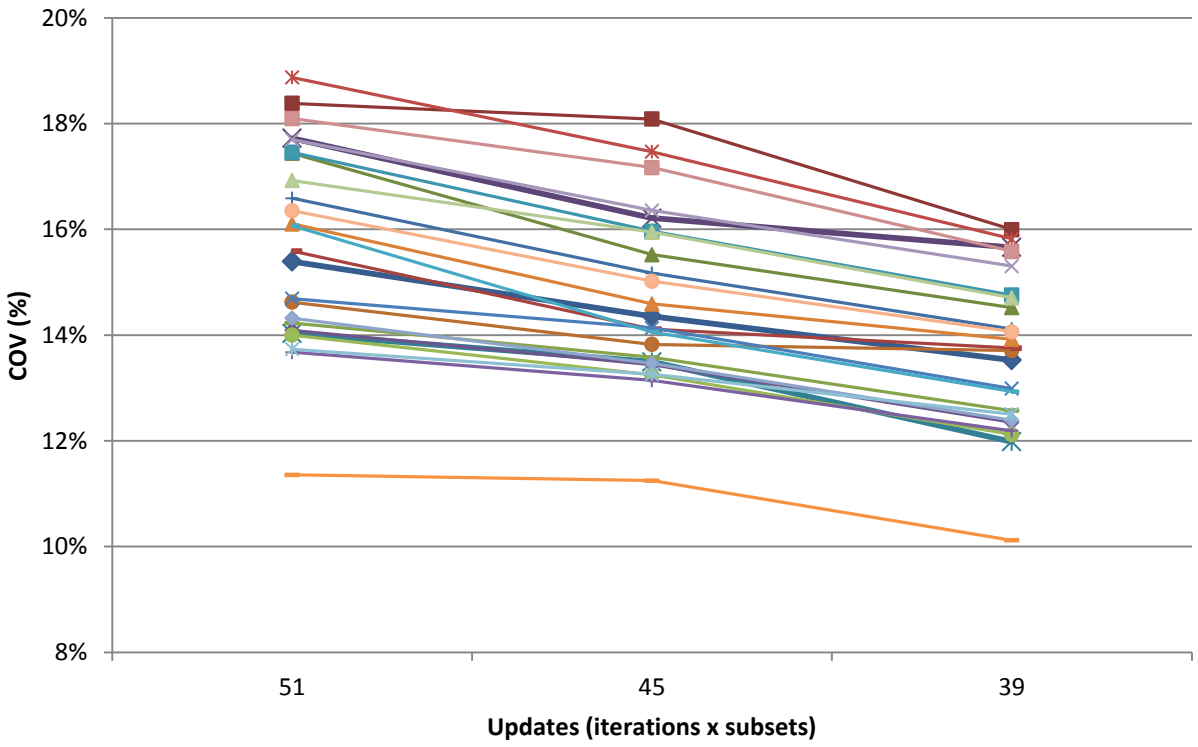


Figure 15 Noise level per reconstruction per patient, each line represents one patient (n = 24). Noise level(COV) decreases by lower number of updates.

Table 10 Quantitative results lesion measurements per update, averaged over all lesions

Average over all lesions (n = 80)	3i17s (51 updates)	3i15s (45 updates)	3i13s (39 updates)
SUV_{max}	6.2 ± 4.6	6.1 ± 4.7	6.1 ± 4.6
ΔSUV_{max}	-	-1.6%	-1.8%
SUV_{mean}	4.7 ± 3.3	4.7 ± 3.3	4.7 ± 3.4
ΔSUV_{mean}	-	-1.7%	-2.0%
SNR_{max}	16.2 ± 10.2	17.0 ± 11.1	18.3 ± 12.1
ΔSNR_{max}	-	+5.1%	+12.6%
SNR_{mean}	12.4 ± 7.5	13.0 ± 8.1	14.0 ± 8.9
ΔSNR_{mean}	-	+5.0%	+12.4%

Table 11 Quantitative results lesion measurements per update, averaged over all small lesions (<0.5 mL)

