Diagnostic implications of a small-voxel reconstruction for loco-regional lymph node characterization in breast cancer patients using FDG-PET/CT

Daniëlle Koopman 1,2*, Jorn A. van Dalen 3, Hester Arkies 1, Ad H. J. Oostdijk 1, Anne Brecht Francken 4, Jos Bart 5, Cornelis H. Slump 2, Siert Knollema 1 and Pieter L. Jager 1

Abstract

Background: We evaluated the diagnostic implications of a small-voxel reconstruction for lymph node characterization in breast cancer patients, using state-of-the-art FDG-PET/CT. We included 69 FDG-PET/CT scans from breast cancer patients. PET data were reconstructed using standard 4 × 4 × 4 mm³ and small 2 × 2 × 2 mm³ voxels. Two hundred thirty loco-regional lymph nodes were included, of which 209 nodes were visualised on PET/CT. All nodes were visually scored as benign or malignant, and SUV max and TB ratio (=SUV max/SUV background) were measured. Final diagnosis was based on histological or imaging information. We determined the accuracy, sensitivity and specificity for both reconstruction methods and calculated optimal cut-off values to distinguish benign from malignant nodes.

Results: Sixty-one benign and 169 malignant lymph nodes were included. Visual evaluation accuracy was 73% (sensitivity 67%, specificity 89%) on standard-voxel images and 77% (sensitivity 78%, specificity 74%) on small-voxel images (p = 0.13). Across malignant nodes visualised on PET/CT, the small-voxel score was more often correct compared with the standard-voxel score (89 vs. 76%, p < 0.001). In benign nodes, the standard-voxel score was more often correct (89 vs. 74%, p = 0.04). Quantitative data were based on the 61 benign and 148 malignant lymph nodes visualised on PET/CT. SUVs and TB ratio were on average 3.0 and 1.6 times higher in malignant nodes compared to those in benign nodes (p < 0.001), on standard- and small-voxel PET images respectively. Small-voxel PET showed average increases in SUV max and TB ratio of typically 40% over standard-voxel PET. The optimal SUV max cut-off using standard-voxels was 1.8 (sensitivity 81%, specificity 95%, accuracy 85%) while for small-voxels, the optimal SUV max cut-off was 2.6 (sensitivity 78%, specificity 98%, accuracy 84%). Differences in accuracy were non-significant.

Conclusions: Small-voxel PET/CT improves the sensitivity of visual lymph node characterization and provides a higher detection rate of malignant lymph nodes. However, small-voxel PET/CT also introduced more false-positive results in benign nodes. Across all nodes, differences in accuracy were non-significant. Quantitatively, small-voxel images require higher cut-off values. Readers have to adapt their reference standards.

Keywords: State-of-the-art PET/CT, Small-voxel reconstruction, Breast cancer, Loco-regional lymph nodes

* Correspondence: d.koopman@isala.nl
1Department of Nuclear Medicine, Isala, Zwolle, the Netherlands
2MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, Enschede, the Netherlands
Full list of author information is available at the end of the article

© The Author(s). 2018 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.
**Background**

In recent years, there has been an increasing role for fluorine-18 fluordeoxyglucose positron emission tomography (FDG-PET) combined with computed tomography (CT) in the diagnostic evaluation of patients with stage II–IV primary breast cancer [1, 2]. FDG-PET/CT has now largely replaced conventional staging that included a bone scan, abdominal echography and chest x-ray. This is due to PET/CT's higher accuracy and the ability to perform whole-body staging in a single session [1, 3, 4].

In patients with stage II–IV primary breast cancer, accurate detection of loco-regional lymph nodes and distant metastasis is crucial for treatment selection and prognosis prediction. However, the sensitivity and detection rate of small lesions and lesions with low metabolism using FDG-PET are restricted, due to the limited PET spatial resolution [5, 6]. For axillary lymph node staging using FDG-PET/CT, a systematic review based on seven studies found an average sensitivity of 56% with 93% specificity [7]. More specifically, for micro-metastatic lesions (diameter ≤ 2 mm), a sensitivity of 11% was reported, while for macro-metastatic lesions (diameter > 2 mm), the sensitivity was 57% [7].

A PET reconstruction setting that possibly improves small lesion detection and sensitivity is the voxel size. In current practice, the image voxel size for whole-body FDG-PET scans is typically around 4 × 4 × 4 mm³ [8]. However, it has been suggested that in combination with new highly sensitive time of flight (TOF) PET/CT cameras, the use of reconstructions with smaller voxels might further improve the detection of small lesions [8, 9]. In a previous study, we have assessed to what extent small lesion detectability is influenced by the voxel size [10]. With the use of small 2 × 2 × 2 mm³ voxels, we found a profound increase in the standardized uptake value (SUV) and an improvement in signal-to-noise ratio (SNR) for small lesions, as compared to the values on PET images reconstructed with default 4 × 4 × 4 mm³ voxels.

