Cardiac Phantom Measurement Validating the Methodology for a Cardiac Multi-Center Trial with Positron Emission Tomography

Johan Nuyts [1], Luc Mortelmans [1], Frans Van de Werf [2], Jacques Djian [3], Gianmario Sambuceti [4], Marcus Schwaiger [5], Paul Touboul [6], Alex Maes[1], and the PSALM study group [7]

[1] Nuclear Medicine, K.U.Leuven, Leuven, Belgium
[2] Cardiology, K.U.Leuven, Leuven, Belgium
[3] Wyeth Research, Paris la Defense Cedex, France
[4] Instituto di Fisiologia Clinica CNR, Pisa, Italy
[5] Nuklearmedizinische Klinik und Poliklinik, TU Muenchen, Muenchen, Germany
[6] Hôpital Cardio-Vasculaire et Pneumologique, Lyon, France

[1]*Correspondence to:* J Nuyts, Nuclear Medicine, UZ Gasthuisberg, Herestraat 49, B3000 Leuven, Belgium. E-mail: Johan.Nuyts@uz.kuleuven.ac.be

[7]The PSALM Study Group comprises: P Mertens (Nuclear Medicine, U.Z. Gasthuisberg, Belgium), P Melon (Cyclotron Research Center, University de Liège, Belgium), N Preumont (Nuclear Medicine, Hôpital Erasme, Brussels, Belgium), M Janier (CERMEP, Lyon, France), JM Ribeiro (CEA, Service Hospitalier Frédéric Joliot, Orsay, France), A Maldonado, (Centro PET Complutense, Madrid, Spain), M Simó (CETIR Unitat PET, Barcelona, Spain). J García Velloso (Servicio Medicina Nuclear, PET, Pamplona, Spain), B Långström (PET centrum, Uppsala, Sweden), M Böttcher, (Skejby Sygehus, Arhus, Denmark), M Schwaiger (Nuklearmedizin, TU Muenchen, Germany), G Sambuceti (Instituto di Fisiologia Clinica CNR, Pisa, Italy), R Chisin (Biophysics and Nuclear Medicine, University Hospital Hadassah, Jerusalem, Israel).

Abstract

In an ongoing international multi-center trial, positron emission tomography (PET) is being used to evaluate the effect of a new P-selectin antagonist on the infarct size in patients with acute myocardial infarction, treated with thrombolysis. Although it is possible to correct for site-dependent factors, it is desirable to reduce these factors to a minimum. Therefore, acquisition and reconstruction protocols have been defined that can be closely followed by all participating centers. The resulting reconstructed images are transferred to the core center for central processing with semi-automatic software. This paper reports on the experiment that was carried out to assess the inter-center reproducibility of defect size determination with this protocol. Also the spatial resolution of the short axis slices was examined. In addition, the analysis procedure was applied to normal PET-studies to evaluate the specificity of perfusion defect detection.

The transmural cold defect in the phantom occupied 14.8% of the left ventricular area. The automated analysis was applied to the phantom measurements from the fourteen participating PET cameras. It yielded an accurate estimate of 15.1 % with a standard deviation of 0.6%, indicating excellent reproducibility. The spatial resolution in the short axis slices was similar for all PET systems: 9.6 mm $\pm$ 0.8 mm. The same procedure produced a defect size of zero in the studies of the normal volunteers.

This study indicates that cardiac studies from multiple PET-systems can be pooled for statistical analysis.

Keywords: Myocardial perfusion - defect size - PET

I. Introduction

In the ongoing international clinical multi-center trial "Psalm" (P-Selectin Antagonist Limiting Myonecrosis), positron emission tomography is being used to evaluate the effect of a new P-selectin antagonist on the infarct size in patients with acute myocardial infarction, treated with thrombolysis. The inclusion criteria select patients with severe myocardial infarctions: persistent ST-segment elevation of at least 0.2 mV in 2 or more limb leads or 0.3 mV in 2 or more contiguous precordial leads. The primary end points of the study are the perfusion and metabolic defect sizes as assessed with positron emission tomography using ^{13}N-ammonia and ^{18}F fluorodeoxyglucose (FDG). The defect size is expressed as a fraction of the total area of the left ventricular wall.