Average over all small lesions < 0.5 mL (n = 52)	3i17s (51 updates)	3i15s (45 updates)	3i13s (39 updates)
SUV_{max}	5.0 ± 2.4	4.6 ± 1.7	4.9 ± 2.7
ΔSUV_{max}	-	-1.2%	-1.7%
SUV_{mean}	3.9 ± 2.0	3.6 ± 1.3	3.8 ± 2.2
ΔSUV_{mean}	-	-1.2%	-1.8%
SNR_{max}	13.7 ± 6.2	13.6 ± 4.6	15.4 ± 7.1
ΔSNR_{max}	-	-0.7%	+12.0%
SNR_{mean}	10.8 ± 5.1	10.7 ± 3.7	12.1 ± 5.7
ΔSNR_{mean}	-	-0.1%	+12.0%

Visual analysis

Based on visual analysis of 24 patients, a small preference was found for the reconstruction with 39 updates for both observers (10/24), as shown in Figure 16. However, agreement between both observers was found in only 8/24 cases. There was no significant difference between both observers ($p = 0.96$). None of the reconstructions was preferred with statistical significance by observer 1 or 2 ($p = 0.69$ and $p = 0.61$, respectively). Image quality was acceptable for all scans that were noted as preferable by the observers. The first observer did see differences between all reconstructions, with the note that differences are only found in contrast and greyscale instead of resolution. However, the second observer noted that this was the case for only three patients. An example of the three different reconstructions in one patient can be found in Figure 17.

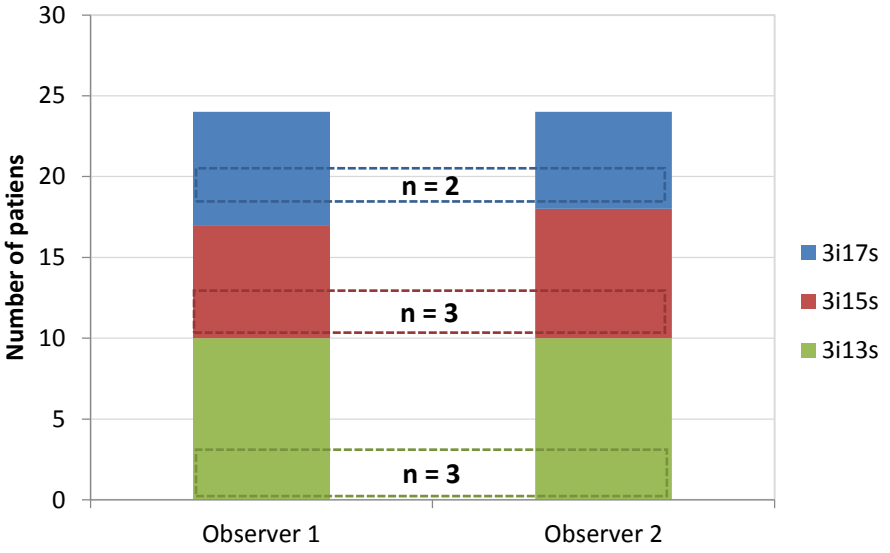


Figure 16 Preferences of both observers for each reconstruction for all patients (n=24). Dotted lines indicate agreement between both observers. The reconstruction with 3 iterations and 13 subsets (39 updates) was chosen slightly more often (10/24) than the other reconstructions.