In clinical trials that assessed the value of FDG-PET/CT in primary breast cancer patients scheduled for neo-adjuvant chemotherapy, the use of small-voxels for loco-regional lymph node characterization has been reported already [11, 12]. However, those studies did not compare standard- and small-voxel reconstructions in a clinical setting. Therefore, the diagnostic implications and potential clinical benefit of small 2 × 2 × 2 mm³ voxels, in terms of small lesion detection and sensitivity, remain unknown. The purpose of this study is to evaluate the diagnostic implications of a small-voxel reconstruction for loco-regional lymph node characterization in breast cancer patients, using state-of-the-art FDG-PET/CT.

**Methods**

**Inclusion**

In this study, we have included 69 consecutive women with primary stage II–IV ductal primary breast cancer who had been referred for a pre-treatment whole-body FDG-PET/CT scan. We only included patients with at least one loco-regional lymph node which could be classified as benign or malignant based on histological or imaging information. We received a waiver from the Medical Ethical Committee of our institution (METC Isala, Zwolle) to perform this retrospective study, as it deals with an evaluation of clinically indicated scans. Informed consent was obtained from all individual participants included in the study.

**PET/CT data acquisition**

Patients fasted for at least 6 h prior to scanning. Before intravenous injection of FDG, blood glucose levels were measured to ensure a value below 10 mmol/L. The mean glucose level was 5.4 mmol/L (range 3.8–9.3 mmol/L). A dedicated dose protocol depending quadratically on patients’ body weight was used. This protocol is described by the formula \( A = 3.8 \times w^2/t \), where \( A \) is the FDG activity to administer (in megabecquerel), \( w \) is the patients’ body weight (in kilogram), and \( t \) is the acquisition time per bed position (in seconds). This approach has been shown to result in an image quality that does not depend on patient's weight [13]. Acquisition times for the patient studies were 1 and 2 min per bed position for patients with body weight ≤ 80 and > 80 kg, respectively. The average administered FDG activity was 331 MBq (range 155–533 MBq).

All scans were acquired with patients in supine position, using a state-of-the-art PET/CT scanner (Ingenuity TF, Philips Healthcare, Cleveland, OH, USA). This fully three-dimensional TOF PET scanner is combined with a 128-slice CT scanner. The PET scan was acquired 60 min post-injection, using a whole-body protocol. Before PET imaging, a CT scan was acquired for attenuation correction. The CT scan parameters were tube voltage 120 kV, dose modulation with an average tube current of 53 mA (range 37–94 mA), slice collimation 64 × 0.625 mm, pitch 0.83 and rotation time 0.5 s.

**PET/CT data reconstruction**

PET data were reconstructed using a list-mode TOF algorithm and line-of-response row-action maximum-likelihood algorithm method [14, 15], called BLOB-OS-TF. Images were reconstructed in two types of matrices: 144 × 144 matrices with voxel size 4 × 4 × 4 mm³ (standard-voxels) and 288 × 288 matrices with voxel size 2 × 2 × 2 mm³ (small-voxels). For the standard-voxel reconstruction, the blob had a 2.5-mm radius with a blob shape parameter of 8.4 mm. The blob radius and
shape parameter for the small-voxel reconstruction were 2.8 and 6.4 mm, respectively. Furthermore, the relaxation parameters for the standard- and small-voxel reconstructions were 1.0 and 0.5, respectively. For both types of voxel reconstructions, 3 iterations and 43 subsets were applied. All reconstruction parameters were default settings recommended by the vendor. Point-spread function modelling was not applied.

CT data were reconstructed using an iterative reconstruction algorithm (iDose, Philips Healthcare, Cleveland, OH, USA) with iDose level 4 and a slice thickness of 3 mm. The administered FDG activity and PET/CT acquisition protocols were consistent with European Association of Nuclear Medicine (EANM) guidelines for tumour PET imaging [16, 17]. Moreover, the reconstructed PET images with standard-voxels fulfilled the EANM research Ltd. (EARL) accreditation specifications [18]. The small-voxel reconstruction does not fulfil the EARL accreditation specifications, because the recovery curves for the small 10- to 13-mm spheres increase up to values above the maximum EARL specifications [10].