Fourteen PET-systems in thirteen centers are being used. They include five different types of PET systems (GE Advance, Ecat Exact, Ecat HR+, Ecat 951 and Posicam HZL/R), and different releases of the system software. Consequently, it cannot be excluded that the data depend, to some extent, on the site where the PET-scan has been performed. With appropriate randomization and statistical analysis, it is always possible to correct for the presence of unwanted factors, but to maximize the statistical power, the influence of such factors should be minimized.

To minimize the influence of the site, the amount of work done in a single core center should be maximized. However, because of the use of cameras from different manufacturers, centralized reconstruction is not feasible, and reconstruction has therefore been done locally. A detailed acquisition and reconstruction protocol has been defined, which can be closely followed by all participating centers. The reconstructed images are then transferred to the core center and are processed and analysed in an identical way.

To validate this approach, a cardiac phantom experiment was designed. The same phantom was sent to all sites, and scanned and reconstructed using the trial protocol. The file with the reconstructed volume was transferred to the core center and analysed with the software used for the trial. The defect is determined by applying a simple thresholding procedure to the polar map representing the myocardial tracer uptake. The phantom experiment permits evaluation of the reproducibility of defect size determination. In addition, it permits estimation of the effect of the entire procedure on the spatial resolution of the final images. Finally, the same analysis procedure was also applied to images of normal volunteers to assess the probability of a false positive outcome.

II. Material and Methods

The protocol includes the dynamic acquisition in 2D mode (septa in the field of view) of ^{13}N-ammonia images, permitting absolute quantification of myocardial blood flow with tracer kinetic modelling analysis. However, the primary end point is relative defect size as determined from baseline ammonia and FDG studies. This requires accurate delineation of the left ventricular wall, and accurate definition of the perfusion defect.

To optimize the performance of the semi-automatic delineation procedure, maximum-likelihood reconstruction has been used rather than filtered backprojection.

Each frame is reconstructed at the site using OSEM (maximum-likelihood expectation-maximization, accelerated with ordered subsets [1]), applying 5 iterations with 8 subsets and no post-smoothing. The files with the reconstructed volumes are transferred to the core center for further centralized processing and analysis. All volumes are first post-smoothed with a 3D Gaussian kernel with a full width at half maximum (FWHM) of 6 mm in all directions.

The system resolution of the PET-camera is position dependent. However, it varies only moderately over the main portion of the field of view and the variation is usually small compared to the final image resolution. The finite resolution of the PET camera is ignored in most clinical implementations (of both filtered backprojection and maximum-likelihood reconstruction), allowing it to propagate into the final reconstruction. Thus, the final resolution in the reconstructed image is a combination of the system resolution and the impulse response of the reconstruction algorithm.

Filtered backprojection is linear and shift invariant, so it does not contribute any position dependence to the spatial resolution. In contrast, maximum-likelihood reconstruction is neither linear nor shift invariant, and at low iteration numbers, its position dependence may dominate that of the system. However, the maximum-likelihood solution is (virtually) unbiased [2], implying that at very high iteration numbers, the impulse response function becomes very sharp (converging to a Dirac impulse). Subsequent convolution with a Gaussian kernel then produces a virtually position independent impulse response for the reconstruction algorithm. It is hypothesized that with 5×8 OSEM iterations and a 6 mm FWHM Gaussian kernel, the resolution dependence of the algorithm is acceptably small, compared with the average resolution of the final reconstruction. This hypothesis was verified with experiments described below.

A. The cardiac phantom

As an introduction to the actual clinical study, a multi-center phantom experiment was carried out for the following purposes:

1. Technical quality assurance: successful completion of the phantom experiment in each of the sites proves that the sites are able to apply the requested reconstruction procedure, and that problems related to file transfer to the core center (possibly different computer platforms and/or different file formats) have been solved.

2. To estimate the inter-center reproducibility of quantification of defect size, which is the primary end-point of the multi-center trial.

