

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a postprint version which may differ from the publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/137023>

Please be advised that this information was generated on 2021-09-26 and may be subject to change.

The influence of SPECT reconstruction algorithms on image quality and diagnostic accuracy in phantom measurements and [^{99m}Tc]-sestamibi parathyroid scintigraphy

Running Head: Influence of SPECT reconstruction on image quality

Authors:

- | | | |
|---|--|-------------------------------|
| 1 | Ralph A. van Hoorn, MSc ^{*,§,‡} | (R.vanHoorn@ebh.umcn.nl) |
| 1 | Dennis Vriens, MD ^{*,§} | (D.Vriens@nucmed.umcn.nl) |
| 3 | Jan W.A. Postema, MD ^{*,**} | (J.Postema@rdgg.nl) |
| 4 | Anne I.J. Arens, MD [*] | (A.Arens@nucmed.umcn.nl) |
| 5 | Andreas Pfestroff, MD [†] | (Pfestrof@med.uni-marburg.de) |
| 6 | Wim J.G. Oyen, MD, PhD [*] | (W.Oyen@nucmed.umcn.nl) |
| 7 | Martin Gotthardt, MD, PhD [*] | (M.Gotthardt@nucmed.umcn.nl) |

*Department of Nuclear Medicine, Radboud University Nijmegen Medical Center, Nijmegen, the Netherlands

†Department of Nuclear Medicine, University Hospital Giessen-Marburg GmbH, Campus Marburg, Germany

§Both authors (RAvH and DV) contributed equally to this study.

‡Author RvH is currently affiliated at the dept. of Epidemiology, Biostatistics and Health Technology Assessment, Radboud University Nijmegen Medical Center, Nijmegen, the Netherlands

**Author JWAP is currently affiliated at the dept. of Nuclear Medicine, Reinier de Graaf Groep, Delft, the Netherlands

Corresponding author: D. Vriens

Competing interest:

RvH: No competing interests.

DV: No competing interests.

JWAP: No competing interests.

AIJA: No competing interests.

AP: No competing interests.

WJGO: No competing interests.

MG: No competing interests.

Authors Contributions:

RvH: Performed phantom studies, designed clinical study, made reconstructions, analyzed data, wrote manuscript

DV: Performed phantom studies, designed clinical study, analyzed data, performed statistics, wrote manuscript, reviewed manuscript.

JWAP: one of the two blinded physicians, wrote manuscript, reviewed manuscript.

AIJA: supervised clinically study, wrote manuscript, reviewed manuscript.

AP: designed clinical study, performed clinical study (scans and clinical data), wrote manuscript, reviewed manuscript.

WJGO: designed clinical study, wrote manuscript, reviewed manuscript.

MG: one of the two blinded physicians, designed clinical study, analyzed data, wrote manuscript, reviewed manuscript, project leader/supervisor.

Abstract & Keywords

Objective: To determine the influence and clinical consequences of different tomographic reconstruction algorithms on image quality and diagnostic accuracy of low count statistics single photon emission computed tomography (SPECT).

Methods: Phantom scans were used to assess image quality (% recovered contrast, relative noise level, spatial resolution). Diagnostic accuracy was assessed in parathyroid SPECT of 60 patients with primary hyperparathyroidism. Phantom and patient SPECT scans were reconstructed by two different algorithms (ReSPECT and HOSEM). Two blinded nuclear medicine physicians interpreted the patient scans in random order, without knowledge of any clinical data. Subjective image quality and certainty of diagnosis were scored.

Results: Significantly lower relative noise levels (0.08 versus 0.13, $p=0.042$) and higher image spatial resolution (6.6 mm versus 17.1 mm full width at half maximum in the transaxial plane) were found in the phantom studies for ReSPECT as compared to HOSEM, respectively. In the clinical scans mean TBR (target to background ratio) was higher for ReSPECT (4.53 versus 2.97, $p<0.001$). There was no significant difference in region-based sensitivity (73% versus 68%, $p=0.538$) or specificity (93% versus 92%, $p=0.815$) for ReSPECT and HOSEM, respectively. Subjective image quality was higher for ReSPECT, as well as certainty of interpretation ($p<0.001$).