**Visual evaluation**

Integrated PET/CT data were reviewed on a dedicated workstation (IntelliSpace Portal 6, Philips Healthcare, Cleveland, OH, USA). First, each PET/CT scan was evaluated by two nuclear medicine (NM) physicians, with more than 5 years of experience in PET/CT viewing. They were blinded to the patient record and histological information and interpreted the PET/CT data by simultaneous viewing of PET, CT and fused PET/CT images. Both the standard- and small-voxel images were evaluated blindly, both separately and randomly. The NM physicians scored all loco-regional lymph nodes showing focal FDG-uptake on the PET/CT images. Both the standard- and small-voxel reconstructions were 1.0 and 0.5, respectively. For both practitioners, the lesion size parameter for the small-voxel reconstruction was 2.8 and 6.4 mm, respectively. Furthermore, the relaxation parameters for the standard- and small-voxel reconstructions were 1.0 and 0.5, respectively. For both types of voxel reconstructions, 3 iterations and 43 subsets were applied. All reconstruction parameters were default settings recommended by the vendor. Point-spread function modelling was not applied.

Next, we calculated the lymph node-to-background ratio (TB\text{ratio}), defined as the ratio between the lymph node SUV\text{max} and the average SUV in the background (SUV\text{background}). To measure the SUV\text{background}, we defined two regions of interest (ROI1 and ROI2) on the axial PET image. ROI1 encircled both the lymph node under study and a surrounding background area of 800 mm², while ROI2 only encircled the lymph node. For both ROIs, the area size and the average SUV (SUV\text{mean}) were collected to calculate the SUV\text{background} in a donut-shaped ROI using Formula 1:

\[
\text{SUV}_{\text{background}} = \frac{(\text{ROI1 SUV}_{\text{mean}} \cdot \text{ROI1 area}) - (\text{ROI2 SUV}_{\text{mean}} \cdot \text{ROI2 area})}{\text{ROI1 area} - \text{ROI2 area}}
\]  

Finally, for all the scored lymph nodes, we measured the short-axis diameter on the axial slice of the attenuation CT scan.

**Final diagnosis**

The final diagnosis for each lymph node was based on histological information, follow-up (FU) imaging (FDG-PET/CT, contrast-enhanced CT or magnetic resonance imaging (MRI)) or additional imaging (contrast-enhanced CT or MRI) in the following way (Fig. 1). For patients who initially underwent a surgical resection that included sentinel lymph node biopsy or axillary lymph node dissection, the final diagnosis was based on the histological information obtained during surgery. Pathology examination was part of the clinical evaluation and was centralised at our institution. Lymph nodes were histologically processed by formalin fixation followed by paraffin embedding, according to standardized procedures.

The lymph nodes were serially sectioned at 250 μm at three levels and stained with both hematoxylin and eosin, with an immune-histochemical cytokeratin staining (panCK). The immune-histochemical procedure was performed by a fully automated procedure, using pre-diluted antibodies on the Ventana Benchmark system (Roche Ventana, Tucson AZ, USA). The sizes of the metastases were measured on a conventional bright-field microscope (Leica, DM4000, Leica Microsystems Germany) using a micro-measuring scale on glass slide (definition = 0.1 mm). In all lymph nodes, the largest diameter of a metastasis was reported.

For patients who were treated with neo-adjuvant chemotherapy, the final diagnosis was based on the response to
therapy as visualised on FU imaging combined with the histological information that was available from the subsequent surgical resection. For these patients, lesions were considered malignant when they showed a decrease in size or FDG-uptake induced by subsequent chemotherapy. Furthermore, lesions that were stable in size and FDG-uptake during neo-adjuvant therapy were considered to be benign unless there was proof of malignancy from histological information obtained during surgery. When histological information or FU imaging was not available, the final diagnosis was based on the results of additional contrast-enhanced CT or MRI.

Additionally, we collected information on all malignant loco-regional lymph nodes that were found during surgical resection but which had not been visualised on FDG-PET or the attenuation CT. For those lesions, we recorded the metastatic deposit size that was measured during a separate pathology examination, performed by one pathologist (JB). For visual PET performance evaluation, these lymph nodes were regarded as benign nodes on PET. Furthermore, for quantitative PET evaluation, these lymph nodes were not taken into account because it was not possible to perform measurements on PET images.

**Lymph node characterization**

Visual evaluation scores were analysed on a lesion-per-lesion basis, by comparing the scores on standard- and small-voxel images for each lymph node. Quantitatively, we calculated average values for SUV\textsubscript{max} and TB\textsubscript{ratio} in both benign and malignant lymph nodes and for both voxel reconstructions. We created receiver operator curves (ROC) and calculated the area under the curve (AUC) with a 95% confidence interval (CI) for SUV\textsubscript{max} and TB\textsubscript{ratio}. For both reconstruction methods, we determined the sensitivity, specificity and accuracy for lymph node characterization from the visual and quantitative PET/CT evaluation, using the final diagnosis as a reference standard. We calculated optimal cut-off values for SUV\textsubscript{max} and TB\textsubscript{ratio} to distinguish benign from malignant lymph nodes on both voxel reconstructions. These cut-off values were based on the highest combined sensitivity and specificity (highest sum).

