

Phantom and Clinical Evaluation for New PET/CT Reconstruction Algorithm: Bayesian Penalized Likelihood Reconstruction Algorithm Q.Clear

Estelle Spasic^{1,2*}, Nina Jehanno¹, Sandy Blondeel-Gomes^{1,3}, Virginie Huchet¹, Marie Luporsi¹ and Thibaut Cassou Mounat¹

¹Institute Curie, PSL Research University, Department of Nuclear Medicine, F-75005 Paris, France

²Institute Curie, PSL Research University, Department of Physics, F-75005 Paris, France

³Institute Curie, PSL Research University, Department of Radiopharmacology, F-75005 Paris, France

*Corresponding author: Estelle Spasic, Institute Curie, PSL Research University, Department of Nuclear Medicine, F-75005 Paris, France, E-mail: estelle.spasic@gmail.com

Received date: August 1, 2018; Accepted date: August 20, 2018; Published date: August 27, 2018

Copyright: ©2018 Spasic E, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Introduction: Positron Emission Tomography (PET) technology is on the rise and improvements in image reconstruction are emerging but they involve changes in our habits, which need to be evaluated and clinically validated. One such potential improvement is the new algorithm Q.Clear, developed by GE Healthcare. Q.Clear uses a Bayesian Penalised Likelihood (BPL) reconstruction algorithm. The aim of this study was to investigate, the performance of the (Q.Clear) algorithm relative to the standard reconstruction algorithm (VPFX-S) Time of Flight Ordered Subsets Expectation Maximization (TF-OSEM) with Point-Spread-Function (PSF) in clinical conditions.

Materials and Methods: The first step was to perform a NEMA IEC Body phantom SetTM with EANM guidelines in order to compare performance of Q.Clear reconstruction versus the gold standard in our health care institution (VPFX-S) and to determine an optimum penalization factor for future clinical use of Q.Clear. The second step was to find the best clinical compromise. Four experienced masked readers independently reviewed 4 types of reconstruction performed on 10 patients FDG PET/CT. Then visual and quantitative analysis were performed.

Results: Regarding the quantitation (SUV_{max}), the set of results is a function of the ratio (contrast). Indeed results show that with the highest ratio, a higher β is preferred whereas with the lowest ratio, a lower β is better. However results have to be balanced with better agreement between a high β , our usual clinical reconstruction and caution with GE concerning the overestimation of the SUV_{max} with this new method of reconstruction. Clinical results show a statistically significant difference between reconstructions with a $p < 0.0001$, without significant reader effect $p < 0.7$.

Conclusion: This study suggests a β value of 550 as an optimum penalization factor for clinical use with FDG PET/CT. Studies are still to be done for the optimization of this β for anatomical localization, the weight of the patient and the PET tracer used.

Keywords: Positron emission tomography; Image quality enhancement; Image reconstruction; Q.Clear; TF-OSEM

Abbreviations: BPL : Bayesian Penalised Likelihood; BMI: Body Mass Index; FDG: FluoroDesoxyGlucose; IQ: Image Quality; OSEM: Ordered Subsets Expectation Maximization; PET: Positron Emission Tomography; PSF: Point Spread Function; SNR: Signal to Noise Ratio; SUV: Standardized Uptake Value; SUV_{max} : Maximum Standardized Uptake Value; VOI: Volume of Interest; TF- OSEM: Time Of Flight Ordered Subsets Expectation Maximization

Introduction

Positron Emission Tomography (PET) technology is on the rise and improvements in image reconstruction are emerging but they involve changes in our habits which need to be evaluated and clinically validated. Today, there are two main types of PET reconstruction available: Analytical and iterative. The most commonly used PET

reconstruction algorithm in clinical practice is an iterative method named the Ordered Subsets Expectation Maximization (OSEM) algorithm. Iterative methods improve signal-to-noise ratio (SNR) compared to analytical methods. They offer the advantage of modelling various system factors and also generate the PET images in clinically relevant times with accelerated convergence through the use of subsets [1-4]. OSEM has been improved in the last few years with new models in the reconstruction process: time-of-flight (ToF) and point spread function (PSF) but with still some limits [5,6].

