Biomedical Physics & Engineering Express

PAPER

CrossMark

NEMA NU 2-2018 evaluation and image quality optimization of a new generation digital 32-cm axial field-of-view Omni Legend PET-CT using a genetic evolutionary algorithm

Rhodri Lyn Smith[1] , Lee Bartley[2], Christopher O'Callaghan[2], Luiza Haberska[1] and Chris Marshall[1]

1 The Wales Research and Diagnostic Positron Emission Tomography Imaging Centre, Cardiff University, School of Medicine, Cardiff, CF14 4XN, Wales, United Kingdom
2 Radiology, Cardiff and Vale University Health Board, University Hospital of Wales, Cardiff, CF14 4XW, Wales, United Kingdom

E-mail: SmithR50@cardiff.ac.uk

Abstract

A performance evaluation was conducted on the new General Electric (GE) digital Omni Legend PET-CT system with 32 cm extended field of view. The first commercially available clinical digital bismuth germanate system. The system does not use time of flight (ToF). Testing was performed in accordance with the NEMA NU2–2018 standard. A comparison was made between two other commercial GE scanners with extended fields of view; the Discovery MI − 6 ring (ToF enabled) and the Discovery IQ (non-ToF). A genetic evolutionary algorithm was developed to optimize image reconstruction parameters from image quality assessments. The Omni demonstrated average spatial resolutions at 1 cm radial offset as 3.9 mm FWHM. The total system sensitivity at the center was 44.36 cps/kBq. The peak NECR was measured as 501 kcps at 17.8 kBq ml^{-1} with a 35.48% scatter fraction. The maximum count-rate error below NECR peak was 5.5%. Using standard iterative reconstructions, sphere contrast recovery coefficients were from $52.7 \pm 3.2\%$ (10 mm) to $92.5 \pm 2.4\%$ (37 mm). The PET-CT co-registration accuracy was 2.4 mm. In place of ToF, the Omni employs software corrections through a pre-trained neural network (PDL) (trained on non-ToF to ToF) that takes Bayesian penalized likelihood reconstruction (Q.Clear) images as input. The optimum parameters for image reconstruction, determined using the genetic algorithm were a Q.Clear parameter, β, of 350 and a 'medium' PDL setting. Using standard iterative reconstructions, the Omni initially showed increased background variability compared to the Discovery MI. With optimized PDL reconstruction parameters selected using the genetic algorithm the performance of the Omni surpassed that of the Discovery MI on all NEMA tests. The genetic algorithm's demonstrated ability to enhance image quality in PET-CT imaging underscores the importance of algorithm driven optimization and underscores the requirement to validate its use in the clinical setting.

1. Introduction

Hybrid Positron Emission Tomography - Computed Tomography (PET-CT) scanners were first conceived in the early 1990s [1], with commercial systems first introduced in early 2001. The GE Discovery LS PET system consisting of 18 rings, each containing 672 bismuth germanate (digital BGO) scintillating crystals (4 mm × 8 mm × 30 mm crystal size) and an axial field of view (AFOV) of 152 mm was the first commercially available PET-CT system. Evolving

technology in the form of both hardware and software [2] has resulted in increasing image quality, improved quantitative accuracy and earlier disease detection from PET images. This has undoubtedly resulted in patient benefit. The purpose of this paper is to evaluate the physical performance of the new Omni Legend PET-CT system, the first commercially available clinical scanner with dBGO crystals integrated with silicon photomultipliers (SiPM). National Electric Manufacturer's Association (NEMA) performance measurements are rigorous tests performed to ensure

that imaging systems are fully operational and perform according to specification. Measurements of performance are assessed before system acceptance and serve as a reference for future tests to ensure that the overall PET performance has not degraded over time. The standard NEMA NU 2–2018 guidelines [3] for PET includes a series of tests for spatial resolution, image quality, scatter fraction, count rate performance, accuracy of correction for count losses and random events, and sensitivity. We report the findings of the NEMA tests and make a comparison to other commercial GE PET-CT system. We further explore into the impact of different reconstruction methods on image quantification, specifically assessing recovery coefficients, background variability, and line profiles using the NEMA image quality phantom. Subsequently, the paper extends to optimizing image reconstruction hyperparameters with a distinct objective in mind: identifying the ideal parameter configuration to enhance the quality evaluations of NEMA images. We navigate the intricate dynamics and interplay of hyperparameters within advanced artificial intelligence-driven reconstructions, introducing an innovative approach through the implementation of a genetic algorithm. The genetic algorithm's use marks a significant departure from conventional methods, as it excels in scenarios where the cost function lacks a readily derivable form. This is particularly beneficial when the optimization objective involves minimizing background variability and maximizing contrast recovery, as is the case with our study. Genetic algorithms offer distinct advantages over heuristic methods in such contexts. They are inherently equipped to explore complex, high-dimensional problem spaces characterized by non-linear and intricate features. Unlike heuristic approaches, genetic algorithms excel at navigating solution spaces with multiple local optima, allowing for a more comprehensive search for the global optimum. Moreover, they hold the potential to yield more optimal solutions. While genetic algorithms have previously demonstrated their efficacy in a diverse array of optimization tasks, including image reconstruction for PET images [4] and various image processing endeavors [5], it's pertinent to underline that, to the best of the authors' knowledge, this marks the first instance of genetic algorithms being employed for the optimization of image reconstruction hyperparameters. This underscores their unique capabilities and potential to address optimization challenges in novel and impactful ways.