**Statistical analysis**

We used the McNemar test for paired samples to compare the visual scores for both reconstructions with the final diagnosis. Quantitative results were presented as mean ± standard deviation (SD). We included ranges in uptake values and lymph node size. Differences in SUV\textsubscript{max} and TB\textsubscript{ratio} between benign and malignant lymph nodes were evaluated using the Mann-Whitney U test. Furthermore, to evaluate differences in characterization performances between standard and small-voxels for SUV\textsubscript{max} and TB\textsubscript{ratio}, we compared the AUCs using a chi-square test. Additionally, the characterization performances for SUV\textsubscript{max} and TB\textsubscript{ratio} using optimal cut-off values were evaluated with the McNemar test for paired samples. A p value less than 0.05 was considered to indicate statistical significance.

**Results**

**Patient characteristics**

Clinical data from 69 patients are shown in Table 1. In total, 230 loco-regional lymph nodes were investigated (mean 3 ± 2 lymph nodes per patient). We have included 61 benign lymph nodes with an average size of 6 mm (range 3–12 mm), and 169 malignant lymph nodes with
Table 1 General characteristics

| Patient characteristics (n = 69) | Age (years) | 53 ± 12 (mean ± SD) |
|---------------------------------|-------------|---------------------|
| Body weight (kilogram)          | 76 ± 14     |                     |
| Hormonal receptor status        |             |                     |
| Oestrogen                      | 51 pos., 16 neg., 2 unknown |
| Progesterone                    | 37 pos., 30 neg., 2 unknown |
| Human epidermal growth factor receptor 2 | 19 pos., 48 neg., 2 unknown |
| Loco-regional lymph nodes (n = 230) | Location | Left region 108 |
|                                 | Right region 122 |
| Final diagnosis                 | Benign 61    |
|                                 | Malignant 169 |
| Final diagnosis based on the following: | Histological proof 128 (56%) |
|                                 | FU imaging 99 (43%) |
|                                 | Additional imaging 3 (1%) |

SD standard deviation, pos. positive, neg. negative, FU follow-up

For 209 loco-regional lymph nodes, which were visualised on PET/CT, were visually and quantitatively evaluated in this study.

Visual evaluation
For 173 out of 209 lymph nodes (83%), visual interpretation scores were exactly similar for both small and standard-voxels. Furthermore, 32 lymph nodes (15%) were scored malignant on the small-voxel images but benign on the standard-voxel images. Contrarily, four lymph nodes (2%) were scored benign on the small-voxel images but malignant on the standard-voxel images.

In Table 2, the final diagnosis of each lymph node is compared with the visual scores on both voxel reconstructions. Sensitivity, specificity and accuracy for standard-voxel PET/CT images in visual evaluation were 67, 89 and 73%, respectively. For small-voxel PET/CT images, we found a sensitivity, specificity and accuracy of 78, 74 and 77%, respectively. Across all lymph nodes, the differences in accuracy were not statistically significant (p = 0.13).

Limiting this analysis to the 148 malignant lymph nodes that were visualised on PET/CT, the small-voxel score was correct in 132 cases (89%) while the standard-voxel score was correct in 113 cases (76%), p < 0.001. In benign lymph nodes (n = 61) only, the small-voxel score was correct in 45 cases (74%) vs. 54 correct scores (89%) on standard-voxel images (p = 0.04).

Quantitative evaluation
SUV\text{max} and TB\text{ratio} across all lymph nodes are shown in Table 3. For both SUV\text{max} and TB\text{ratio}, and in both types of voxel reconstructions, uptake values for malignant lymph nodes were averagely a factor of 3.0 and 1.6 higher, respectively, as compared to those for benign nodes (p < 0.001). Furthermore, the use of small-voxels resulted in SUV increases of typically 40% (Table 3).

Lymph node characterization by quantitative evaluation
Table 4 shows the sensitivity, specificity and accuracy of lymph node characterization for standard- and small-voxel reconstructions.
small-voxel PET images, using optimal cut-off values for SUV\textsubscript{max} and TB\textsubscript{ratio}. Furthermore, receiver operator curves (ROC) for SUV\textsubscript{max} and TB\textsubscript{ratio} are shown in Fig. 2. TB\textsubscript{ratio} had a significantly lower AUC as compared to the SUV\textsubscript{max} parameter, for both standard and small-voxels \((p = 0.003\) and \(p = 0.002\)). AUC values for standard- and small-voxels images were comparable for both SUV parameters, with \(p = 0.71\) for SUV\textsubscript{max} and \(p = 0.61\) for TB\textsubscript{ratio}. Additionally, no significant differences were found in characterization performances based on the accuracy between standard- and small-voxel images, with \(p = 0.11\) for SUV\textsubscript{max} and \(p = 0.29\) for TB\textsubscript{ratio}. Table 4 shows that the use of small-voxels required higher SUV cut-offs for accurate lymph node characterization.