There are still some drawbacks; one of them is that the noise in the image increases with each iteration. The OSEM algorithm is, in general, stopped after 2 (typically used in clinical practice) or 4 iterations [3,4]. A new iterative PET reconstruction algorithm, Bayesian Penalised Likelihood (BPL) has been developed by GE Healthcare, named Q.Clear. This algorithm integrates the time of flight process and is aimed at providing more accuracy in PET quantitation as well as improving SNR over the OSEM. The main advantage is an additional term in the objective function of the OSEM algorithm, which enables it to reach full convergence without increasing the noise,

and with improved accuracy in quantitation. This additional term is controlled by the factor beta (β), which is the only user-input variable in the algorithm [7-11].

The aim of this study was to investigate, through phantom and clinical evaluations, performance in determining disease response and the quantitation of the BPL reconstruction algorithm relative to the standard reconstruction algorithm in clinical conditions. As a result of this study, we propose an optimum penalization factor (β) for future set up of this new algorithm in our health care institution.

Materials and Methods

Phantom

A NEMA IEC Body Phantom SetTM image quality phantom [12] was scanned on a Discovery 710 PET/CT (GE Healthcare) [13]. The 6 spheres were filled once with a Fluorodesoxyglucose (FDG) (theoretical volume of each sphere is 26.5; 11.5; 5.58; 2.57; 1.15 and 0.5 mL). The ratio of activity concentration of the sphere to background activity was 8:1. Then, by filling the background with FDG activity in order to achieve two other experiments with the same conditions of filling the spheres but with two other ratios 6:1 and 4:1, the background activity concentration was modified. The amount of activity and the time were accurately determined, taking into account the radioactive decay for each ratio. The phantom was scanned for each ratio, following the EANM guidelines [14]. Our clinical standard TF-OSEM with PSF (2 iterations, 24 subsets, 6.4 mm gaussian filter) protocol, called VPFX-S by GE, was used as the standard for image reconstruction and compared with the new method of reconstruction Q.Clear using 7 different β value reconstructions between 350 and 650 in increments of 50.

The data was analysed on an Advantage Workstation (AW 4.6; GE Healthcare), using 3 parameters commonly used in clinical analysis: the volume of the spheres, the maximum Standardized Uptake Value (SUV_{max}) and the contrast. These parameters were measured and evaluated from the one experiment, which included 3 results from the 3 ratios between the background and the spheres 8:1, 6:1 and 4:1. The SUV_{max} and volume of each sphere were recorded using a standard volume of interest (VOI) tool according to the Standardized Uptake Value (SUV) body weight model with a "patient" weight equal to 9.82 kg [12] and the threshold set to 42%, which estimated a volume depending on true and observed contrast [15-18]. The results of spheres and background were respectively the value of the VOI of each sphere and the mean of 8 VOIs of same volume (40 cc) in the uniform zone. The contrast between each sphere and the background were estimated according to the equation 1:

$$Contrast = \frac{SUV_{sphere}}{SUV_{background}}$$

Knowing the 3 ratios, these results were compared to the expected values for volume, contrast and SUV_{max} . In addition, for SUV_{max} the standard reconstruction was compared to the 7 Q.Clear reconstructions, in order to estimate the deviation between our usual quantitation and the new one with a bayesian reconstruction method. Figure 1 shows the results of phantom evaluation regarding parameters of reconstruction used for the clinical evaluation and with the nearest ratio to the mean of SUV_{max} lesions in clinical evaluation (8:1).