2. Methods

2.1. PET-CT system
The Omni Legend PET-CT consits of the Revolution Maxima integrated CT component (80–140 kV) with 64 slices (detector element thickness of 0.625 mm).

The PET component has 32 cm axial FOV with Light-Burst BGO crystals (4.1 mm $\times$ 4.1 mm $\times$ 30 mm) enclosed in 72 detector rings, with a total of 38016 crystals in 528 blocks, backed to 9504 silicon photomultiplier channels (SiPM). The digital BGO detector has the advantages of high density and stopping power, resulting in improved sensitivity. Time of Flight (ToF) capabilities are not provided by the Omni Legend; an image derived ToF correction, from non-ToF images is provided in the form of a pre-trained convolutional neural network, to enhance non-ToF images to their ToF equivalent [6]. This is termed 'precision Deep Learning (PDL)'. The Omni Legend's digital BGO detector provides up-gradeability of the FOV and is designed to support future extended axial FOV upgrades. The Omni Legend is also commercially available with a 16 cm axial FOV. The system provides advanced quantitative reconstructions in the form of VUE Point HD (VPHD—3D ordered subset expectation maximization—OSEM [7]) and Q.Clear (BSREM - Bayesian penalized-likelihood reconstruction [8]). The PDL algorithm takes as input the Q.Clear reconstructed images.

2.2. Measurements
Following the NEMA NU 2–2018 standard, the physical PET performance of the Omni Legend was assessed. The tests included spatial resolution, sensitivity, scatter fraction, count-rate performance, accuracy of count losses, random corrections and image quality. In benchmarking against GE acceptance testing specifications, we regarded these as the minimal performance thresholds, with the expectation that our measured results would often surpass these baselines. The impact of image reconstruction on image quality metrics is also presented. Prior to NEMA acceptance testing it was ensured the dose calibrator used was calibrated to a secondary standard and the clock from the calibrator synchronised to the image workstation.

2.3. Spatial resolution
The spatial resolution test assesses the full width half maximum (FWHM) in air of a reconstructed ^{18}F point source to assess the point spread function (psf). Three-point sources were prepared using narrow capillary tubes, with an inside diameter of less than 1 mm and an outside diameter of less than 2 mm. These tubes were filled with ^{18}F at a high activity concentration > 20 MBq ml^{-1} ensuring the filled length was kept under 1 mm. The tubes were securely sealed and positioned using a specialized holder, aligning them within the scanner's field of view at three different locations: 1 cm, 10 cm, and 20 cm vertically from the center. The spatial resolution was measured in two transaxial planes: one was at the center, the other was at the 1/8 position from the edge, in the axial field of view (AFOV). An initial position check scan was used and position of the source holder adjusted until the x, y and

z positional error was <1.0 mm. Image acquisition time was selected to acquire an image of >500,000 prompt counts. The images were reconstructed using the filtered back projection algorithm (FBP) and the VPHD algorithm (matrix size 384 × 384 with 6 iterations, 22 subsets, 2 mm cut off Gaussian filter) with voxel size of $0.65 \times 0.65 \times 2.07$ mm with a field of view of 250×250 mm The transverse spatial resolution is the average of the radial and tangential values. Profiles across the point source response functions in all three directions (radial, tangential and axial) were generated using manufacturer provided software. A parabolic fit, with linear interpolation, is fitted to the profiles; and the FWHM calculated using the voxel size. The radial and tangential resolutions were averaged along the axial positions to give the transverse resolution.