Conclusion: There is significant difference in image quality between commercially available algorithms for tomographic reconstruction of SPECT scans, which appears to have an impact on diagnostic accuracy and interpreter certainty. Therefore, we believe that systematic comparison of different SPECT reconstruction algorithms should be conducted to ensure a reproducible high image quality in clinical SPECT.

Key words: SPECT, hyperparathyroidism, image quality, image reconstruction, parathyroid adenoma.

Introduction

There are many different commercially available algorithms for iterative tomographic reconstruction of Single Photon Emission Computed Tomography (SPECT). Although the manufacturers provide data about the achievable resolution and other factors, direct comparison of different algorithms is usually not done. However, the use of certain reconstruction software packages could considerably influence the image quality. These outcome measures can be of interest as they may lead to more accurate diagnosis, especially in low lesion-to-normal tissue ratios. Also, algorithms optimized for reconstruction of low-count images may allow administration of lower activity doses or decrease the scanning time, directly influencing the patient's radiation burden.

Parathyroid scintigraphy is used to pre-operatively localize the adenoma causing primary hyperparathyroidism (pHPT). pHPT is one of the most common endocrine diseases and the number one cause of hypercalcaemia in non-hospitalized patients. In 80-85% of patients with pHPT the underlying cause is a single parathyroid hormone-secreting adenoma. The treatment of choice for pHPT is surgery; using bilateral neck exploration, 92-98% of patients with pHPT can be cured [1-5].

Preoperative localization of the lesions, for example by [^{99m}Tc]-sestamibi or [^{99m}Tc]- tetrofosmin SPECT [6,7], can be used to limit surgery to a unilateral or even a minimally invasive approach under local anesthesia with similar high success rates as bilateral neck exploration, leading to lower morbidity, shorter hospital stays and lower costs [8,9]. For this purpose, high sensitivity of SPECT in the detection of pathological parathyroid tissue is essential [10,11]. The sensitivity is inversely linked to the size of the

lesions, which is attributable to spatial resolution, tracer uptake of lesions, and the partial volume effect [12].

The purpose of this study was to determine the impact of different commercially available SPECT algorithms on clinical sensitivity in low-contrast imaging, using parathyroid SPECT with [^{99m}Tc]-sestamibi. We aimed to determine whether clinical test characteristics (sensitivity and specificity), certainty of interpretation and inter-observer consistency vary between these algorithms. We have therefore compared image quality in patient parathyroid scans and have confirmed the results by phantom scans.

Methods

Image quality

Phantom scans were acquired with decreasing target-to-background ratios (TBRs). The NEMA (National Electrical Manufacturers Association) NU-2 PET chest phantom was used, containing six fillable spheres (inner diameter of 10, 13, 17, 22, 28 and 37 mm) as well as a fillable background volume, considered ideal for simulating parathyroid lesions and cervical/thoracic background activity. The phantom background volume was filled with [^{99m}Tc]-pertechnetate using an estimated physiologic background activity concentration of 8.5 kBq/ml based on biodistribution [13] and body statistics [14] data. The range of target activity concentrations was determined on previously described clinical TBRs [15], ranging from 1.3 to 2.9. For determination of image resolution, a 1-millimeter droplet of undiluted [^{99m}Tc]-pertechnetate (37 MBq) placed in the exact center of the field of view was scanned using the same settings.

Clinical test characteristics

To determine the diagnostic potential of each reconstruction algorithm, 60 consecutive [^{99m}Tc]-sestamibi SPECT scans of patients with biochemical pHPT (table 1) who consecutively underwent surgery between 01-01-2001 and 12-31-2002 in a university hospital serving as tertiary center for endocrine diseases were used. The field of view included the neck and upper mediastinum. Furthermore, individual patient clinical data was collected from patient files which included biochemical data, surgical and histopathological reports and clinical follow-up.

Acquisition Settings

All phantom scans were acquired on a Siemens dual head E-cam gamma camera (Siemens Healthcare, Erlangen, Germany) using the local parathyroid protocol, in accordance with the SNM (Society of Nuclear Medicine) and EANM (European Association of Nuclear Medicine) guidelines[16,17]: 360 degrees rotation, 128 projections, 30sec/projection, 128*128 matrix, non-circular orbit, high-resolution low-energy parallel-hole collimator. For the acquisitions used to determine image resolution, a circular orbit with a radius of 30 cm was chosen, which was comparable to the other phantom scans.