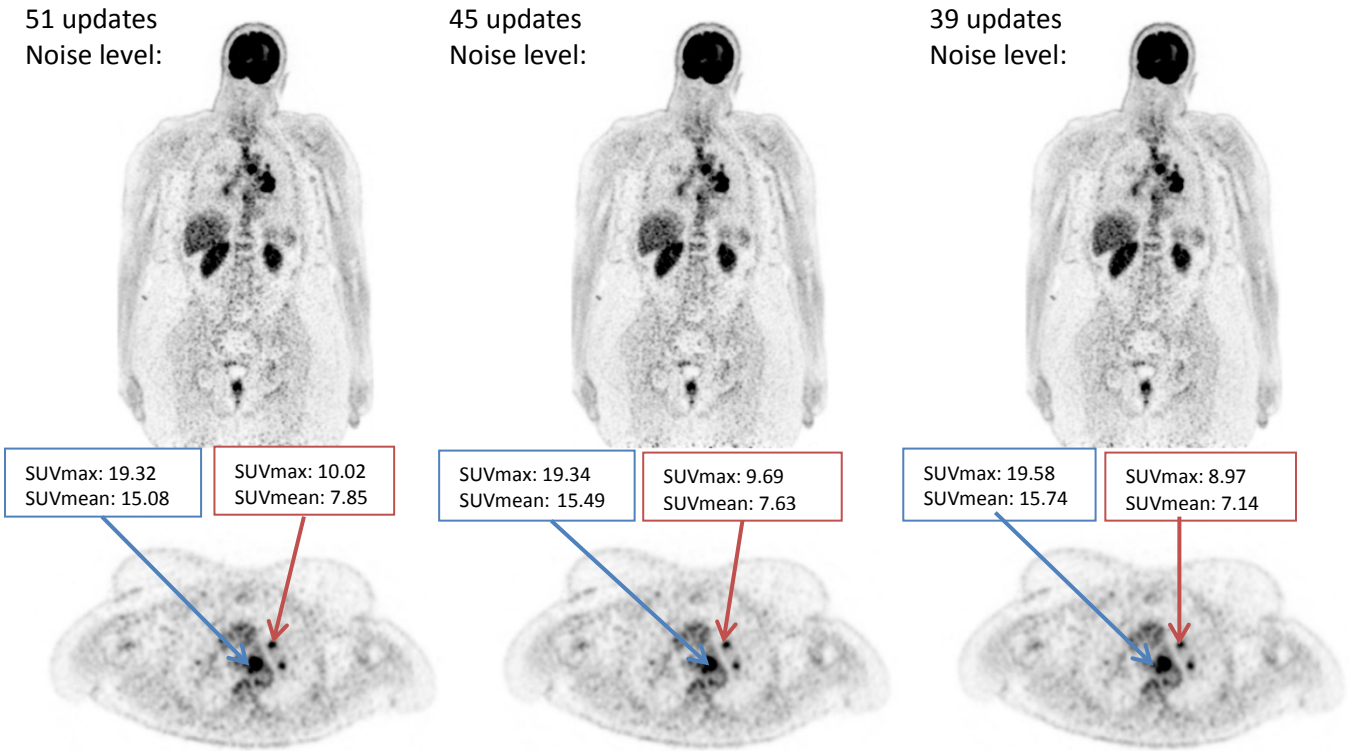


Figure 17 Three tested reconstruction settings on one patient. No differences in lesion detectability are seen. A slightly lower noise level for the 39 updates can be seen. This patient was scanned with her arms along her body.

Discussion

In this study, we have determined reconstruction settings for a small-voxel reconstruction for digital PET, resulting in the best image quality with high tumor-to-background ratio and an acceptable noise level. A reconstruction with 3 iterations and 13 subsets without post-Gaussian filtering showed a noise reduction of 13% and comparable lesion detectability in patient data, compared to 3 iterations and 17 subsets without filtering, which are the default settings recommended by the manufacturer. In phantom data, we found a significant decrease in noise level by lowering the number of updates ($p < 0.01$). This is consistent with results Morey et al. showed in their research about varying number of subsets [54]. Adding a 2 mm or a 4 mm post-smoothing Gaussian filter resulted in noise reduction as well ($p \leq 0.03$). With regard to RCs, only lowering updates up to 39 updates resulted in stable signal ($p > 0.12$), whereas going even lower or adding filters resulted in decrease of signal ($p < 0.01$), especially for small spheres. As we wanted to enhance small lesion detection, this partial volume effect is not preferable. Therefore, on patient data we only tested the three highest updates (51, 45 and 39) without added filtering. We have found that noise levels decreased by lower updates with 6% ($p < 0.01$) and 13% ($p < 0.01$) respectively for 45 and 39 updates, and SUVs remained similar, as can be found in Table 5. On the basis of visual assessment, no differences in preference were found ($p = 0.61$ and $p = 0.69$). However, based on quantitative measurements, noise level will reduce significantly when using 39 updates, without hampering signal levels.

A remarkable observation was the difference in noise reduction between the phantom study and the patient study. In the NEMA phantom, we obtained a noise reduction of 22%, as for the patients study it was 13%, both for 39 updates. The main reason for the difference in noise levels, could be the large variation between 24 patients, compared to just one measurement of both phantoms. Therefore, phantom data should not be used one to one to patient data.