2.4. System sensitivity

The sensitivity determines the count rate (true coincidences) as measured by the imaging system versus the activity within the FOV. A standardized source configuration was employed. This comprises of a line source surrounded by concentric aluminum sleeves. This setup included five sleeves with progressively increasing diameters, each sleeve was 1.25 mm thick. The co-axial alignment of these sleeves created a controlled attenuation path. Whilst these sleeves facillitate photon annihilation, they also partially absorb emitted photons, complicating the direct measurment of absolute sensitivity. Therefore, measurements of true coincidences are extrapolated to zero absorption to enable an estimate of sensitivity in the absence of attenuation. For this sensitivity phantom, a polyethylene tube was placed inside the aluminum sleeves with a a central 70 cm section filled with ^{18}F to serve as a line source. The source activity was chosen to maintain dead time losses below 5% while ensuring at least 10,000 true counts per slice were acquired For the Omni scanner, the activity at imaging time was kept below 4 MBq. We positioned the phantom in air, supported at each end by low-density materials: a phantom holder and a custom-designed plastic hook. This arrangement minimized scatter in the FOV. The phantom was carefully aligned along the scanner's z-axis. Five, one-minute acquisitions were performed, starting with the smallest sleeve and gradually adding larger sleeves to increase the wall thickness. To account for random coincidence, events were subtracted from the prompts by using a delayed coincidence window. Using the manufacturers software the count rate for each acquisiton is corrected for decay to the imaging start time. The corrected count rates for each sleeve is fitted to a linear function to determine the unattenuated count rate ($R_{corr\ O}$) in the absence of any sleeves. The total system senstivity (S_{tot}) is calculated using equation (1) where A_{cal}, represents the activity at time of imaging.

$$S_{tot} = R_{corr\ o}/A_{cal} \qquad (1)$$

We repeateded this procedure with the phantom and line source offset by 10 cm from the central axis. This provided a sensitivity profile that varied with slice number along the axial offset.

2.5. Scatter fraction, count losses, and randoms

The scatter fraction, or the ratio of scattered to total coincidences, indicates the scanner's response to scattering, with random events being minimal at low count rates. Count rate performance, which evaluates dead time losses and random event rates, varies with radioactivity. The Noise Equivalent Count Rate (NECR) measures PET scanner performance by integrating true, scattered, and random events detection. NECR values guide optimal radioactivity levels for patient administration in clinical settings. The experiment utilized a cylindrical polyethylene scatter phantom (70 cm × 20 cm) with a 45 mm off-axis bore, holding a line source (165 MBq ml^{-1}, 5.15 ml). This assembly was centrally placed on the patient bed, positioning the line source close to the bed. Image frames were acquired as a decay series to sample the NECR curve, each acquisition achieving at least 500,000 prompt counts and continuing until true event losses were less than 1%. The entire PET scan protocol comprised 24 timing frames over approximately 12 h. PET images were reconstructed with VPHD (22 subsets, 3 iterations, 5 mm cutoff Gaussian filter).

For each acquisition (j) with duration $T_{acq,j}$ we obtained prompt and random sinograms for each slice I, denoted as ($C_{tot,i,j}$) and ($C_{r,i,j}$) respectively. Scattered plus random event ($C_{r+s,i,j}$) are estimated from profiles drawn across the prompt sinogram, as per NEMA section 4. This enables the calculation of the true (equation (2)), random (equation (3)), and scatter (equation (4)) event rates for each slice and acquisition. The scatter fraction (SF) for each slice and acquisition is calculated by averaging (equation (5)).

$$\textit{True\quad Event\quad Rate}: R_{t,i,j} = \frac{C_{tot,i,j} - C_{r+s,i,j}}{T_{acq,j}} \qquad (2)$$

$$\textit{Random\quad Event\quad Rate}: R_{r,i,j} = \frac{C_{r,i,j}}{T_{acq,j}} \qquad (3)$$

$$\textit{Scatter\quad Event\quad Rate}: R_{s,i,j} = \frac{C_{r+s,i,j} - C_{r,i,j}}{T_{acq,j}} \qquad (4)$$

$$\textit{Scatter\quad Fraction}: SF = \frac{C_{r+s} - C_r}{C_{tot} - C_r} \qquad (5)$$

The analysis was performed using GE provided software. Count rate curves are calculated for total, true, random and scatter events; the NECR for each slice i is determined by equation (6), where $R_{Tot,i,j} = C_{Tot,i,j}/T_{acq,j}$ is the total event rate. The system NECR is the sum over all slices.

$$NECR: R_{NEC,i,j} = \frac{R_{t,i,i}^2}{R_{Tot,i,j} + R_{r,i,j}} \qquad (6)$$

2.6. Accuracy of count losses and random corrections

This test measures the accuracy of count losses and random corrections in accordance with NEMA NU2–2018 section 6 This test uses the data acquired in section 1.3, scatter fraction, count losses and randoms. At low count rates it is assumed that the count loss and random events are negligible. Least squares fit of true rates at low count rates versus effective activity concentration is determined. This is extrapolated to higher count rates (below the peak NECR) to determine the effectiveness of corrections. The mean, minimum and maximum deviations from the extrapolated fit are observed and the maximum error below the peak NECR reported.