Clinical examples
Figure 3 shows FDG-PET/CT images from a patient with breast cancer. The visual score of the axillary lymph node altered from benign on standard-voxel PET/CT to malignant on small-voxel PET/CT. Furthermore, SUV\textsubscript{max} of this lymph node increased by 57\% on small-voxel images. Follow-up imaging showed that this lymph node responded to chemotherapy, which indicated that this lesion was malignant. This confirmed the small-voxel score and the classification by the optimal SUV\textsubscript{max} cut-off shown in Table 4.

Figure 4 shows FDG-PET/CT images from a breast cancer patient, with a small axillary lymph node. The visual score was benign on standard-voxel images while it was scored malignant on small-voxel images. In this case, SUV\textsubscript{max} increased with 64\% on small-voxel images. However, during sentinel node biopsy, no malignancy was found. This indicated that this lymph node was benign and the standard-voxel score was correct.

Discussion
This study shows that small-voxel PET reconstructions improve the diagnostic sensitivity in the detection of lymph node metastases in breast cancer, at the expense of an impaired specificity. Based on visual evaluation of PET/CT images, the lymph node characterization accuracy did not change when using small-voxels instead of standard-voxels \((p = 0.13)\). Nevertheless, limiting the analysis to malignant lymph nodes only, the small-voxel images were correct in 89\% while the standard-voxel images provided concordant scores in only 76\% of the cases \((p < 0.001)\). This shows that the visual evaluation and detection of malignant axillary lymph nodes improves using small-voxel PET/CT. Contrarily, benign lymph nodes were more often correctly scored on standard-voxel images as compared to those on the small-voxel images \((74 \text{ vs. } 89\%, p = 0.04)\).

Table 4 Sensitivity, specificity and accuracy for SUV\textsubscript{max} and TB\textsubscript{ratio} at optimal cut-offs, determined for standard- and small-voxel PET

|          | Optimal cut-off | Sensitivity | Specificity | Accuracy |
|----------|----------------|-------------|-------------|----------|
| SUV\textsubscript{max} | Standard-voxels | 1.8 | 81\% | 95\% | 85\% |
|          | Small-voxels  | 2.6 | 78\% | 98\% | 84\% |
| TB\textsubscript{ratio} | Standard-voxels | 2.4 | 80\% | 82\% | 80\% |
|          | Small-voxels  | 3.3 | 84\% | 77\% | 82\% |

Table 4: The use of small-voxel images requires higher SUV cut-offs for accurate lymph node characterization. Furthermore, SUV\textsubscript{max} showed a higher performance as compared to TB\textsubscript{ratio} with \(p = 0.04\) for standard-voxels and \(p < 0.001\) for small-voxels. However, the characterization performances were similar for standard- and small-voxel images, with \(p = 0.11\) for SUV\textsubscript{max} and \(p = 0.29\) for TB\textsubscript{ratio}.

SUV\textsubscript{max} maximum standardized uptake value, TB\textsubscript{ratio} ratio between the lymph node SUV\textsubscript{max} and the lymph node background uptake, PET positron emission tomography.
Fig. 3 FDG-PET/CT images of a patient with proven breast cancer. 

- **a** Axial PET image, standard-voxels. 
- **b** Axial-fused PET/CT image, standard-voxels. 
- **c** Axial PET image, small-voxels. 
- **d** Axial-fused PET/CT image, small-voxels. 

SUV\textsubscript{max} for this small lymph node (red arrows) increased from 2.1 on standard-voxel PET with visual score benign to SUV\textsubscript{max} 3.3 on small-voxel PET and visual score malignant. On follow-up imaging after chemotherapy, this lymph node showed regression, which indicated that the lymph node was malignant. This confirmed the small-voxel score and the classification using the optimal SUV\textsubscript{max} cut-off.

Fig. 4 FDG-PET/CT images of a patient with proven breast cancer. 

- **a** Axial PET image, standard-voxels. 
- **b** Axial-fused PET/CT image, standard-voxels. 
- **c** Axial PET image, small-voxels. 
- **d** Axial-fused PET/CT image, small-voxels. 