3. To estimate the site-dependence of spatial resolution, which may affect kinetic modelling results.

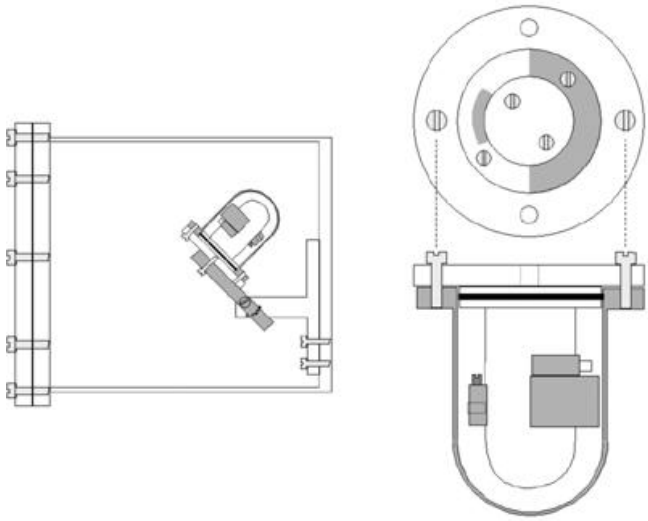

Figure 1: The cardiac phantom with the large transmural and small subendocardial inserts.

Figure 1 shows the phantom (Data Spectrum Corporation). It consists of a left ventricle with fillable wall and cavity, mounted in a 20 cm diameter cylinder. The left ventricular wall is 10 mm thick and contains two inserts, filled with non-radioactive water: a large transmural infarction (10 mm thick, 14.8 % of the total wall volume), and a small subendocardial infarction (5 mm thick, 4.8% of the total wall volume). The cavity is filled with non-radioactive water. The wall and the cylinder were filled with water containing a solution of ^{18}F. The concentration of the wall was five times higher than that of the cylinder, and the total radioactivity was about 20 MBq. The phantom was placed in the PET-system, and a 40 min emission scan and a 30 min transmission scan were performed. The transmission scan was done after overnight decay of the radioactivity, or immediately after the emission scan if the software provided a post-injection transmission protocol.

B. Defect size determination

As described above, a transaxial volume is reconstructed at the PET site and transferred to the core center, where it is post-smoothed with a 3D Gaussian filter with 6 mm FWHM.

Then, the long axis of the left ventricle is indicated manually in the transaxial volume image, the image is resampled into radial slices and the left ventricular wall is automatically delineated as described previously [3, 4]. This results in a three-dimensional delineation of the left ventricular wall. From the delineation, a polar map is computed [3, 4]. For each pixel of the polar map, a single value is computed, representing the regional transmural tracer uptake. In addition, the left ventricular wall area represented by each pixel is computed from the delineation. Thus, we can directly compute the area of the left ventricular wall corresponding to a region in the polar map, and the reported value is unaffected by the deformations applied to compute that map. By reporting the area rather than the volume, we eliminate the need to determine the wall thickness accurately, which is difficult to do in emission tomography.

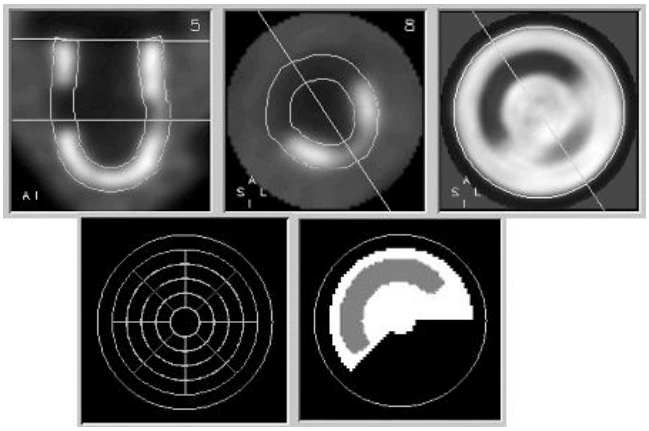

Figure 2: Radial slice, short axis slice and polar map of a phantom measurement. The first image of the bottom row shows the grid which is used to define the normalization value and the perfusion territory. The last image shows the perfusion territory used to exclude the subendocardial defect, and the perfusion defect defined with a threshold of 50%.

In figure 2, a grid consisting of 33 segments is shown. This grid is used for two purposes. Firstly, the mean tracer uptake in every segment is computed, and the maximum of those values is used to normalize the entire polar map (so the mean tracer uptake in this reference segment is regarded as 100%). Secondly, the operator can define perfusion territories by combining multiple segments into a single region of interest.

Finally, a threshold is applied to the pixels (not the segments) of the perfusion territory to define the perfusion defect. The procedure is illustrated in figure 2 for the analysis of a phantom experiment. The last image shows the perfusion territory in white, and the perfusion defect (for the large insert) obtained with a threshold of 50% in grey. To analyse the other defect, a second perfusion territory must be defined.