2.7. Image quality and accuracy of attenuation, and scatter correction and quantification

NEMA image quality tests use an IQ phantom and line source in a scatter medium to replicate total body imaging, with external radioactivity simulated by scatter medium activity. The IQ phantom contains spheres of varying diameters (10 mm to 37 mm) and a 5 cm diameter cylindrical insert filled with Styrofoam to simulate lung attenuation. Image quality is quantified by calculating image contrast and background variability ratios for the spheres and assessing the accuracy of attenuation and scatter corrections, with the test repeated three times for quantification of variability. The IQ phantom background (9729 ml volume) was filled with 5.3 kBq ml^{-1} of ^{18}F and spheres ($\sim$60 ml^{-1} total volume) with 21 kBq ml^{-1} creating a 4:1 concentration ratio . The scatter phantom contained 120MBq of ^{18}F. The IQ phantom was placed and centred on the couch with the spheres facing away from the gantry with the scatter phantom positioned distally but outside the PET FOV.

A PET-CT acquisition lasting initially 7 min and 5 s, simulating a 100 cm, 30 min whole body scan was performed and repeated twice whilst accounting for decay compensation. A CT scans (120 kVp tube voltage, 115 mAs exposure, 0.95 mm pitch) was used for attenuation correction. Images were reconstructed with VPHD into a 384×384 matrix, with 6 iterations, 22 subsets and 2 mm gaussian filter cut-off. For each sphere 's', a circular region-of-interest (ROI) was drawn on the central slice, with average ROI counts in each sphere defined as C_s Background variability (BV) and contrast recovery coefficients (CRC) (equation (7)) were computed for each sphere size with 60 ROIs drawn in the background over 5 slices. Automated ROI positioning was performed with GE software. The average background ROI counts C_{Bs} per sphere were also measured. The concentration ratio is defined as the ratio between the spheres and background and BV is the coefficient of variation for background spheres of each size.

$$Contrast \quad recovery \quad coeffs: CRCs$$
$$= \frac{(C_s / C_{Bs}) - 1}{conc.ratio - 1} \qquad (7)$$

Additionally, 12 ROIs (C_{BL}) (37 mm diameters) and a ROI in the centre of the lung insert (30 mm diameters) (C_{Li}) for each slice (i) were defined in the axial range (for 5 slices). This allowed evaluation of the accuracy of corrections by the ratio of C_{Li}/C_{BL}. The residual lung error (RLE) is the average of this ratio over all slices. Furthermore, the CRC and BV were assessed using the Bayesian penalized likelihood reconstruction Q.Clear and the integrated PDL algorithm . The extent of penalization when using the Q.Clear algorithm is controlled by a single parameter, β. The PDL algorithm can provide three levels of contrast-enhancement to noise trade off, 'low',' medium' and 'high.'

2.8. Optimization of image reconstruction parameters

Our experimentation focused on determining the optimal hyperparameters of the Q.Clear image reconstruction method, specifically the β parameter and the PDL setting. Previous studies have not provided a consistent, universally optimum β value for NEMA phantom Image quality assessments across all sphere sizes [9]. Generally, lower β values favour small structure detectability and quantification, while higher β values are preferable for larger structures. However, an inherent trade-off arises when optimizing for both image contrast and background variability, as lowering the β value increases both the contrast recovery coefficients and the background variability [9]. To address this gap in research, we reconstructed NEMA IQ images with a range of Q.Clear values ($\beta = 100$, 200, 300, 350, 400, 500, 600, 700, 800 and 900) and processed them using various PDL settings, including 'low,' 'medium,' and 'high.' We employed the novel use of a genetic algorithm to identify the 'optimal' hyperparameters for Q.Clear and PDL. The genetic algorithm was configured with a population size of 100, a mutation rate and crossover rate of 0.5, and was run for 50 generations. Optimization was halted upon reaching a minimum of the cost function. Throughout our study, three cost functions (equation (8) i-iii) were used for each individual sphere size and across all sphere sizes, employing normalized values of CRC (Contrast Recovery Coefficient) and BV (Background Variability). To prevent bias from absolute values, we applied Max-Min normalization across all reconstruction settings.