Clinical imaging was performed as a dual-phase single tracer examination after intravenous injection of 600-700 MBq of [^{99m}Tc]-sestamibi. The scans were acquired on three different gamma cameras equipped with a high-resolution low-energy parallel-hole collimator (Siemens Multispect, 3 heads, 64 projections per head (n=45), Siemens E-cam, 2 heads, 64 projections per head (n=9) and Siemens Diacam with 1 head and 48 projections (n=6), Siemens Healthcare, Erlangen, Germany).

Reconstruction

Two commercially available iterative reconstruction algorithms were used: HOSEM version 3.11 (Hermes Medical Solutions, Stockholm) and ReSPECT version 2.1 (Scivis GmbH, Göttingen). In both cases, the most recent versions available at the time point of the study were used. HOSEM (Hermes-OSEM) is based on a fast iterative tomographic reconstruction method (Ordered Subsets Expectation Maximization, OSEM). It detects ellipsoidal body contours for attenuation correction (AC) [18] using thresholds of tracer concentrations for soft tissue and air distinction. ReSPECT is an iterative reconstruction algorithm with ordered subsets and weighted back projection. It defines free-form body contours based on thresholds for AC [18]. Furthermore, its algorithm includes corrections for small gantry movements, truncated data (the total attenuation coefficient is different in case parts of a subject are not inside the measuring boundaries) and utilizes a specific noise suppression algorithm. In contrast, HOSEM does not use resolution recovery.

The settings for both reconstruction algorithms were optimized by the manufacturers Scivis (ReSPECT) and Hermes (HOSEM) based on a random sample of 6 patient data sets. The aim of the optimization was the best subjective image quality, best reflecting the clinical situation. These optimized settings were used for the reconstruction of all SPECT acquisitions including the phantom scans. All scans were tomographically reconstructed including the entire scan region cranial to the upper border of the left ventricle. For automatic determination of body contours the best suiting threshold values based on the first ten scans was selected and used for all further scans. Slice thickness and pixel size were identical; the differences between both

reconstruction algorithms were negligible (4.795mm for HOSEM and 4.7951999mm for ReSPECT).

Objective image quality measures

Image activity recovery was quantified using a 3D volume-of-interest (VOI) of the exact volume of the phantom's spheres. Since at lower TBRs the exact location of the smaller spheres proved difficult to determine, these VOIs were drawn on the images with the highest TBR. All other scans were first co-registered with this scan and the defined VOIs were copied. The percentage recovered contrast was defined according to the NEMA NU 2-2007 Performance Measurements of PET standard:

$$\text{Percent Recovered Contrast} = \frac{T_{\text{mean}} / B_{\text{mean}} - 1}{TBR_{\text{actual}} - 1} * 100\%$$

Where T_{mean} is the mean tracer concentration in the target VOI, B_{mean} is the mean signal of the 2 background VOIs nearest to the target and TBR_{actual} is the actual ratio of tracer concentrations between the target and background compartment of the phantom.

Image noise in both the phantom scans and the clinical scans was determined by the relative noise level, ($RNL = B_{SD} / B_{\text{mean}}$) where B_{SD} is the standard deviation of the tracer concentration in a VOI drawn in the image background and B_{mean} is the mean tracer concentration in the same VOI. The background area for the phantom scans was defined as a 37 mm spherical volume in the background area of the phantom (figure 1). In the patient data, 37 mm spheres were drawn in the trapezius muscle as a background area because of its relative homogeneity.

To determine the image resolution, the full width at half maximum (FWHM) of the reconstructed images of the point-source scan was determined by non-linear least

squares fitting of a Gaussian function to a profile through the center of the point source in each direction (Inveon Research Workplace version 2.2 (Siemens, Knoxville, TS)). As this definition of resolution in iterative algorithms is highly dependent on the number of iterations, reconstruction was repeated along a range of MLEM (maximum likelihood expectation maximization) equivalent number of iterations, which is defined as the product of the number of OSEM-iterations and –subsets.