C. Reproducibility of defect size

The same phantom was scanned by the fourteen participating PET-cameras and analyzed as described above. Thresholds ranging from 30% to 80% were applied and the corresponding defect size was computed for both inserts. The polar map value in the transmural defect should be 0%, so assuming that the edge response function is sufficiently sharp, a good estimate of the defect should be obtained with a threshold of approximately (100+0)/2 = 50%. Similarly, a good estimate for the subendocardial defect is expected with a threshold of about (100 + 50)/ 2 = 75%. Moreover, if the behavior of all PET-systems is similar, they should all produce a similar relationship between threshold value and estimated defect size.

D. Reproducibility of the spatial resolution

The spatial resolution of the post-smoothed OSEM reconstructions was estimated from the short axis slices. The short axis slice was modelled as the convolution of a high resolution image with a Gaussian point spread function. Least squares fitting was applied to determine the parameters of the high resolution image and the width of the Gaussian kernel. The width of the kernel (characterized by its full width at half maximum or FWHM) was used as the estimate of the spatial resolution. The high resolution image consisted of an annulus with known dimensions (representing the wall). The position of the annulus and the activities in the wall, the cavity and the background were determined in the fitting procedure, together with the width of the Gaussian. In the calculation of the high resolution image, partial volume effects due to the finite pixel size were taken into account. The procedure is illustrated in figure 3. For each study, we analyzed two or three short axis slices in which the inserts were not visible, and the mean FWHM was reported.

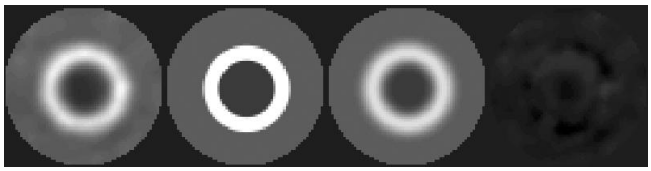

Figure 3: From left to right: 1) the short axis slice, 2) the synthetic high resolution image, 3) same after convolution with the Gaussian point spread function, 4) the difference between 1 and 3.

E. Effect of number of iterations on the spatial resolution

As stated above, we hypothesized that with post-smoothing after 5 $\times$ 8 OSEM iterations, the reconstruction algorithm will have a virtually position independent local impulse response. This is only true if the spatial resolution after 5 $\times$ 8 iterations is close to the resolution of the ML-solution, that is, the resolution after an infinite number of iterations. This is obviously difficult to verify exactly, but a very strong indication is obtained by comparing the resolution after 5 $\times$ 8 iterations with that obtained after "many" iterations. The Leuven phantom experiment was reconstructed twice, once with 5 $\times$ 8 and once with "many" iterations, implemented as 24 iterations of 8 subsets, 2 iterations of 4 subsets and two iterations of 1 subset. For both images, the spatial resolution was determined with the procedure described above.

F. Evaluation on normal volunteers

Maes et al [5] studied the myocardial tissue perfusion with PET and ^{13}N-ammonia early after successful thrombolysis. They found that in 11 of 30 patients with normalized myocardial flow below 50%, no recovery of left ventricular function occurred at 3 months. One of the aims of the ongoing trial is to test the hypothesis that the new agent reduces the occurrence of this no-reflow phenomenon. For that reason, a threshold of 50% was proposed to detect perfusion defects.

A second argument in favour of this threshold is produced by the phantom experiments: as shown below, a threshold of 50% yields an accurate estimate of the defect size produced by the transmural insert.

Because even with normal perfusion, the tracer uptake in the human heart is not as uniform as it is in the phantom, the error on the estimated defect size is expected to be larger than in the phantom measurements. In principle, this could lead

to a misclassification of normally perfused myocardium as a perfusion defect. Therefore, a final test of the 50% threshold was done by applying the procedure to a series of archived PET-studies of normal volunteers. We analysed 19 baseline ^{13}N-ammonia studies and 8 ^{18}F-FDG studies.

III. Results

A. Reproducibility of defect size

Figure 4 shows the relation between the threshold and the estimated defect size for the two inserts and the 14 PET-systems. Except for extremely low or high thresholds, the estimated defect sizes were very similar for all systems.