$$(i) - 1 \times CRC_{normalized} \qquad (8)$$

$$(ii) 1 \times BV_{normalized}$$

$$(iii) - 0.5 \times CRC_{normalized} + 0.5 \times BV_{normalized}$$

Figure 1. Axial sensitivity profile at the 10cm offset profile.
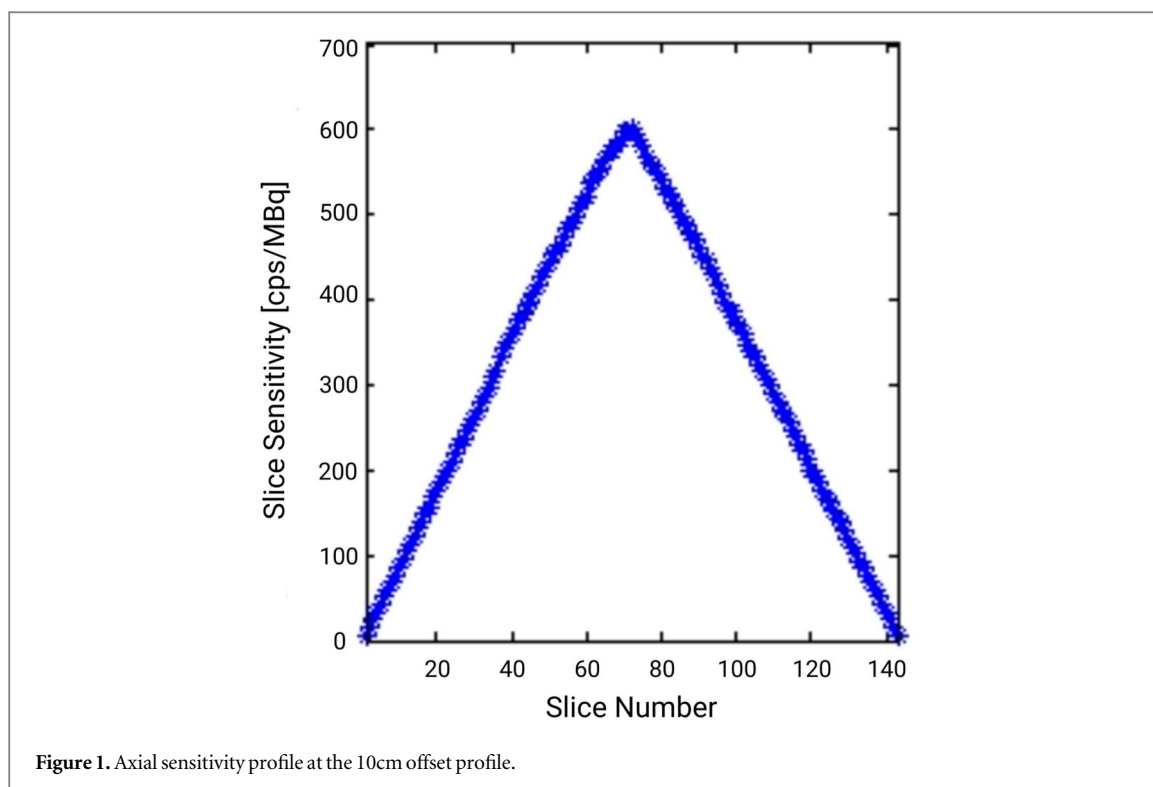
2.9. PET-CT Co-registration accuracy

An assessment of the alignment between the PET and CT data was made to assess any co-registration error. Data was acquired with PET and CT fiducial markers at 3 locations within the PET and CT field of view at two different axial locations (20 cm from tip of table and 100 cm from tip of table). This is performed using a customized jig. IQ Spheres with dimensions as specified in NEMA section 9.4.6 (17 mm, 22 mm and 28 mm) were filled with 3 MBq ml^{-1} of ^{18}F and CT contrast medium. The spheres were attached to the jig at positions (0,1), (0,20) and (20,0) (x(cm),y(cm)) respectively. 115 Kg of weight were distributed along the table in a uniform manner. Two acquisitions were performed, firstly with the phantom placed at 20 cm from the tip of the table and secondly with the phantom placed at 100 cm from the tip of the table. The centroids of the fiducial markers were calculated within the PET and CT data, and the co-registration error was determined by calculating the distance between the centroids.

3. Results

3.1. Spatial resolution

At 1 cm radial offset the tangential, radial and axial spatial resolutions were measured using both FBP and VPHD. For VPHD these values are 3.73 mm, 3.7 mm and 4.25 mm FWHMs respectively. At 10 cm radial offset, the resolutions were 5.11 mm, 3.85 mm and 4.22 mm. The FBP values along with the results at 20 cm radial offset are displayed in table 1, GE specification are displayed where available.

Table 1. Assessments of spatial resolution.

	Spatial resolution			
	Using VPHD FWHM mm		Using FBP FWHM mm	
Scan Type	Spec.	Measured	Spec.	Measured
Radial @1 cm	N/A	3.73	N/A	4.17
Tangential @1 cm	N/A	3.76	N/A	4.25
Transverse @1 cm	4.29	3.75	4.84	4.21
Axial @1 cm	4.56	4.25	5.28	5.12
Radial @10 cm	N/A	5.11	N/A	5.91
Tangential @10 cm	N/A	3.85	N/A	4.56
Transverse @10 cm	5.06	4.48	5.61	5.23
Axial @10 cm	5.04	4.22	6.84	6.07
Radial @20 cm	N/A	7.67	N/A	7.9
Tangential @20 cm	N/A	4.19	N/A	4.94
Transverse @20 cm	N/A	5.93	N/A	6.42
Axial @20 cm	N/A	4.21	N/A	6.3

Table 2. System sensitivity.