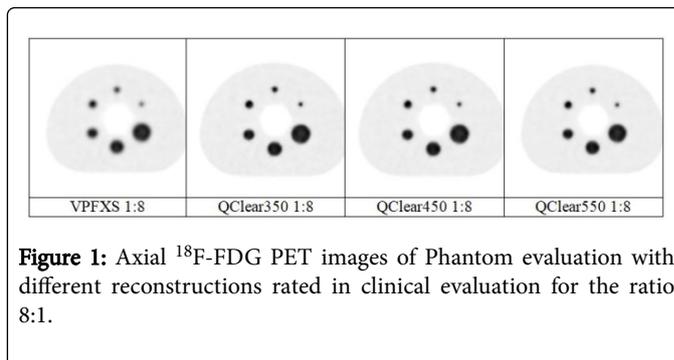


Figure 1: Axial ¹⁸F-FDG PET images of Phantom evaluation with different reconstructions rated in clinical evaluation for the ratio 8:1.

Clinical Evaluation

Based on first results with phantom and manufacturer requirements, four different reconstructions (VPFX-S, Q.Clear 350, 450 and 550) were performed for ten consecutive ¹⁸F-FDG PET/CT scans from patients with an anatomopathologic proof of cancer. Five patients had a body mass index (BMI) \leq 25 and five a BMI > 25. Four experienced nuclear medicine physicians (A, B, C and D) independently reviewed the 4 different reconstructions for each case on an Advantage Workstation (AW 4.6; GE Healthcare), blinded for the reconstruction parameters. Patients and reconstructions were anonymized and labelled from #1 to #10 and from A to D respectively in a randomized order. Six different parameters were reviewed and scored from 1 (best) to 4 (worst) for each reconstruction: overall Image Quality (IQ), background Liver IQ, background mediastinum, background marrow, noise level and lesion detectability [19].

Reconstructions were then ranked according to these 6 parameters, and a final global score was given from the best (1) to the worst (4) image that could be used in clinical conditions for interpretation. The lesion detectability was realized with a blind analysis, in order to evaluate the number of lesion compared to VPFX-S as the reference result. All analyses were done per case per reconstruction and for the final score the 4 reconstructions from the same case were visualised at the same time (Figure 2).

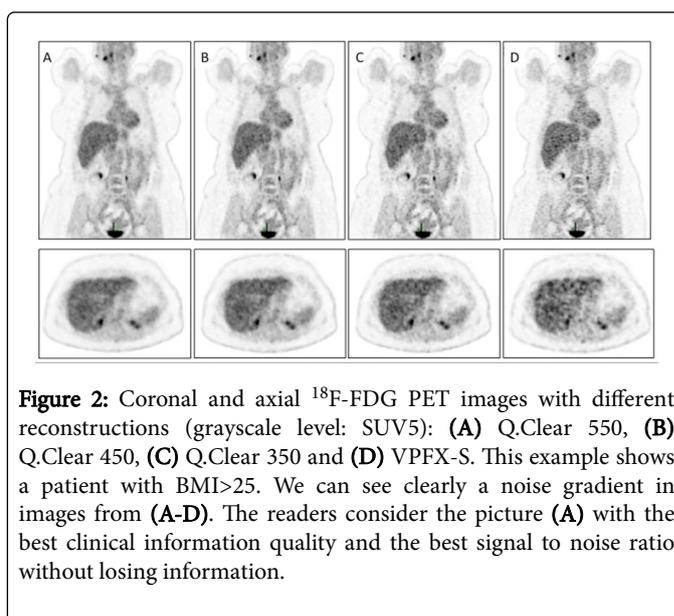


Figure 2: Coronal and axial ¹⁸F-FDG PET images with different reconstructions (grayscale level: SUV5): (A) Q.Clear 550, (B) Q.Clear 450, (C) Q.Clear 350 and (D) VPFX-S. This example shows a patient with BMI > 25. We can see clearly a noise gradient in images from (A-D). The readers consider the picture (A) with the best clinical information quality and the best signal to noise ratio without losing information.