As expected, accurate estimates were obtained with a threshold of 50% for the transmural infarct, and with a threshold of 75% for the subendocardial infarct. The true relative area of the large insert was 14.8%, mean and standard deviation of the estimates with the 50% threshold were 15.1 $\pm$ 0.6. The true relative area of the small insert was 4.80%, and with the 75% threshold we obtained 4.89 % $\pm$ 0.51 %.

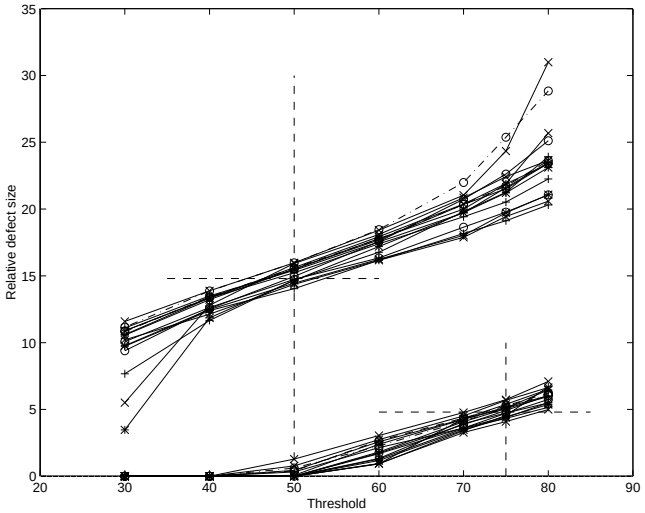

Figure 4: The phantom defect size as a function of the threshold, for the two inserts and the 14 PET-systems. The horizontal dashed lines show the true defect sizes.

B. Reproducibility of the spatial resolution

The spatial resolution of the short axis slices was 9.6 mm $\pm$ 0.8 mm, the maximum was 10.9 mm, and the minimum, 8.5 mm. We found no relation between the resolution and the system type. There was a weak but significant correlation between resolution and pixel size (r = 0.53, P = 0.025), and no significant correlation between resolution and plane separation.

C. Effect of number of iterations on the spatial resolution

The spatial resolution of the Leuven phantom image was 9.2 mm with post-smoothing after 5 $\times$ 8 iterations, and decreased to 8.1 mm when the procedure with increased iteration number was applied.

D. Evaluation on normal volunteers

Figure 5 shows the results obtained in the normal ammonia and FDG studies. With a threshold of 50%, defect size was zero in all cases. With thresholds above 60%, relative defect sizes exceeding 10% are obtained.

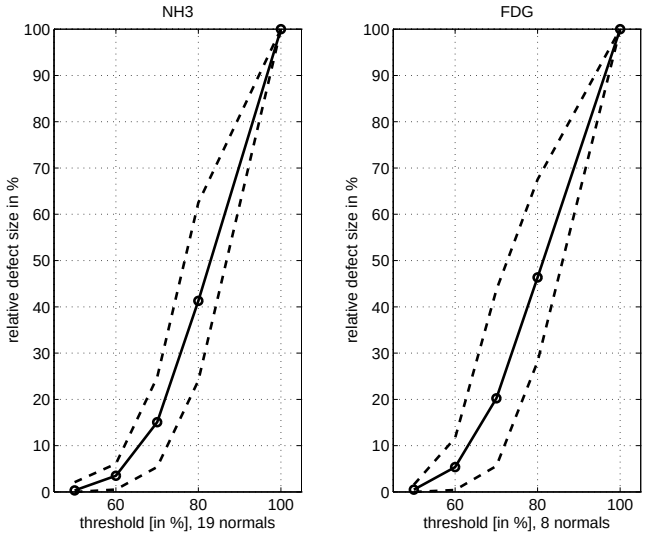

Figure 5: The defect size as a function of the threshold in normals. The left panel shows the relation in 19 normal ammonia studies, the right panel is for 8 normal FDG studies. The three curves show the mean, the minimum and the maximum defect size obtained for each threshold.