	Sensitivity		
Sensitivity	Units	Specification	Measured
Center	cps/kBq	41.4	44.36
10 cm off center	cps/kBq	38.7	44.63

3.2. System sensitivity

The total system sensitivity with the line source at the center of the FOV was 44.36 cps/kBq. At 10 cm off-center, this result was 44.63 cps/kBq. The axial sensitivity profile at 10 cm offset is shown in figure 1. Table 2 shows the result in comparison to the GE specification.
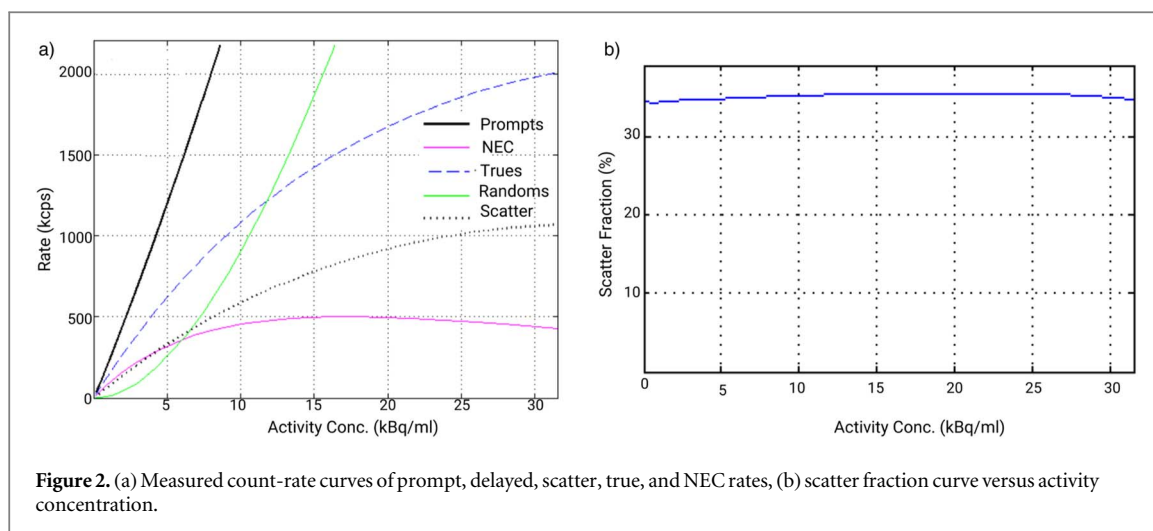
Figure 2. (a) Measured count-rate curves of prompt, delayed, scatter, true, and NEC rates, (b) scatter fraction curve versus activity concentration.
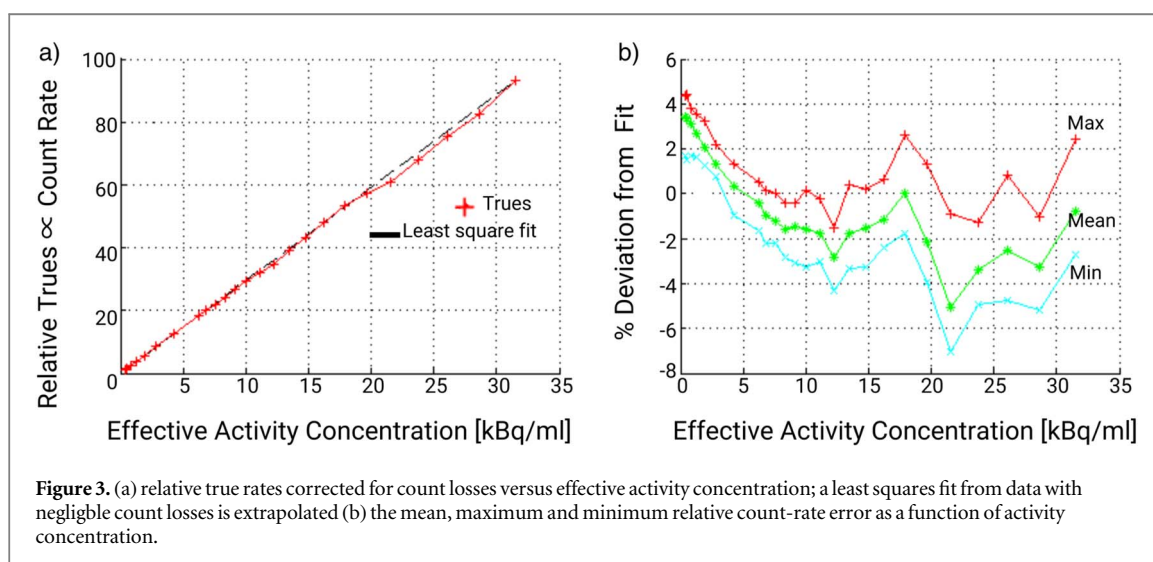

Figure 3. (a) relative true rates corrected for count losses versus effective activity concentration; a least squares fit from data with negligble count losses is extrapolated (b) the mean, maximum and minimum relative count-rate error as a function of activity concentration.