In addition, for each patient, the SUV_{max} of the liver background noise was identified and analyzed according to the group of BMI. Furthermore, for each reconstruction of each patient, the SUV_{max} of 1 lesion was noted and the trend was compared to the phantom results (Figure 3). Statistical analyses were performed using SAS[®] (Version 9.4) software.

A generalized linear model for ordinal data with a cumulative logit link function (PROC GLIMMIX) was chosen to study the preference reconstruction order and the impact of the readers in the classification. For the analysis of background noise, a Mann Whitney test was used.

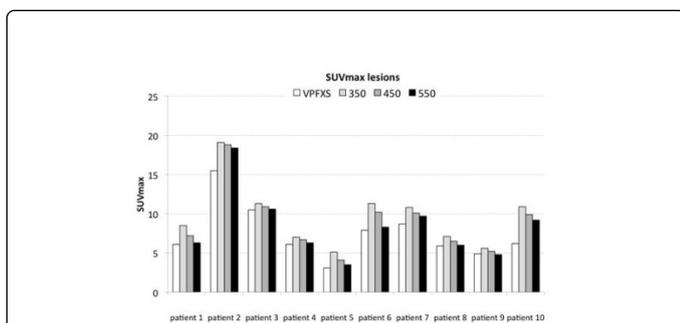


Figure 3: SUV_{max} value of one lesion per patient, showing an overestimation with Q.Clear 350 as the phantom evaluation. Within the three reconstructions, Q.clear 550 has the nearest results from VPFX-S.

Results

The results for the phantom study are summarized in Figures 1-4 and Table 1. They are represented as the mean of all spheres for each parameter. The mean contrast deviation and metabolic volume compared to theoretical values show better results with Q.Clear reconstructions whatever the ratio and the β coefficient. Regarding the metabolic volume Q.Clear reconstructions have results on average 30% better than VPFX-S and for the contrast it is on average 5% to 10% better.

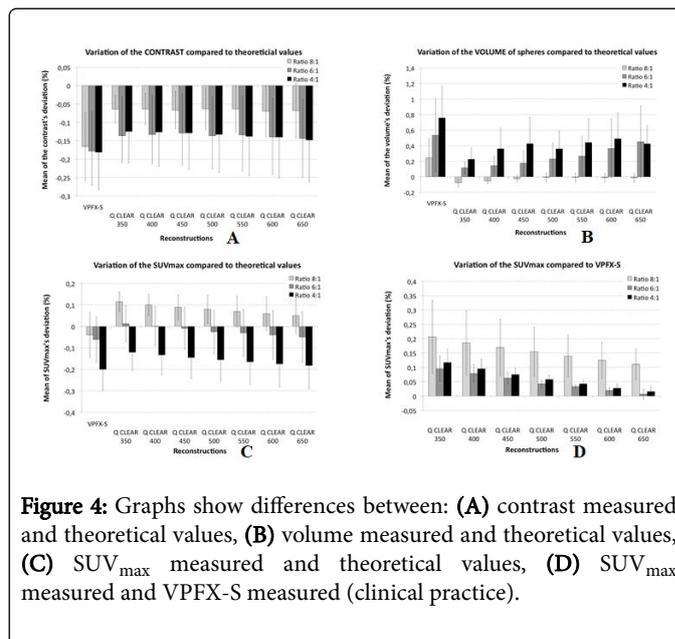


Figure 4: Graphs show differences between: (A) contrast measured and theoretical values, (B) volume measured and theoretical values, (C) SUV_{max} measured and theoretical values, (D) SUV_{max} measured and VPFX-S measured (clinical practice).