IV. Discussion

The primary end points of the clinical multi-center trial are the relative defect sizes as determined from baseline ammonia and FDG studies. This requires accurate delineation of the left ventricular wall, and accurate definition of the perfusion defect. To optimize the performance of the semi-automatic delineation procedure, it was decided to use maximum-likelihood reconstruction rather than filtered backprojection. No adverse effect of this choice on the kinetic modelling (secondary end-points) is expected. Although many groups still prefer filtered backprojection for kinetic modelling analysis, others report good results with statistical reconstruction as well. Schiepers et al [6, 7] found that with OSEM, the dynamic curves were less noisy, which resulted in a smaller residue after fitting. Boellaard et al [8] studied the effect of OSEM on the kinetic analysis of dynamic FDG-images. They reported a better signal to noise ratio for the tissue curves, and a reduced noise contribution from the transmission scan with OSEM as compared to FBP. The tissue curves from both algorithms produced equivalent quantitative data. However, they also observed positive bias in the later time points of the input curve. This was probably due to zeroing the negative values resulting from precorrection for randoms and scatter. This effect is most pronounced for a low count region (the blood pool at late time points) surrounded by regions of high activity (the myocardial wall). Fortunately, in our study, the quantification of myocardial blood flow was based on the first few minutes of the study, where the bias was negligible.

The threshold of 50% produced accurate defect size estimates for the transmural phantoms on all cameras. This is in good agreement with the SPECT and PET cardiac phantom study of Matsunari et al [9]. In addition, the evaluation in normal volunteers showed that with this threshold, no false positives are generated. However, the phantom experiment also showed that the same threshold fails to detect the subendocardial infarction. Maes et al reported that in about one third of patients, persistent severe perfusion defects were seen after successful thrombolysis [5]. In patients with a flow index of less than 50%, no functional recovery was observed after three months. Since the aim of the trial is to test the hypothesis that the new agent reduces the occurrence of this no-reflow phenomenon, the defect size obtained with a 50% threshold was used as the primary end point. Figure 4 shows that the relation between defect size and threshold is linear and similar for all systems, indicating that the exact value of the threshold is not very critical.

The difference in spatial resolution obtained with 5×8 and with (24×8, 2×4, 4×1) iterations was 1.1 mm. This shows that the effect of incomplete convergence at 5×8 iterations is small compared to that of other resolution degrading factors. When applied to patients, the heart is moving during the dynamic scan, which further degrades the resolution, so it is concluded that 5×8 iterations is sufficient for our purposes. Although more sophisticated methods exist to obtain uniform resolution with statistical reconstruction [10, 11], post-smoothed OSEM is currently probably the only one that can be easily implemented with widely available software.

The spatial resolution (FWHM) in the short axis slices was 9.6 mm with a standard deviation of less than a mm, indicating that by using the same reconstruction algorithm and the same post-smoothing for all cameras, the site dependence of the resolution is nearly eliminated. This may also be partly due to a common strategy used on high resolution cameras, which is to sacrifice some of the resolution to speed up the reconstruction process.

The global spatial resolution can be represented as:

$$F = \sqrt{F_s^2 + F_r^2 + F_p^2 + F_i^2} \qquad (1)$$

where F is the global FWHM in the short axis slice, F_s represents the sinogram resolution, F_r is the effective point spread function of the reconstruction algorithm, F_p is the FWHM of the post-smoothing filter and F_i describes the smoothing effect of the interpolation (reorientation to produce the short axis slices).

The sinogram resolution F_s is position dependent. It allows reconstruction with a resolution of 4-5 mm near the center, and at 20 cm the resolution is about 5-7.5 mm tangentially, 7-9.5 mm radially and 6.5-8 mm axially for the PET systems included in this clinical trial [12–15]. The effective point spread function F_r of the reconstruction algorithm can be estimated by comparing the resolution at 5×8 iterations with that at the high number of iterations, assuming that the latter is effectively converged: $F_r = \sqrt{9.2^2 - 8.1^2}$ = 4.4 mm. The FWHM of the post-smoothing filter is fixed at F_p = 6 mm. The FWHM describing the smoothing effect of the interpolation is in the order of the pixel size, which was around 2.2 mm transaxially and 2.9 mm axially for most systems. Setting F_s to 4.5 mm and F_i to 2 mm the expected spatial resolution is 8.9 mm. With F_s = 7 mm and F_i = 3 mm the spatial resolution increases to 10.6 mm. This is in good agreement with the numbers obtained with the fitting procedure. It also suggests that the position in the camera may be more important than the effects of interpolation, explaining the poor correlation with the pixel size.