A limitation of our study was the relative small size of our patient population. Consequently, the results could be more prone to variability within the population. However, the results of the quantitative study are still statistically significant, so it is to be expected that the results will not become different for a larger patient cohort. Moreover, the results of the visual analysis did not show significant results. The main cause for that could be that the differences in noise levels were, although quantitative different, not different enough for visual analysis. Another aspect was the fact that only two physicians performed the visual analysis.

In our study, we only took into account the OSEM reconstruction method, because this is the most used iterative reconstruction method nowadays, and in our hospital the only reconstruction we perform [14]. Other techniques enhancing better image quality are based on different steps within the iterative reconstruction methods or based on post reconstruction filtering, such as the PSF compensation [14]. In our research project, we did not test PSF correction because earlier research in our hospital showed the use of PSF is not preferable [56]. However, it is possible that this would result in less noise and increased detection of small lesions, because PSF increases spatial resolution, even by lower number of updates [14, 57]. On the other hand, PSF could lead to more false-positives, as all lesions are emphasized by the PSF [58]. A technique based on PSF that may improve image quality is a Bayesian penalized likelihood algorithm. This algorithm uses prior knowledge to improve image quality, for example by comparing voxels in each iteration to investigate the noise level and using a penalization factor to reduce the noise in following iterations and enhance edge enhancement [13, 59].

To further improve image quality, research could be done on the topic of longer uptake times between injection and scanning. Nowadays, our scanning protocol is based on a 60 minutes uptake time. However, malignant tissue continues to accumulate FDG for a longer time than benign tissue, thus by enlarging the uptake time, the lesion detectability for malignant lesions can be improved

[15]. Al-Faham et al. observed that lesion detectability in the liver will be improved by a uptake time of 80 minutes, when the signal for malignant lesions is still increasing and for benign lesions the signal will reach a plateau [60]. A disadvantage of longer uptake time could be a higher FDG uptake in muscles, due to motion. Besides, a longer uptake time will lead to longer delay in clinical workflow.

Overall, we can conclude that for a OSEM reconstruction with 2x2x2 mm^3 voxels on a digital PET system, a decrease in number of updates from 51 to 39 updates is preferably, as less subsets will result in decreased noise levels and therefore better lesion detectability. Another consequence of using 39 updates could be reduction in FDG dose or in scan duration, resulting in comparable noise level as for 51 updates. For other PET systems and other voxel sizes, it is preferable to optimize the image reconstruction protocol as well, as it is possible to reach an optimum between lesion detectability and noise level.

Conclusion

We have demonstrated that adequate image quality for a digital PET system using the OSEM reconstruction algorithm requires to select a number of updates. The lowest noise-level for 2x2x2 mm^3 voxels without impairing small lesion detectability can be obtained using 3 iterations x 13 subsets (39 updates) without post-smoothing filter. Post-smoothing filtering results in a reversed effect, with not only reduced noise levels but reduced small lesion detectability as well.