		SUVmax deviation vs. Exected value		SUVmax deviation vs. VPFX-S		VOLUME of SPHERES		CONTRAST	
		Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
RATIO 8:1	VPFX-S	-3.90%	21.10%			24.60%	49.20%	-16.60%	18.30%
	Q CLEAR 350	11.50%	9.00%	20.60%	25.30%	-7.70%	9.90%	-6.40%	7.60%
	Q CLEAR 400	10.00%	10.00%	18.60%	22.40%	-5.30%	7.20%	-6.40%	8.50%
	Q CLEAR 450	8.90%	11.70%	16.90%	19.60%	-2.60%	8.30%	-6.70%	10.00%
	Q CLEAR 500	8.00%	13.10%	15.40%	17.00%	-0.70%	10.90%	-6.30%	11.40%
	Q CLEAR 550	6.90%	14.50%	13.90%	14.60%	-0.90%	11.10%	-6.30%	12.70%
	Q CLEAR 600	5.90%	15.70%	12.50%	12.50%	-1.20%	11.30%	-7.00%	13.80%
	Q CLEAR 650	5.00%	16.80%	11.10%	10.50%	-1.30%	11.50%	-6.80%	14.90%
RATIO 6:1	VPFX-S	-6.10%	21.20%			53.40%	93.10%	-17.80%	18.60%
	Q CLEAR 350	1.30%	16.90%	9.50%	8.90%	11.60%	15.70%	-13.60%	14.40%
	Q CLEAR 400	10.00%	18.40%	7.80%	6.30%	14.40%	23.90%	-13.30%	15.90%
	Q CLEAR 450	-0.90%	19.70%	6.20%	3.90%	17.50%	31.80%	-12.90%	17.40%
	Q CLEAR 500	-2.50%	20.40%	4.20%	2.70%	23.00%	41.50%	-13.60%	18.10%
	Q CLEAR 550	-3.00%	21.90%	3.30%	1.20%	26.50%	51.80%	-13.40%	19.50%

	Q CLEAR 600	-4.00%	22.80%	1.90%	2.20%	36.50%	74.90%	-14.00%	20.40%
	Q CLEAR 650	-4.90%	23.60%	0.60%	3.50%	45.10%	92.30%	-14.40%	21.20%
RATIO 4:1	VFPX-S	-20.00%	19.90%			75.80%	82.20%	-18.10%	20.30%
	Q CLEAR 350	-12.00%	17.20%	11.70%	9.40%	22.60%	29.50%	-12.50%	17.10%
	Q CLEAR 400	-13.30%	18.40%	9.50%	6.70%	36.00%	53.90%	-12.60%	18.50%
	Q CLEAR 450	-14.50%	19.40%	7.50%	4.60%	42.60%	67.40%	-12.90%	19.80%
	Q CLEAR 500	-15.50%	20.20%	5.80%	3.20%	36.00%	46.00%	-13.30%	20.70%
	Q CLEAR 550	-16.50%	20.80%	4.20%	2.60%	44.10%	60.10%	-13.80%	21.50%
	Q CLEAR 600	-17.50%	21.40%	2.70%	2.90%	48.90%	67.20%	-14.00%	22.30%
	Q CLEAR 650	-18.20%	21.90%	1.60%	3.50%	42.40%	47.30%	-14.80%	22.80%

Table 1: Results of phantom evaluation: mean and standard deviation of contrast, volume and SUV_{max} from the 3 ratios. Quantitation (SUV_{max}) results compared to theoretical values are influenced by the ratios. Theoretical values were calculated based on activity and concentration injected during each step of the phantom evaluation. ($SUV_{max, theoretical} = 8.02, 5.94$ and 3.93 respectively for the ratio 8:1, 6:1 and 4:1). Indeed, VFPX-S shows the best results with a ratio 8:1 and this begins to break down with decrease of the contrast. The same results were observed using Q.Clear. For the 8:1 ratio, the best results were obtained with a high β whereas the trend changes for lower contrast to better results with a low β .

Now, if we look for a comparison between Q.Clear results and the clinical standard VFPX-S, it is important to note that Q.Clear results are closer to the standard with a high β (Figure 4D). In the clinical

evaluation, results can be summarized as the final score as shown in Table 2.