Table 3. NECR.

	Noise equivalent count rate	
Units	Specification NECR Peak Lower Limit	Measured
kcps	450	501

Table 4. Count rate accuracy assessed with maximum absolute error below the peak NECR.

	Noise equivalent count rate	
	Specification Max absolute error	
Units	(below peak NECR)	Measured
%	5.5	4.4

3.3. Scatter fraction, count losses and randoms

The NEC rates, prompts, trues, randoms, and scatter varied with radioactivity concentration and are plotted as curves shown in figure 2. The peak NEC rate is measured as 501 kcps at 17.8 kBq ml^{-1} activity concentration. The scatter fraction at Peak NECR was 35.48%, this is below the acceptance criteria of 38.5%. Figure 3(a) shows the relative true coincident rate (trues) calculated using count loss and random corrections. The true rate is extrapolated using a linear fit from measurements with negligible count losses and randoms. This highlights the accuracy of count loss corrections as a function of activity concentration. The minium , maximum and mean deviation of the extrapolated versus the calculated / corrected is shown in figure 3(b). The maximum absolute error below peak NECR peak was determined as 5.5%. These values are compared against GE specification in tables 3 and 4 respectively.

3.4. Image quality and optimization of image reconstruction parameters.

The averaged sphere-to-background ratio at 4:1 revealed a range of Contrast Recovery Coefficients (CRC) from 52.7 ±3.2% (for 10 mm spheres) to 92.5 ±2.4% (for 37 mm spheres), averaged across three acquisitions. Additionally, a lung residual measurement of 11.4 ±0.1% was obtained. A comprehensive

Table 5. Image quality.

							Lung error
			Hot spheres				
Diameter	10 mm	13 mm	17 mm	22 mm	28 mm	37 mm	50 mm
Manufacturer specified image quality metrics for acceptance testing.							
Specified Contrast VPHD %	30	40	50	60	60	60	19
Specified Background %	12	10	9	7	6	5	—
Measured Image Quality Metrics							
Measured Contrast VPHD %	52.7 ±3.2	63.7 ± 5.3	76.7 ± 2.1	78.8 ± 1.1	84.9 ± 2.5	92.5 ± 2.3	11.4 ± 0.1
Measured Background VPHD%	8.7 ± 0.5	6.8 ± 0.2	5.3 ± 0.2	4.2 ± 0.4	3.1 ± 0.2	2.5 ± 0.1	n/a
Measured Contrast ($\beta = 350$) Q. clear %	39.8 ± 2.5	58.2 ± 0.2	72.0 ± 0.2	75.6 ± 0.3	87.5 ± 0.2	93.4 ± 0.1	6.3 ± 0.2
Measured Background ($\beta = 350$) Q. clear %	2.4 ± 0.3	2.0 ± 0.3	1.7 ± 0.2	1.6 ± 0.2	1.4 ± 0.2	1.1 ± 0.1	
Measured Contrast ($\beta = 350$) Q.clear 'Medium' PDL %	52.0 ± 3.1	65.0 ± 0.9	78.7 ± 0.5	75.9 ± 1.0	88.3 ± 1.3	96.5 ± 0.2	2.2 ± 0.4
Measured Background ($\beta = 350$) Q. clear 'Medium' PDL %	2.8 ± 0.5	2.5 ± 0.3	2.3 ± 0.2	2.2 ± 0.1	2.0 ± 0.1	1.9 ± 0	n/a

For acceptance the measured contrast should be greater than or equal to the specified values. The measured lung error should be less than or equal to the specified values. The measured background values should be less than or equal to the specified values.

Table 6. Image quality optimization.

							Across all spheres
			Hot spheres				
Diameter	10 mm	13 mm	17 mm	22 mm	28 mm	37 mm	
Optimized to Maximize Contrast							
PDL Setting	H	H	H	L	H	M	H
Q. Clear Beta	100	100	100	700	800	900	100
Optimized to Minimize Background Variability							
PDL Setting	L	L	L	M	M	M	L
Q. Clear Beta	700	800	300	100	100	100	900
Optimized to Both Maximize Contrast Recovery and Minimize Background Variability: (Cost Function = 0.5 x—CR + 0.5 x BV)							
PDL Setting	H	L	L	L	M	M	H
Q. Clear Beta	400	200	700	700	100	100	100

Optimization is conducted on normalized contrast recovery and background variability. In this process, every image quality metric within each region is systematically normalized, assigning a scale of 0 to 1 based on the observed range across all image reconstruction parameters. Averaging the outcomes across all spheres and optimizing for both contrast recovery and background variability, the resultant average optimized setting, rounded to the nearest option, is Flavor 'M' Q.Clear Beta = 350.

summary of all CRCs, corresponding Background Variability (BV), and the GE specifications is presented in table 5. This overview encompasses the VPHD reconstruction, Q.Clear processing, and the results of PDL processing using the optimal setting as determined by experimentation. Figure 6 provides a visual representation of the CRC and BV for various combinations of settings across different sphere sizes. The optimal parameters, as discerned through the genetic algorithm, are detailed in table 6. The mean parameter configuration, calculated by averaging across the individual spheres and their settings across the entire set, was found to be $\beta = 350$ and PDL = 'medium.' The central slice of the image quality phantom for the 4:1 measurement, the CRC curves, BV and lung residual, for a single VPHD

Table 7. PET-CT Co-registration.