References

1. García-Figueiras R, Baleato-González S, Padhani AR, Luna-Alcalá A, Vallejo-Casas JA, Sala E, et al. How clinical imaging can assess cancer biology. Insights Imaging 2019;10:1–35.
2. Ben-Haim S, Ell P. 18F-FDG PET and PET/CT in the Evaluation of Cancer Treatment Response. J Nucl Med 2009;50:88–99.
3. Baratto L, Park SY, Hatami N, Davidzon G, Srinivas S, Gambhir SS, et al. 18F-FDG silicon photomultiplier PET/CT: A pilot study comparing semi-quantitative measurements with standard PET/CT. PLoS One 2017;12:1–13.
4. Nguyen NC, Vercher-Conejero JL, Sattar A, Miller MA, Maniawski PJ, Jordan DW, et al. Image Quality and Diagnostic Performance of a Digital PET Prototype in Patients with Oncologic Diseases: Initial Experience and Comparison with Analog PET. J Nucl Med 2015;56:1378–1385.
5. Boellaard R, Delgado-Bolton R, Oyen WJG, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging 2015;42:328–354.
6. Boellaard R, O'Doherty MJ, Weber WA, Mottaghy FM, Lonsdale MN, Stroobants SG, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: Version 1.0. Eur J Nucl Med Mol Imaging 2010;37:181–200.
7. Bertero L, Massa F, Metovic J, Zanetti R, Castellano I, Ricardi U, et al. Eighth Edition of the UICC Classification of Malignant Tumours: an overview of the changes in the pathological TNM classification criteria—What has changed and why? Virchows Arch 2018;472:519–531.
8. Wright CL, Binzel K, Zhang J, Knopp M V. Advanced functional tumor imaging and precision nuclear medicine enabled by digital PET technologies. Contrast Media Mol Imaging 2017;2017:1–7.
9. Koopman D, van Dalen JA, Arkies H, Oostdijk AHJ, Francken AB, Bart J, et al. Diagnostic implications of a small-voxel reconstruction for loco-regional lymph node characterization in breast cancer patients using FDG-PET/CT. EJNMMI Res 2018;8:1–10.
10. Slomka PJ, Pan T, Germano G. Recent Advances and Future Progress in PET Instrumentation. Semin Nucl Med 2016;46:5–19.
11. Conti M. Focus on time-of-flight PET: The benefits of improved time resolution. Eur J Nucl Med Mol Imaging 2011;38:1147–1157.
12. Kadrmas DJ, Casey ME, Conti M, Jakoby BW, Townsend DW. Impact of Time-of-Flight on PET Tumor Detection. J Nucl Cardiol 2009;50:1315–1323.
13. van der Vos CS, Koopman D, Rijnsdorp S, Arends AJ, Boellaard R, van Dalen JA, et al. Quantification, improvement, and harmonization of small lesion detection with state-of-the-art PET. Eur. J. Nucl. Med. Mol. Imaging
14. Tong S, Alessio AM, Kinahan PE. Image reconstruction for PET/CT scanners: Past achievements and future challenges. Imaging Med 2010;2:529–545.
15. Adams MC, Turkington TG, Wilson JM, Wong TZ. A systematic review of the factors affecting accuracy of SUV measurements. Am J Roentgenol 2010;195:310–320.
16. Koopman D, van Dalen JA, Lagerweij MCM, Arkies H, de Boer J, Oostdijk AHJ, et al. Improving the Detection of Small Lesions Using a State-of-the-Art Time-of-Flight PET/CT System and Small-Voxel Reconstructions. J Nucl Med Technol 2015;43:21–27.
17. Jager PL, Koopman D, Van Dalen JA, Van Osch JAC, Blaauwbroek A, Knollema S. Research protocol Intra-individual patient-based comparison of conventional and digital PET/CT scanners (PETPET study). 2017.
18. Poeppel TD, Krause BJ, Heusner TA, Boy C, Bockisch A, Antoch G. PET/CT for the staging and follow-up of patients with malignancies. Eur J Radiol 2009;70:382–392.
19. Lv YL, Yuan DM, Wang K, Miao XH, Qian Q, Wei SZ, et al. Diagnostic performance of integrated positron emission tomography/computed tomography for mediastinal lymph node staging in non-small cell lung cancer: A bivariate systematic review and meta-analysis. J Thorac Oncol 2011;6:1350–1358.
20. Tournoy KG, Maddens S, Gosselin R, Van Maele G, Van Meerbeeck JP, Kelles A. Integrated FDG-PET/CT does not make invasive staging of the intrathoracic lymph nodes in non-small cell lung cancer redundant: A prospective study. Thorax 2007;62:696–701.
21. Takamochi K, Yoshida J, Murakami K, Niho S, Ishii G, Nishimura M, et al. Pitfalls in lymph node staging with positron emission tomography in non-small cell lung cancer patients. Lung Cancer 2005;47:235–242.
22. Soret M, Bacharach SL, Buvat I. Partial-Volume Effect in PET Tumor Imaging. J Nucl Med 2007;48:932–945.
23. Kolthammer JA, Su K-H, Grover A, Narayanan M, Jordan DW, Muzic RF. Performance evaluation of the Ingenuity TF PET/CT scanner with a focus on high count-rate conditions. Phys Med Biol 2014;59:3843–3859.