For the clinical scans image, contrast was quantified as the target-to-background ratio ($TBR = T_{mean} / B_{mean}$) where T_{mean} is the average tracer concentration of a VOI (3D isocontour of 50% of maximum activity concentration) covering the lesion pointed out by the most experienced observer and B_{mean} is the average signal intensity of the background VOI in the trapezius muscle.

Clinical interpretation

All clinical scans were randomly interpreted by two independent blinded physicians: one experienced nuclear physician (M.G. more than 500 parathyroid scans) and one resident in his last year of training (J.W.A.P.). The observers were asked to score the images on contrast, noise, resolution and interpretive certainty, using a 5-point Likert scale. Higher scores mean better contrast, lower noise, higher resolution and higher certainty of interpretation, respectively.

The localizations of the parathyroid adenomas were compared to surgical, histopathological and follow-up data (serving as standard of reference) to determine the sensitivity (ratio of true positives of all diseased) and specificity (ratio true negatives of

all healthy). Patient-based specificity could not be determined since no patients without a parathyroid adenoma were included. Therefore, region-based sensitivity and specificity were determined using four regions: left, right, mediastinum and ectopic, based on clinical relevance (unilateral surgical exploration). Because of the expected difficulties with correlations of descended or ascended adenomas with the result of surgery (a descended upper parathyroid may be considered a lower parathyroid, while for the surgeon it remains an upper gland), it was chosen not to differentiate between superior and inferior parathyroid glands.

Finally, the occurrence of any reconstruction artifacts identified in the patient SPECT scans were described and compared both in frequency and interference with diagnostics.

Statistical analysis

For normally distributed data (assessed by the Shapiro-Wilk statistic, New Brunswick, NJ, USA [19]), the means (\pm SD) are presented. Otherwise, the medians (interquartile range, IQR) are presented.

Pair-wise comparison of phantom TBRs, RNLs and image resolution was performed using the non-parametric Wilcoxon test. Agreement between both observers was determined using the κ statistic and McNemar's test was used to compare sensitivity and specificity, both inter-observer and inter-regional. Correlations between subjective image contrast and noise on the one hand and TBR and RNL on the other hand were quantified using Spearman's ρ . Multivariate logistic regression modeling was performed to assess any effects of thyroid nodules and size, lesion size, previous neck surgery and the type of camera used on the true-positive ratio of SPECT scans using an

iterative step-wise (forward and backward) model based on likelihood ratio with a cut-off for variable inclusion of $p=0.10$.

All statistical analyses were performed using SPSS (Statistical Package for the Social Sciences) version 16.0 (SPSS Inc, Chicago, United States). Two-sided statistical significance was set at $p<0.05$.

Results

Phantom scans: Image quality

All fillable spheres could be visualized adequately for TBRs ≥ 1.7 . For all TBRs used, reconstructions by ReSPECT led to significantly higher percentages recovered contrast than HOSEM (figure 2). All RNLs of the phantom images were lower for ReSPECT than for HOSEM. For the phantom with the highest TBR this was 0.09 for ReSPECT versus 0.15 for HOSEM. Comparing the RNLs for all TBRs, there was a significant difference in mean RNL for ReSPECT versus HOSEM (mean 0.09 (SD \pm 0.008) versus 0.13 (SD \pm 0.027), Wilcoxon $p = 0.042$).

The axial FWHM was 16.7 mm for HOSEM and 7.6 mm FWHM for ReSPECT at the manufacturer's recommended number of iterations/subsets. The transaxial FWHM was 17.1 mm for HOSEM and 6.6 mm for ReSPECT (figure 3). Altering the number of MLEM-equivalent iterations used in the reconstruction positively affected the resolution for ReSPECT (figure 4).

Clinical scans: Image quality

A total of 60 lesions were quantified. Although the correlation between subjective and objective measures of image contrast was weak but significant ($\rho=0.385$, $p<0.001$), both the objective TBR (median 4.53 (IQR 2.60-4.2) versus 2.97 (IQR 2.06-4.09), $p<0.001$) as the subjective contrast score (mean 3.37/5 versus mean 2.93/5, $p<0.001$) was higher for ReSPECT. The concordance of the subjective contrast determined by both observers, however, was weak ($\kappa=0.098$, $p=0.078$).