For this lymph node, with short-axis diameter 5 mm in the right axillary region (blue arrows), SUV\textsubscript{max} values were 1.4 and 2.3 (increase 64%) on standard- and small-voxel images, respectively. Furthermore, the visual scores were benign on standard-voxel PET and malignant on small-voxel PET. The sentinel node biopsy procedure did not reveal any malignancy. This indicates that this lymph node was benign, confirming the standard-voxel score and confirming the classification using the optimal SUV\textsubscript{max} cut-off.
Quantitatively, we found significant differences between benign and malignant lymph nodes for both parameters $S_{U V}^{\max}$ and $T_{B}^{\text{ratio}}$, and both PET voxel sizes ($2 \times 2 \times 2 \text{ mm}^3$ and $4 \times 4 \times 4 \text{ mm}^3$). Furthermore, the use of small-voxels resulted in higher FDG-uptake values of typically 40% for both benign and malignant lymph nodes. This increase is comparable with previously published values [10].

This study shows that the visual evaluation of malignant lymph nodes results in a higher accuracy and sensitivity when using small-voxel PET/CT. However, when looking quantitatively, both reconstruction voxel types can be used interchangeably for accurate loco-regional lymph node characterization. Furthermore, across all lymph nodes, both the quantitative and visual evaluation showed comparable performances in disease characterization. Moreover, when using small-voxel reconstructions, PET readers have to be aware of the risk of false-positives and the need of higher cut-off values to distinguish benign from malignant nodes. In the past, all visually FDG-positive axillary lymph nodes were considered as metastatic, because that approach resulted in high specificity [11, 12]. However, we demonstrated that on small-voxel PET images, benign loco-regional lymph nodes can also show some increased FDG-uptake. In the literature, this phenomenon has been described as well for the visualisation of benign nodes on PET images that incorporated point-spread function (PSF) modelling [19, 20]. Clinically, the main feature of small-voxel PET is the ability to detect more malignant axillary lymph nodes. This could lead to a higher disease stage and thereby potentially changes treatment. For example when the N-stage based on PET changes from N1 to N2, this could give rise to treat the patient with neo-adjuvant chemotherapy before surgery. Generally, with an improved detection of malignant axillary lymph nodes, the role of PET as an additional tool in the treatment decision plan can be further extended.

Between the two parameters, $S_{U V}^{\max}$ and $T_{B}^{\text{ratio}}$, which were evaluated in this study, we found a significant difference in clinical performance. For both voxel reconstruction types, the $S_{U V}^{\max}$ parameter was significantly better in distinguishing benign from malignant lymph nodes. This indicates that during quantitative PET evaluation, it is not required to take the surrounding background uptake around the lymph node into account for accurate characterization of loco-regional lymph nodes.

Despite the use of a small-voxel PET reconstruction, the detection rate with FDG-PET/CT for small metastases remains restricted. In our study, 21 out of 230 lymph nodes were not visualised on PET/CT images but were only found during the histological procedures after surgery. It is very likely that the size and/or the metabolic volume of those undetected lymph node metastasis was not large enough to be visualised with PET [2]. Currently, several developments in PET technology are taking place to further improve the diagnostic performance of PET/CT systems [21]. Together with the development of more specific radionuclide tracers for breast cancer, this may further improve the detection rate of PET/CT [22].

A limitation of this study is that the analysis was based on PET images together with low-dose CT only. Possibly, we did not incorporate all additional findings in case a diagnostic contrast-enhanced CT was obtained. It is likely that adding more CT aspects of lymph nodes, such as spherical shape and the absence of intra-nodal fat, will further improve the sensitivity, specifically of small-voxel reconstructions, and further reduce the false-positive rate. Moreover, there were some differences in visual scores between the two observers as consensus scoring was needed in 20% of the cases. This can be partly explained by the use of a five-point scoring system, which easily introduces small differences in visual scores between observers. Furthermore, the lack of a gold standard for the relatively new small-voxel PET images may have led to differences between observers. There is a learning curve involved in this new detailed method of reading the axilla in breast cancer patients. Also, possibly, results will improve by rigorous comparison with non-involved small nodes in the contralateral axilla as a normal reference.

The present study has some other limitations. The retrospective study design may have led to some bias in our study population, since only patients with a known lymph node status were included in this study. Additionally, since not all patients had a lymph node dissection, some microscopic nodes may have been missed in some patients. Moreover, the final diagnosis per lymph node in this study was based on three different references: histological data, FU imaging and additional imaging. Although the value of additional imaging as a reference method can be difficult, the number of final diagnoses that was based on this method was very limited (1%). Moreover, FDG-PET may give a false-positive result, e.g. in case of an inflammation after biopsy which would also disappear after, but not due to, neo-adjuvant chemotherapy.