		VFPX-S					Q.Clear 350					Q.Clear 450					Q.Clear 550				
		#1	#2	#3	#4	#5	#1	#2	#3	#4	#5	#1	#2	#3	#4	#5	#1	#2	#3	#4	#5
BMI ≤ 25	A	3	2	3	2	3	4	4	4	4	4	2	3	2	3	2	1	1	1	1	1
	B	1	1	1	1	1	4	4	4	4	4	3	3	3	3	3	2	2	2	2	2
	C	4	4	4	4	3	3	3	3	3	4	2	2	2	2	1	1	1	1	1	1
	D	3	3	2	3	3	4	4	4	4	4	2	1	3	2	2	1	1	1	1	1
		#6	#7	#8	#9	#10	#6	#7	#8	#9	#10	#6	#7	#8	#9	#10	#6	#7	#8	#9	#10
BMI > 25	A	4	2	2	4	3	3	4	4	3	4	2	3	3	2	1	1	1	1	1	1
	B	3	2	2	3	2	4	4	4	4	4	2	3	3	2	3	1	1	1	1	1
	C	4	4	4	4	4	3	3	3	3	3	1	2	1	2	2	2	1	2	1	1
	D	3	4	3	3	4	2	2	4	4	3	4	3	1	2	2	1	1	2	1	1

Table 2: Final scores (1: excellent; 4: worst) from clinical evaluation for the 10 patients (from #1 to #10) presented in function of the body mass index (BMI) analysed by four readers (A, B, C and D)

Clinically the results show a statistically significant difference between reconstructions with a $p < 0.0001$, without significant “reader effect” $p < 0.7$. Statistically the probability of reconstruction classification in order of preference is Q.Clear 550 first, Q.Clear 450 second, VFPX-s in third and Q.Clear 350 last.

For one of the readers there is an apparent disagreement between VFPX-S and Q.Clear 550. Scorer “B” prefer VFPX-S for patients with $BMI \leq 25$ compared to $BMI > 25$ with the best results linked to Q.Clear 550. In contrast, the others scorer always chose Q.Clear 550 whatever the BMI. But statistically there is no significant reader effect.

Moreover, there is a statistically significant difference between SUV_{max} hepatic background noise of the groups according to the BMI $p < 0.00012$. Figure 3 shows the tendency of the SUV_{max} according to the reconstructions and we see that it is the same as that obtained with Phantoms (Figure 4). SUV_{max} values are overestimated with Q.Clear 350 and the values of Q.Clear 550 are closer to VFPXS. In addition, the average of the SUV_{max} lesions was 7.6, which appears closer than the 1:8 phantom evaluation.

Discussion

The aim of this study was to investigate the performance of the new reconstruction algorithm (Q.Clear) regarding the influence on the penalization factor β compared to the standard reconstruction algorithm VPFX-S used in clinical practice.

The study started with the phantom evaluation in order to evaluate a wide range of factors β . Because of the difficulty in accurately filling all spheres, 3 scans were performed in 1 experiment. In this way, uncertainty in volume was avoided. The presentation of our results as the mean of all sphere's size enables us to summarise all the data and at the same time to show an overall result regardless of the size of the spheres. We wanted the phantom study to be closer to a clinical evaluation. It can be seen that contrast, metabolic volume and quantitation are all clinical parameters and the use of 3 scans with 3 ratios aims to be representative of clinical situations.

As a transition to the clinical evaluation, when we are looking for the SUV_{max} results compared to our clinical standard (Figure 4D and Figure 3), it appeared that Q.Clear with high β factor presents closer SUV_{max} results to VPFX-S than low β factor, whereas with high β the reconstructed images were visually less close than the standard method. Furthermore, the tendency of the SUV_{max} according to the reconstructions is the same between clinical and phantom analysis. SUV_{max} values are always overestimated with Q.Clear 350 and the values of Q.Clear 550 are closer to VPFX-S.