There was a low but significant correlation between subjective noise ratings and quantified noise levels ($\rho=0.214$, $p=0.014$). In opposite to the phantom scans, the clinical reconstructions yielded slightly higher noise levels for ReSPECT (median 0.28 (IQR 0.21-0.35)) versus HOSEM (median 0.22 (IQR 0.14-0.38)), $p<0.001$). However, both observers found ReSPECT reconstructions to have less noise (mean subjective noise score 3.15/5 versus 3.71/5, $p<0.001$, $\kappa=0.040$, $p=0.441$).

Subjective resolution ratings were significantly higher for ReSPECT (mean 3.1/5 versus mean 2.5/5, $p<0.001$, $\kappa=0.056$, $p=0.282$).

Clinical scans: Diagnostic accuracy

Region-based sensitivity and specificity (table 2) was higher for ReSPECT (not significant). The agreement between both observers was moderate for both HOSEM ($\kappa=0.405$, $p=0.001$) and ReSPECT ($\kappa=0.130$, $p=0.351$).

The interpreters were significantly more confident in their readings when using ReSPECT (mean 2.98/5, HOSEM mean 2.62/5, $p<0.001$). Comparing only true positive lesions ($n=44$), the certainty was not significantly different. No relations between interpretive certainty and subjective contrast, noise or resolution could be detected.

Artifacts

Both interpreters reported significantly more reconstruction artifacts with HOSEM (92% versus 41%, $p < 0.001$) with a higher rating of impact on the diagnosis ($p < 0.001$). These artifacts were mostly stripe artifacts in the transversal plane for HOSEM, while ReSPECT artifacts were photopenic areas and occasionally a single stripe at chin height (figure 6). None of these artifacts could be directly linked to the type of camera used.

Multivariate analysis

None of the following factors investigated by multivariate analysis showed an independent significant relation with the true positive rate of the scans: existence of thyroid nodules, thyroid enlargement, previous neck surgery, lesion size, camera type, existence of image artifacts.

Discussion

To the best of our knowledge, this is the first clinical study to directly compare different commercially available reconstruction algorithms for objective and subjective image quality, diagnostic test characteristics and observer variability. Our results clearly demonstrate the potential clinical impact of different commercially available algorithms for iterative tomographic reconstruction on image quality (noise, contrast recovery and spatial resolution). Most healthcare professionals in nuclear medicine believe that the camera will have the most important influence on image quality. Many vendors of camera systems stress the improved image quality of their newest products in sales brochures. Our results indicate that the reconstruction software package may have a

much larger influence on the image quality than the camera system. It is well-known that the type of reconstruction (FBP versus OSEM) has a considerable effect on image quality [20]. No differences between reconstruction software from different vendors had been found in a previous study using numeric phantom simulations [21]. In contrast, we have used images with a low count rate and have evaluated the clinical impact. We therefore believe that the differences between the reconstruction software packages may be most apparent in low count-rate scans such as parathyroid scans, somatostatin receptor scans etc.

ReSPECT performed better in this study, raising the question whether this may be due to a favorable trade-off in the reconstruction settings for the specific study design. Based on our results, we believe that this is not the case. Although image noise and contrast recovery are often positively related (smoothing, for example, will lead to lower noise but also to lower contrast recovery), this was not the case in our phantom studies: ReSPECT reconstruction yielded both better results in noise and contrast, rendering strategies such as smoothing unlikely to be responsible for the observed differences.

The applied definition for image resolution by determination of the FWHM of reconstructed images of a point source scan is susceptible to the number of iterations in iterative tomographic reconstruction algorithms. In order to describe this relation we have determined the resolution in a wide range of MLEM-equivalent numbers of iterations. As expected, the resolution improves for ReSPECT as a function of the number of iterations. This was not the case for HOSEM, which may be due to ReSPECT's resolution recovery (which HOSEM lacked). ReSPECT's higher resolution might explain the improved percentage recovered contrast in especially the smaller

spheres, as a higher resolution will result in a decreased influence of partial-volume effects. We are aware that iterative reconstruction algorithms when applied to point sources or line sources do not yield unambiguous values of the observed FWHMs of these sources. The results depend e.g. on the number of iterations, subsets, post recon filtering, specific choices dealing with the projection matrix, etc. However, we found it nevertheless useful to present the results for a point source as well and to investigate the effect of the number of iterations.