V. Conclusion

The same phantom experiment was repeated on 14 scanners in 13 PET centers. Excellent inter-center reproducibility of defect size and global spatial resolution was observed, indicating that myocardial images from different PET-centers can be pooled for statistical analysis.

VI. Acknowledgments

The authors are grateful to dr. Sylvie Jouve of Wyeth Research, the study team leader of the clinical trial, to dr. Mika Teras (Turku PET Centre in Finland) and to Michel Feron (ESAT-PSI, KULeuven), for their help with the file conversion procedures, and to Stefaan Vleugels and Peter Vermaelen (dept. nuclear medicine, KULeuven) for taking care of the logistics of the phantom experiment.

VII. References

[1] MH Hudson, RS Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging* 1994; 13: 601-609.

[2] HH Barrett, JL Denny, RF Wagner, KJ Myers. Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance. *J Opt Soc Am A* 1995; 12: 834-852.

[3] Nuyts J, Suetens P, Oosterlinck A, De Roo M, Mortelmans L. Delineation of ECT images using global constraints and dynamic programming. *IEEE Trans Med Imaging* 1991; 10: 489 - 498

[4] A Maes, W Flameng, J Nuyts, M Borgers, B Shivalkar, J Ausma, G Bormans, C Schiepers, M De Roo, L Mortelmans. Histological alterations in chronically hypoperfused myocardium, correlation with PET findings. *Circulation* 1994; 90: 735-745.

[5] A Maes, F Van de Werf, J Nuyts, G Bormans W Desmet, L Mortelmans. Impaired myocardial tissue perfusion early after successful thrombolysis, impact on myocardial flow, metabolism and function at late follow-up. *Circulation* 1995; 92: 2072-2078.

[6] C Schiepers, J Nuyts, C Wu, R Verma. PET with F-18 fluoride: effects of iterative versus filtered backprojection reconstruction on kinetic modeling. *IEEE Trans Nucl Sci* 1997, 44:1591-1593.

[7] C Schiepers, C Wu, M Seltzer, M Phelps, M Dahlbom, C Hoh, J Nuyts. Factor analysis for delineation of organ

structures and automatic generation of in- and output functions in PET studies of prostate cancer. *Proceedings of the IEEE NSS and MIC*, November 6-11, 2001, San Diego, CA, USA.

[8] R Boellaard, A van Lingen, AA Lammertsma. Experimental and clinical evaluation of iterative reconstruction (OSEM) in dynamic PET: quantitative characteristics and effets on kinetic modeling. *J Nucl Med* 2001; 42: 808-817.

[9] I Matsunari, T Yoneyama, S Kanayama, M Matsudaira, K Nakajima, J Taki, S Nekolla, N Tonami, K Hisada. Phantom studies for estimation of defect size on cardiac ^{18}F SPECT and PET: implications for myocardial viability assessment. *J Nucl Med* 2001; 42: 1579-1585.

[10] JW Stayman, FA Fessler. Regularization for uniform spatial resolution properties in penalized-likelihood image reconstruction. *IEEE Trans Med Imaging* vol 19, pp 601-615, 2000.

[11] S Matej, RM Lewitt. Practical considerations for 3D image reconstruction using spherically symmetric volume elements. *IEEE Trans Med Imaging* 1996; 15: 68-78.

[12] G Brix, J Zaers, LE Adam, ME Bellemann, H Ostertag, H Trojan, U Haberkorn, J Doll, F Oberdorfer, WJ Lorenz. Performance evaluation of a whole-body PET scanner using the NEMA protocol. *J Nucl Med* 1997; 38: 1614-1623.

[13] TR DeGrado, TG Turkington, JJ Williams, CW Stearns, JM Hoffman, RE Coleman. Performance characteristics of a whole-body PET scanner. *J Nucl Med* 1994; 35: 1398-1406.

[14] K Wienhard, L Eriksson, S Grootoonk, M Casey, U Pietrzyk, WD Heiss. Performance evaluation of the positron scanner ECAT EXACT. *J Comput Assist Tomogr* 1992; 16: 804-813.

[15] NA Mullani, KL Gould, RK Hartz, RE Hitchens, WH Wong, D Bristow, S Adler, EA Philippe, B Bendriem, M Sanders. Design and performance of POSICAM 6.5 BGO positron camera. *J Nucl Med* 1990; 31: 610-616.