24. López-Mora DA, Flotats A, Fuentes-Ocampo F, Camacho V, Fernández A, Ruiz A, et al. Comparison of image quality and lesion detection between digital and analog PET/CT. Eur J Nucl Med Mol Imaging 2019;46:1383–1390.
25. van Sluis J, de Jong J, Schaar J, Noordzij W, van Snick P, Dierckx R, et al. Performance characteristics of the digital Biograph Vision PET/CT system. J Nucl Med 2019;60:1031–1036.
26. Fuentes-Ocampo F, López-Mora DA, Flotats A, Paillahueque G, Camacho V, Duch J, et al. Digital vs. analog PET/CT: intra-subject comparison of the SUVmax in target lesions and reference regions. Eur J Nucl Med Mol Imaging 2019;46:1745–1750.
27. Hsu DFC, Ilan E, Peterson WT, Uribe J, Lubberink M, Levin CS. Studies of a Next-Generation Silicon-Photomultiplier–Based Time-of-Flight PET/CT System. J Nucl Med 2017;58:1511–1518.
28. Alfonso López-Mora D, Flotats A, Fuentes-Ocampo F, Camacho V, Fernández A, Ruiz A, et al. Comparison of image quality and lesion detection between digital and analog PET/CT. Eur J Nucl Med Mol Imaging 2019;46:1383–1390.
29. Koopman D, Jager PL, van Dalen JA. Small-voxel reconstructions significantly influence SUVs in PET imaging. Eur J Nucl Med Mol Imaging 2019;46:1751–1752.
30. Koopman D, van Osch JAC, Jager PL, Tenbergen CJA, Knollema S, Slump CH, et al. Technical note: how to determine the FDG activity for tumour PET imaging that satisfies European guidelines. EJNMMI Phys 2016;3:1–10.
31. Sridhar P, Mercier G, Tan J, Truong MT, Daly B, Subramaniam RM. FDG PET metabolic tumor volume segmentation and pathologic volume of primary human solid tumors. Am J Roentgenol 2014;202:1114–1119.
32. Froud R, Abel G. Using ROC curves to choose minimally important change thresholds when sensitivity and specificity are valued equally: The forgotten lesson of pythagoras. Theoretical considerations and an example application of change in health status. PLoS One 9:
33. Cheng G, Torigian DA, Zhuang H, Alavi A. When should we recommend use of dual time-point and delayed time-point imaging techniques in FDG PET? Eur J Nucl Med Mol Imaging 2013;40:779–787.
34. Chin BB, Green ED, Turkington TG, Hawk TC, Coleman RE. Increasing uptake time in FDG-PET: Standardized uptake values in normal tissues at 1 versus 3 h. Mol Imaging Biol 2009;11:118–122.
35. Koopman D, van Dalen JA, Stigt JA, Slump CH, Knollema S, Jager PL. Current generation time-of-flight 18F-FDG PET/CT provides higher SUVs for normal adrenal glands, while maintaining an accurate characterization of benign and malignant glands. Ann Nucl Med 2015;1–8.
36. Yang DD, Mirvis E, Goldring J, Patel ARC, Wagner T. Improving diagnostic performance of 18F-FDG-PET/CT for assessment of regional nodal involvement in non-small cell lung cancer. Clin Radiol 2019;74:818.e17-818.e23.
37. Goense L, Ruurda JP, Carter BW, Fang P, Ho L, Meijer GJ, et al. Prediction and diagnosis of interval metastasis after neoadjuvant chemoradiotherapy for oesophageal cancer using 18F-FDG PET/CT. Eur J Nucl Med Mol Imaging 2018;45:1742–1751.
38. Kaalep A, Sera T, Rijnsdorp S, Yaqub M, Talsma A, Lodge MA, et al. Feasibility of state of the art PET/CT systems performance harmonisation. Eur J Nucl Med Mol Imaging 2018;45:1344–1361.
39. Nederlandse Kankerregistratie (NKR) IKNL. Cijfers over kanker. Available at: https://www.cijfersoverkanker.nl/selecties/dataset_1/img5d4921be5ba95.
40. Ji X, Cai J, Chen Y, Chen L-Q. Lymphatic spreading and lymphadenectomy for esophageal carcinoma. World J Gastrointest Surg 2016;8:90.
41. Li CL, Zhang FL, Wang Y Di, Han C, Sun GG, Liu Q, et al. Characteristics of recurrence after radical esophagectomy with two-field lymph node dissection for thoracic esophageal cancer. Oncol Lett 2012;5:355–359.
42. Schlottmann F, Barbetta A, Mungo B, Lidor AO, Molena D. Identification of the Lymphatic Drainage Pattern of Esophageal Cancer with Near-Infrared Fluorescent Imaging. J Laparoendosc Adv Surg Tech 2017;27:268–271.
43. Quint LE, Hepburn LM, Francis IR, Whyte RI, Orringer MB. Incidence and distribution of distant metastases from newly diagnosed esophageal carcinoma. Cancer 1995;76:1120–1125.
44. Nederlandse Kankerregistratie (NKR) IKNL. Cijfers over kanker - Overleving. Available at: https://www.cijfersoverkanker.nl/selecties/Dataset_3/img5d8206e21d1cd.
45. Shaheen O, Ghibour A, Alsaid B. Esophageal Cancer Metastases to Unexpected Sites: A Systematic Review. Gastroenterol Res Pract
46. Van Westreenen HL, Westerterp M, Bossuyt PMM, Pruim J, Sloof GW, Van Lanschot JJB, et al. Systematic review of the staging performance of 18F- fluorodeoxyglucose positron emission tomography in