The subjective resolution, contrast and noise correlated well with the objectively measured values from the clinical scans as well as the phantom scans. The results for lowest image noise in the phantom scans (ReSPECT) compared to the clinical scans (HOSEM) are contradictory. However, the higher spatial resolution of ReSPECT would lead to a higher recovery of tissue heterogeneity. As the definition used for noise in our study assumes homogenous tracer distribution in the VOI under investigation, the use of inhomogeneous tissues (trapezius muscle) would lead to an artificially higher value for the RNL for ReSPECT in the patient scans as compared to HOSEM. A lower resolution leads to smoothing of tissue heterogeneity due to partial-volume-effects, resulting in a lower RNL for HOSEM.

The improved image quality characteristics found in the phantom studies indicated towards a potential positive influence of ReSPECT on both, sensitivity and specificity of detection of parathyroid pathology on late [^{99m}Tc]-sestamibi parathyroid SPECT scans. The actually detected differences in sensitivities between both algorithms would have led to the avoidance of one futile bilateral neck exploration in 20 patients. This means that in 5% of the cases, minimally invasive neck surgery would suffice instead of bilateral exploration. However, the observed difference did not reach statistical significance.

Based on the largest difference in sensitivities between both algorithms, we have calculated a required sample size of at least 746 pHPT patients to obtain sufficient statistical power to reach statistical significance. In view of the limited number of patients available for parathyroid scintigraphy, this does not seem to be a realistic aim. For further studies comparing different reconstruction algorithms, it may be advantageous to choose other clinical scans than parathyroid scans.

Several advanced iterative reconstruction software packages are currently available, including not only ReSPECT and Hybrid Recon, but also Flash 3D (Siemens), Astonish (Philips), Evolution (GE) and others. The vendors claim an improved image quality in low-count rate scans including the option of shortened acquisition times. We believe that, in view of our results, it is important to compare different reconstruction methods with variable count statistics in phantom and clinical studies, as clinically relevant differences may exist. This is especially interesting in respect to shortened acquisition times or the use of lower tracer activities.

Finally, we would like to remark that both vendors recently brought new versions of their reconstruction software to the market. The results presented here are therefore not applicable to the new software packages.

Conclusion

The algorithm used for image reconstruction directly influences the image quality and may therefore have an impact on patient management. Although differences in image quality are attributable to the performance of gamma cameras in the perception of many professionals working in nuclear medicine, our study indicates that the reconstruction

algorithm may have a higher influence on image quality. We should be aware that while gamma cameras have to be certified for approval as medical devices in the EU, this is not the case for image reconstruction software. Therefore, although a certain level of performance can be guaranteed for gamma cameras, this is not the case for reconstruction software. In view of our results, we believe that systematic comparison of image reconstruction software is advisable.

Acknowledgements:

D.Vriens' PhD project is funded in part by The Netherlands Organization for Health Research and Development (ZonMW).