Regarding the main studies [7,10], our study was performed with the same phantom but not the same procedure of filling spheres and background ratio. We filled all spheres and performed 3 ratios in order to be closer to the clinical evaluation. This choice enabled us the performance of phantom acquisition with ratio closer to the clinical ratio as we can see on the SUV_{max} clinical evaluation where the mean value is around 8.

Indeed, the 4 masked readers preferred significantly the reconstruction with the highest β factor (Q.Clear 550). This result is the main discrepancy with GE recommendations [11] and the studies mentioned above which respectively suggest using 350 and 400 as β factor. This point can be balanced regarding conditions of evaluation that are not the same. Their studies are based on 4:1 ratio (usually used in NEMA's and EANM's procedures) [14] compared to our study using 3 ratios 4:1, 6:1 and 8:1. Results change according to the ratio used, meaning relative to the contrast.

On the other hand, this is the first study to our knowledge that shows the influence of BMI in β factor selection according to the contrast, which impacts overweight patients. Results clearly showed an agreement between the four independent readers for the highest β factor (Q.Clear 550) for patients with a BMI > 25.

As said above, this is understandable because the contrast is altered for the overweight patients. In overweight patients the noise is increased because fewer events are collected and noise is amplified by the increased attenuation correction factors. For proof our statistically significant results between the SUV_{max} liver background noise of the groups according to the BMI ($p < 0.00012$).

The use of a β 550 penalization factor leads to smooth images, removing noise from images and actually improving the quality without losing data. In patients with BMI ≤ 25 , Q.Clear 550 remains the best choice for three of the readers, whereas for one of the reader it is VPFX-S.

We explain this exception by the influence of usual clinical practice and the reader's experience with a variety of images. Although not significant, this result is interesting because it suggests that an adaptation of the acquisition and reconstruction techniques must be customized to the BMI. Thus, we have to consider a compromise between theoretical and clinical results. Using Q.Clear gives better results than VPFX-S but both are impacted by the contrast. In clinical practice, the contrast varies according to both anatomical location and weight of the patient (BMI) and using a high Beta Q.Clear seems to minimize the defaults generated by the noise in the image. So we still recommend using Q.Clear 550, even if it requires an adaptation period for some doctors.

Our second suggestion would be to choose different β factors in the same reconstruction according to the contrast in the image and the anatomical localisation, although such reconstructions (step-dependant β factors) are difficult to implement and time-consuming.

On the other hand, we need to carefully choose the β factor in order to obtain acceptable quantitation results for clinical interpretation. In medical imaging, the objective is to obtain a compromise between signal to noise and contrast, without degrading the images. GE warns against overestimating the SUV value with this new algorithm reconstruction [20], which means new adaptations for clinical interpretation.

From those results, we should consider gradually setting up of this new Q.Clear algorithm in our health care institution, choosing a β factor of 550 for FDG PET/CT examinations.

Conclusion

It appears that the Q.Clear algorithm improves image quality and facilitates the interpretation of PET imaging without decreasing lesion detectability compared to our standard reconstruction (VPFX-S). Regarding the influence of the penalization factor (β) on contrast and the overestimation of the SUV_{max} with low β values, we suggest β 550 as an optimum penalization factor for a start for clinical use with FDG PET/CT. More studies are necessary to determine optimal β value for use with Q.Clear according to the anatomical regions and other PET radionuclides.

Ethics Approval

Not applicable

Consent for Publication

All patient gave their written consent to our Health Care Institute for use data including anonymized pictures.

Availability of Data and Materials

The datasets supporting the conclusions of this article are included within the article.

Acknowledgements

We thank the biostatistics department from the Curie Institute and Matthieu Carton for his assistance in the statistical analysis.