Some recommendations can be made for further research. The small-voxel reconstruction protocol could be further optimized to improve image quality and lesion detectability. Recently, Bellevre et al. [19] demonstrated that the use of PSF modelling improved the performance of axillary staging in breast cancer patients. A combination of small-voxels and PSF modelling could be explored. Furthermore, the number of iterations and subsets could be optimized [20, 23, 24]. Meanwhile, other evaluation methods could be considered. For example, the value of combined standard- and small-voxel PET in visual evaluation can be explored, to further improve the accuracy. Also, the use of
SUV\textsubscript{peak} could potentially decrease differences between both reconstructions as it is less influenced by image noise as compared to SUV\textsubscript{max} [25]. Furthermore, the added value of the optimal SUV cut-offs as proposed in this study, onto the visual evaluation of small-voxel PET, can be studied. Also, optimal SUV cut-offs can be different in other types of breast cancers and may be influenced when incorporating data on the tumour proliferation index (ki67) or the hormone receptor status.

Apart from this, it can be interesting to study the clinical impact of small-voxel images in other body regions and for other diseases. For example, we previously evaluated the impact of a state-of-the-art PET/CT scanner on the visualisation and quantification of adrenal glands, using a standard-voxel reconstruction [26]. This could be extended towards an evaluation of the impact of a small-voxel reconstruction on adrenal gland quantification and classification with FDG-PET/CT.

Conclusions
We evaluated the diagnostic properties of a small-voxel reconstruction for loco-regional lymph node staging in 69 patients with breast cancer, using state-of-the-art TOF FDG-PET/CT. The use of small-voxels improves the sensitivity of visual PET/CT evaluation in malignant lymph nodes, as compared to standard-voxel analysis. However, it also introduced more false-positive results for benign nodes. Across all nodes, differences in accuracy were non-significant. Quantitatively, a small-voxel reconstruction implicates higher SUV cut-off values when differentiating benign from malignant axillary lymph nodes. In case small-voxel images are used, readers have to adapt their reference standards visually and quantitatively.

Abbreviations
AUC: Area under the curve; CI: Confidence interval; CT: Computed tomography; EANM: European Association of Nuclear Medicine; EARL: EANM research Ltd.; FDG: Fluor-18 fluorodeoxyglucose; FU: Follow-up; LYSO: Lutetium yttrium orthosilicate; MRT: Magnetic resonance imaging; NM: Nuclear medicine; PET: Positron emission tomography; ROC: Receiver operator curve; ROI: Region of interest; SD: Standard deviation; Small-voxels: 2 × 2 × 2 mm\textsuperscript{3} voxels; SNR: Signal-to-noise ratio; Standard-voxels: 4 × 4 × 4 mm\textsuperscript{3} voxels; SUV: Standardized uptake value; SUV\textsubscript{background}: SUV in the background; SUV\textsubscript{max}: Maximum SUV; TB ratio: Ratio between SUV\textsubscript{max} and SUV\textsubscript{background}; TOF: Time of flight

Acknowledgements
We are grateful to Amarisina Blauwsbroek from Isala for her assistance in the data collection, and we gratefully acknowledge the staff from Isala, Zwolle and MIRA, Enschede for their overall support and kind collaboration.

Funding
There were no sources of funding for this research.

Authors’ contributions
DK, JvD and PJ designed the study. HA and AO performed the visual PET evaluation. JB performed an additional pathology analysis on the samples obtained from surgery. DK performed the quantitative PET evaluation and analysed the data. DK wrote the concept manuscript, which was revised by JvD, HA, AO, AF, JB, CS, SK and PJ. All authors approved the final manuscript.

Ethics approval and consent to participate
All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. For this type of study, formal consent is not required. We received a waiver from the Medical Ethical Committee (METC, Isala, Zwolle) of our institution to perform this retrospective study, as it deals with an evaluation of clinically indicated scans. This research involved human participants. Informed consent was obtained from all individual participants included in the study.

Consent for publication
Not applicable.

Competing interests
The Department of Nuclear Medicine, Isala, has established a research cooperation with Philips Healthcare regarding new PET technologies. The content of the article was solely the responsibility of the authors. No other potential conflicts of interest relevant to this article were reported.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1Department of Nuclear Medicine, Isala, Zwolle, the Netherlands. 2MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, Enschede, the Netherlands. 3Department of Medical Physics, Isala, Zwolle, the Netherlands. 4Department of Surgery, Isala, Zwolle, the Netherlands. 5Department of Pathology, Isala, Zwolle, the Netherlands.