References

1. Adami, S, Marcocci, C, Gatti, D. Epidemiology of primary hyperparathyroidism in Europe. *J Bone Miner Res* 2002; **17 Suppl 2**:N18-23.
2. Palestro, CJ, Tomas, MB, Tronco, GG. Radionuclide imaging of the parathyroid glands. *Semin Nucl Med* 2005; **35**:266-276.
3. Rose, DM, Wood, TF, Van Herle, AJ, Cohan, P, Singer, FR, Giuliano, AE. Long-term management and outcome of parathyroidectomy for sporadic primary multiple-gland disease. *Arch Surg* 2001; **136**:621-626.
4. Udelsman, R. Six hundred fifty-six consecutive explorations for primary hyperparathyroidism. *Ann Surg* 2002; **235**:665-670; discussion 670-662.
5. Wang, CA. Surgical management of primary hyperparathyroidism. *Curr Probl Surg* 1985; **22**:1-50.
6. Shafiei, B, Hoseinzadeh, S, Fotouhi, F, Malek, H, Azizi, F, Jahed, A, *et al*. Preoperative ^{99m}Tc-sestamibi scintigraphy in patients with primary hyperparathyroidism and concomitant nodular goiter: comparison of SPECT-CT, SPECT, and planar imaging. *Nucl Med Commun* 2012; **33**:1070-1076.
7. Taieb, D, Urena-Torres, P, Zanotti-Fregonara, P, Rubello, D, Ferretti, A, Henter, I, *et al*. Parathyroid scintigraphy in renal hyperparathyroidism: the added diagnostic value of SPECT and SPECT/CT. *Clin Nucl Med* 2013; **38**:630-635.
8. Chen, H, Sokoll, LJ, Udelsman, R. Outpatient minimally invasive parathyroidectomy: a combination of sestamibi-SPECT localization, cervical block anesthesia, and intraoperative parathyroid hormone assay. *Surgery* 1999; **126**:1016-1021; discussion 1021-1012.
9. Goldstein, RE, Blevins, L, Delbeke, D, Martin, WH. Effect of minimally invasive radioguided parathyroidectomy on efficacy, length of stay, and costs in the management of primary hyperparathyroidism. *Ann Surg* 2000; **231**:732-742.
10. Gotthardt, M, Lohmann, B, Behr, TM, Bauhofer, A, Franzius, C, Schipper, ML, *et al*. Clinical value of parathyroid scintigraphy with technetium-99m methoxyisobutylisonitrile: discrepancies in clinical data and a systematic metaanalysis of the literature. *World J Surg* 2004; **28**:100-107.
11. Dasgupta, DJ, Navalkisoor, S, Ganatra, R, Buscombe, J. The role of single-photon emission computed tomography/computed tomography in localizing parathyroid adenoma. *Nuclear Medicine Communications* 2013; **34**:621-626.
12. Stephen, AE, Roth, SI, Fardo, DW, Finkelstein, DM, Randolph, GW, Gaz, RD, *et al*. Predictors of an accurate preoperative sestamibi scan for single-gland parathyroid adenomas. *Arch Surg* 2007; **142**:381-386.
13. Wackers, FJ, Berman, DS, Maddahi, J, Watson, DD, Beller, GA, Strauss, HW, *et al*. Technetium-99m hexakis 2-methoxyisobutyl isonitrile: human biodistribution, dosimetry, safety, and preliminary comparison to thallium-201 for myocardial perfusion imaging. *J Nucl Med* 1989; **30**:301-311.
14. Rush, EC, Freitas, I, Plank, LD. Body size, body composition and fat distribution: comparative analysis of European, Maori, Pacific Island and Asian Indian adults. *Br J Nutr* 2009; **102**:632-641.
15. Ruf, J, Seehofer, D, Denecke, T, Stelter, L, Rayes, N, Felix, R, *et al*. Impact of image fusion and attenuation correction by SPECT-CT on the scintigraphic detection of parathyroid adenomas. *Nuklearmedizin* 2007; **46**:15-21.

16. Greenspan, BS, Dillehay, G, Intenzo, C, Lavelly, WC, O'Doherty, M, Palestro, CJ, *et al.* SNM practice guideline for parathyroid scintigraphy 4.0. *J Nucl Med Technol* 2012; **40**:111-118.
17. Hindie, E, Ugur, O, Fuster, D, O'Doherty, M, Grassetto, G, Urena, P, *et al.* 2009 EANM parathyroid guidelines. *Eur J Nucl Med Mol Imaging* 2009; **36**:1201-1216.
18. Chang, LT. Method for Attenuation Correction in Radionuclide Computed Tomography. *Ieee Transactions on Nuclear Science* 1978; **25**:638-643.
19. Shapiro, SS, Wilk, MB. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* 1965; **52**:591-611.
20. Knoll, P, Kotalova, D, Kochle, G, Kuzelka, I, Minear, G, Mirzaei, S, *et al.* Comparison of advanced iterative reconstruction methods for SPECT/CT. *Z Med Phys* 2012; **22**:58-69.
21. Seret, A, Forthomme, J. Comparison of different types of commercial filtered backprojection and ordered-subset expectation maximization SPECT reconstruction software. *J Nucl Med Technol* 2009; **37**:179-187.