References

1. Tong S, Alessio AM, Kinahan PE (2010) Image reconstruction for PET/CT scanners: Past achievements and future challenges. *Imaging Med* 2: 259-545.
2. Alessio AM, Stearns CW, Tong S (2010) Application and evaluation of a measured spatially variant system model for PET image reconstruction. *IEEE Trans Med Imaging* 29: 938-949.
3. Hudson HM, Larkin RS (1994) Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging* 13: 601-609.
4. Nuyts J, Beque D, Dupont P, Mortelmans L (2002) A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography. *IEEE Transactions on Nuclear Science* 49: 56-60.
5. Armstrong IS, Kelly MD, Williams HA, Matthews JC (2014) Impact of point spread function modelling and time of flight on FDG uptake measurements in lung lesions using alternative filtering strategies. *EJNMMI Phys* 1: 99.
6. Prieto E, Domínguez P, García VMJ, Peñuelas I, Richter JÁ, et al. (2013) Impact of time-of-flight and point-spread-function in SUV quantification for oncological PET. *Clin Nucl Med* 38: 103-109.
7. Teoh EJ, McGowan DR, Belcher E, Black E, Gleeson FV (2016) Novel penalised likelihood reconstruction of PET in the assessment of histologically verified small pulmonary nodules. *Eur Radiol* 26: 576-584.
8. Chism CB, Ravizzini GC, Macapinlac HA, Pan T (2017) Quantitative comparison between regularized time-of-flight and OSEM PET reconstructions for small 18F-FDG-avid lesions. *Nucl Med Commun* 38: 529-536.
9. Howard BA, Morgan R, Thorpe MP, Turkington TG, Oldan J, et al. (2017) Comparison of Bayesian penalized likelihood reconstruction versus OSEM for characterization of small pulmonary nodules in oncologic PET/CT. *Ann Nucl Med* 31: 623-628.
10. Ahn S, Ross SG, Asma E, Miao J, Jin X, et al. (2015) Quantitative comparison of OSEM and penalized likelihood image reconstruction using relative difference penalties for clinical PET. *Phys Med Biol* 60: 5733-5751.
11. Ross S (2013) Q.Clear (GE Healthcare White Paper). Accessed 16.
12. NEMA NU-2-2012 Performance measurements of positron emission tomography. (2013) NEMA; National Electrical Manufacturers Association.
13. Andersen T, Hoiland CPF (2016) The Q.Clear PET reconstruction algorithm: Evaluation using the NEMA IQ phantom. *J Nucl Med* 57 : 1
14. Boellard R, O'Doherty MJ (2010) FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging* 37: 181-200.
15. Burger IA, Vargas HA, Apte A, Beattie BJ, Humm JL, et al. (2014) PET quantification with an histogram derived total activity metric: superior quantitative consistency compared to total lesion glycolysis with absolute or relative SUV thresholds in phantoms and lung cancer patients. *Nucl Med Biol* 41: 410-418.
16. Seol YM, Kwon BR, Song MK, Choi YJ, Shin HJ, et al. (2010) Measurement of tumor volume by PET to evaluate prognosis in patients with head and neck cancer treated by chemo-radiation therapy. *Acta Oncol* 49: 201-208.
17. Chung HH, Kim JW, Han KH, Eo JS, Kang KW, et al. (2011) Prognostic value of metabolic tumor volume measured by FDG-PET/CT in patients with cervical cancer. *Gynecol Oncol* 120: 270-274.
18. Chung HH, Kwon HW, Kang KW, Park NH, Song YS, et al. (2012) Prognostic value of preoperative metabolic tumor volume and total lesion glycolysis in patients with epithelial ovarian cancer. *Ann Surg Oncol* 19: 1966-1972.
19. Teoh EJ, McGowan DR, Macpherson RE, Bradley KM, Gleeson FV (2015) Phantom and Clinical Evaluation of the Bayesian Penalized Likelihood Reconstruction Algorithm Q.Clear on an LYSO PET/CT system. *J Nucl Med* 56: 1447-1452.
20. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33: 159-174.