Received: 8 November 2017 Accepted: 9 January 2018
Published online: 16 January 2018

References
1. Rieger C, Herrmann J, Nagarajah J, Hecktor J, Kueimmel S, Otterbach F, et al. Whole-body FDG PET/CT is more accurate than conventional imaging for staging primary breast cancer patients. Eur J Nucl Med Mol Imaging. 2012;39(5):852–63.
2. Groheux D, Espé M, Giacchetti S, Hindle E. Performance of FDG PET/CT in the clinical management of breast cancer. Radiology. 2013;266(2):388–405.
3. Groheux D, Giacchetti S, Delord M, Hindle E, Vercellino L, Cuvier C, et al. 18F-FDG PET/CT in staging patients with locally advanced or inflammatory breast cancer: comparison to conventional staging. J Nucl Med. 2013;54(1):5–11.
4. Koolen BB, M-JTV P, Aukema TS, Vogel WV, Oldenburg HS, van der Hage JA, et al. Improving the detection of small lesions using a state-of-the-art time-resolution. Eur J Nucl Med Mol. 2011;38(6):1147–58.
5. Cooper K, Haman S, Meng Y, Ward S, Fitzgerald P, Papaionannou D, et al. Pet emission tomography (PET) for assessment of axillary lymph node status in early breast cancer: a systematic review and meta-analysis. Eur J Surg Oncol. 2011;37(3):187–98.
6. Conti M. Focus on time-of-flight PET: the benefits of improved time resolution. Eur J Nucl Med Mol. 2011;38(6):1147–57.
7. Morey AM, Nio F, Kadmas DJ. Effect of using 2 mm voxels on observer performance for PET lesion detection. IEEE Trans Nucl Sci. 2016;63(3):1359–66.
8. Koopman D, van Dalen JA, Lagerweij MC, Arkies H, de Boer J, Oostdijk AH, 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(1):21–7.
9. Aukema TS, Strafer ME, M-JTV P, Russell NS, Gilhuis KG, Vogel WV, et al. Detection of extra-axillary lymph node involvement with FDG PET/CT in patients with stage II–III breast cancer. Eur J Cancer. 2010;46(18):3205–10.
10. Koolen BB, RAV O, Elkhuisen PH, Vogel WV, M-JTV P, Rodenhuis S, et al. Locoregional lymph node involvement in 18F-FDG PET/CT in breast cancer.
13. de Groot EH, Post N, Boellaard R, Wagenaar NR, Willemse AT, van Dalen JA. Optimized dose regimen for whole-body FDG-PET imaging. EJNMMI Res. 2013;3(1):63.

14. Lewitt RM. Multidimensional digital image representations using generalized Kaiser–Bessel window functions. J Opt Soc Am. 1990;7(10):1834–46.

15. Matej S, Lewitt RM. Practical considerations for 3-D image reconstruction using spherically symmetric volume elements. IEEE Trans Med Imaging. 1996;15(1):68–78.

16. 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(2):328–54.

17. 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(1):181–200.

18. EARL. FDG-PET/CT Accreditation. EARL; [cited 2016]; Available from: http://earl.eanm.org/cms/website.php?id=/en/projects/fgd_pet_ct_accreditation/accreditation specifications.htm.

19. Bellevre D, Fournier CB, Switters O, Dugué AE, Levy C, Allouache D, et al. Staging the axilla in breast cancer patients with 18F-FDG PET: how small are the metastases that we can detect with new generation clinical PET systems? Eur J Nucl Med Mol Imaging. 2014;41(6):1103–12.

20. Lason C, Hicks RJ, Beauregard J-M, Milner A, Pacienca M, Guizard A-V, et al. Impact of point spread function reconstruction on thoracic lymph node staging with 18F-FDG PET/CT in non-small cell lung cancer. Clin Nucl Med. 2012;37(10):971–6.

21. 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 Imag. 2017;44(1):14–16.

22. Linden HM, Dehdashti F. Novel methods and tracers for breast cancer imaging. Semin Nucl Med. 2013;43:324–9.

23. Schaefferkoetter J, Casey M, Townsend D, El Fakhri G. Clinical impact of time-of-flight and point response modeling in PET reconstructions: a lesion detection study. Phys Med Biol. 2013;58(5):1465.

24. Vogel MV, Wensing BM, van Dalen JA, Kabbe PF, van den Hoogen FJ, Oyen WJ. Optimised PET reconstruction of the head and neck area: improved diagnostic accuracy. Eur J Nucl Med Mol Imaging. 2005;32(11):1276–82.

25. Akaishi G, Ikari Y, Nishida H, Nishio T, Ohnishi A, Maebatake A, et al. Influence of statistical fluctuation on reproducibility and accuracy of SUVmax and SUVpeak: a phantom study. J Nucl Med Technol. 2015;43(3):222–6.

26. Koopman D, van Dalen JA, Sigrd JA, Stump 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. 2016;30(2):145–52.