esophageal cancer. J Clin Oncol 2004;22:3805–3812.
47. Teoh AYB, Chiu PWY, Yeung WK, Liu SYW, Wong SKH, Ng EKW. Long-term survival outcomes after definitive chemoradiation versus surgery in patients with resectable squamous carcinoma of the esophagus: Results from a randomized controlled trial. Ann Oncol 2013;24:165–171.
48. Gwynne S, Hurt C, Evans M, Holden C, Vout L, Crosby T. Definitive Chemoradiation for Oesophageal Cancer - a Standard of Care in Patients with Non-metastatic Oesophageal Cancer. Clin Oncol 2011;23:182–188.
49. Mariette C, Balon JM, Piessen G, Fabre S, Van Seuningen I, Triboulet JP. Pattern of recurrence following complete resection of esophageal carcinoma and factors predictive of recurrent disease. Cancer 2003;97:1616–1623.
50. Ohta M, Tokuda Y, Suzuki Y, Kubota M, Makuuchi H, Tajima T, et al. Whole body pet for the evaluation of bony metastases in patients with breast cancer: Comparison with 99 tc m -mdp bone scintigraphy. Nucl Med Commun 2001;22:875–879.
51. Meyer M, Gast T, Raja S, Hubner K. Increased F-18 FDG accumulation in an Acute Fracture. Clin Nucl Med 1994;19:13–14.
52. Kaalep A, Sera T, Oyen W, Krause BJ, Chiti A, Liu Y, et al. EANM/EARL FDG-PET/CT accreditation - summary results from the first 200 accredited imaging systems. Eur J Nucl Med Mol Imaging 2018;45:412–422.
53. Rausch I, Ruiz A, Valverde-Pascual I, Cal-González J, Beyer T, Carrio I. Performance Evaluation of the Vereos PET/CT System According to the NEMA NU2-2012 Standard. J Nucl Med 2019;60:561–567.
54. Morey AM, Kadrmas DJ. Effect of Varying Number of OSEM Subsets on PET Lesion Detectability NIH Public Access. J Nucl Med Technol 2013;41:1–15.
55. Miller MA. Vereos White paper - Focusing on high performance. 2016.
56. Assink N, Van Dalen JA, Koopman D, Stevens H, Slump CH, Jager PL. Impact of point-spread function modelling as a post-reconstruction deconvolution on FDG-PET image quality. Annu. Congr. Eur. Assoc. Nucl. Med.
57. Ciappuccini R, Desmonts C, Licaj I, Blanc-Fournier C, Bardet S, Aide N. Optimization of a dedicated protocol using a small-voxel PSF reconstruction for head-and-neck 18FDG PET/CT imaging in differentiated thyroid cancer. EJNMMI Res 2018;8:1–8.
58. Lasnon C, Hicks RJ, Beauregard J-M, Milner A, Paciencia M, Guizard A-V, et al. Impact of Point Spread Function Reconstruction on Thoracic Lymph Node Staging With 18 F-FDG PET/CT in NonYSmall Cell Lung Cancer. Clin Nucl Med 2012;37:971–976.
59. Chen C-T, Johnson VE, Wong WH, Hu X, Metz CE. Bayesian image reconstruction in positron emission tomography. IEEE Trans Nucl Sci 1990;37:636–641.
60. Al-Faham Z, Jolepalem P, Rydberg J, Wong C-YO. Optimizing 18F-FDG Uptake Time Before Imaging Improves the Accuracy of PET/CT in Liver Lesions. J Nucl Med Technol 2016;44:70–72.