Tables

Table 1: Population Characteristics. ULN: upper limit of normal in laboratory. IQR: Interquartile range *Lower number due to missing data

Population characteristics	n=60
Female	81 %
Median age in years (range)	60.9 (16.5-78.8)
Mean preoperative PTH in pmol/l (ULN: 6.80) (IQR)	13.7 (9.3–22.0) n=52*
Mean preoperative Ca ²⁺ in mmol/l (ULN: 2.65) (IQR)	2.80 (2.73-3.00)
Previous neck surgery	
For Thyroid Disease	5.0%
For Parathyroid Disease	6.7%
Thyroid anomalies on ultrasonography	
Enlarged (♀ ULN: 18ml, ♂ ULN: 25ml)	31.7%
Nodules	38.3%
Peroperative localization adenomas	n=60
Left parathyroid	49%
Right parathyroid	48%
Mediastinum	3%
Mean weight of removed adenomas in g (range)	0.9 (0.2–11) n=57*
Mean postoperative total Ca ²⁺ in mmol/l (ULN: 2.65) (range)	2.30 (2.00-2.60)

Table 2: Region-based Sensitivity, Specificity and Accuracy. Based on four regions (left parathyroid, right parathyroid, mediastinum, ectopic) based on 60 patient **scans**.

*CI: Confidence Interval, †McNemar test.

N=240	HOSEM (95% CI*)	ReSPECT (95% CI*)	Significance†
Observer 1			
Sensitivity	68.3% (57.0-78.0)	73.0% (62.1-82.1)	0.581
Specificity	91.5% (87.1-94.7)	92.7% (88.5-95.6)	0.815
Accuracy	85.4% (80.8-89.3)	87.5% (83.1-91.1)	0.473
Observer 2			
Sensitivity	65.1% (53.7-75.3)	66.7% (55.3-76.7)	1.000
Specificity	86.4% (81.1-90.6)	88.7% (83.7-92.5)	0.557
Accuracy	80.8% (75.7-85.2)	82.9% (78.0-87.1)	0.542

Figure captions:

Figure 1: Phantom studies definitions for volumes of interest. Left: transverse section. Right: sagittal section. Black circles represent the fillable spheres of the phantom. The dark grey circles represent volumes of interest in the background (light gray) of the phantom. The air-containing cylinder in the middle of the phantom is displayed in a black/white pattern.

Figure 2: Contrast recovery as function of sphere size and target-to-background ratio for both algorithms. Inset: A single transversal plane of the three scans through the center of the spheres. As TBR increases, the fillable spheres become better distinguishable from the background, especially the larger ones, but at lower TBR they are lost in the noise of the background. TBR: tumour-to-background ratio.

Figure 3: Recovered signal of a point-source scan using the recommended settings by each manufacturer fitted with a Gaussian curve, through the center of the point-source in the transaxial plane, normalized for maximum signal. The width of these curves is a measure of the resolution of the images, where a smaller width equals a higher resolution. Resolution was defined as the width at 50% of maximum height (FWHM). In this graph ReSPECT's FWHM (6.6mm) is less than half of that of HOSEM (17.1mm).

Figure 4: Influence of the number of maximum-likelihood expectation maximization (MLEM) equivalent iterations on resolution for reconstructing a point source. FWHM: full

width at half maximum. As the number of MLEM-equivalent iterations increase, ReSPECTS' resolution decreases from 9-13 mm to 2-4 mm. HOSEMS' resolution however stays roughly the same at around 17mm.

Figure 5: Example scan showing a single parathyroid nodule in the upper mediastinum of the patient. Two reconstructions of the same SPECT image. Left: HOSEM. Right: ReSPECT. The parathyroid adenoma can more clearly be delineated on the ReSPECT reconstructed images although it is also visible on the HOSEM reconstructed images.

Figure 6: Example scans with typical artifacts. Two reconstructions of the same SPECT image, showing severe stripe artifacts on the HOSEM reconstruction (left) and a single stripe artifact at the chin on the right reconstruction (ReSPECT).