PET-CT Co-registration			
Scan Type	Units	Specification	Measured
Max Co-registration Error	mm	5	2.4

reconstructions are shown in figure 4. Figure 5 displays the central slice of the image quality phantom for reconstructions using VPHD, Q.Clear and PDL. Also presented in figure 7 are the horizontal and vertical line profiles for the spheres within the central slice; displayed are line profiles for reconstructions with VPHD, Q.Clear ($\beta = 350$) and PDL 'medium' ($\beta = 350$).
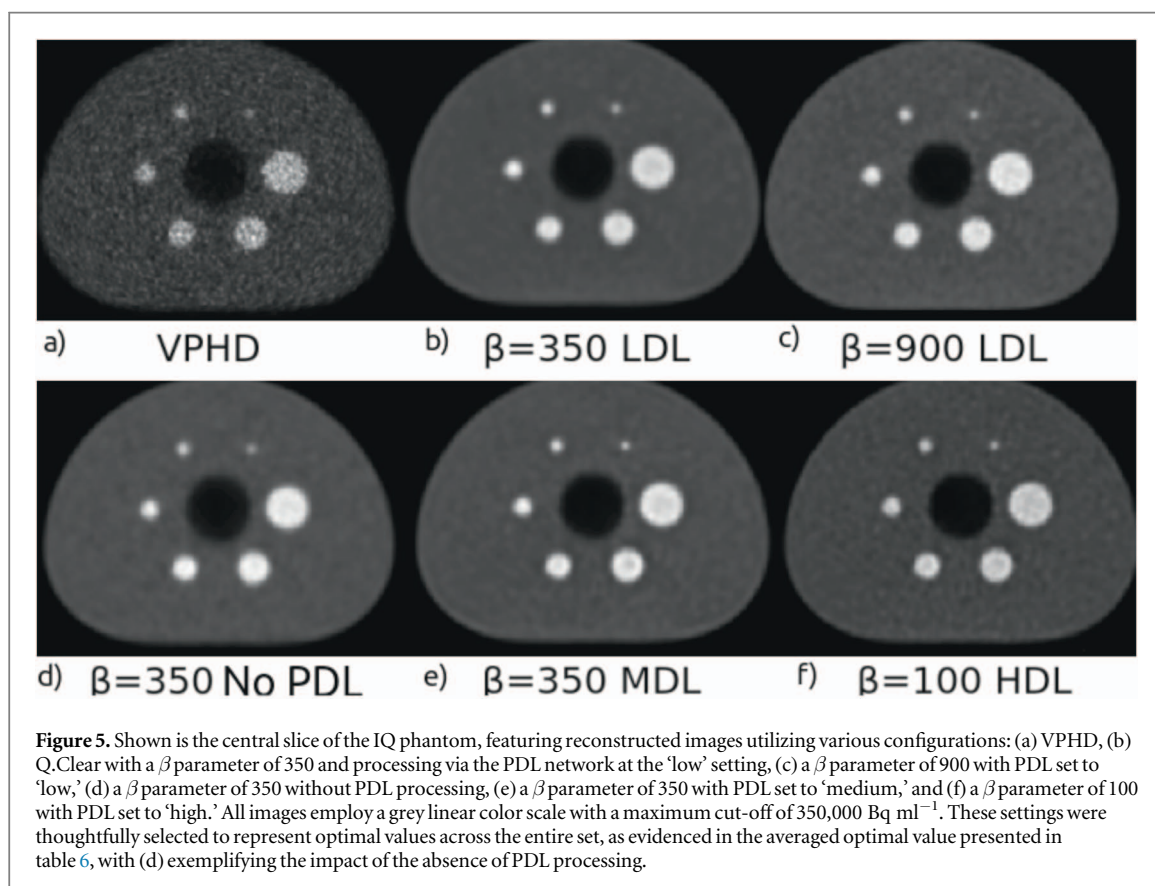
Figure 4. (a) Contrast recovery curves from a NEMA IQ acquisition reconstructed using VPHD, (b) displays the background variability, (c), the lung error versus slice number and (d) the central reconstructed slice showing region of interest position.

Figure 5. Shown is the central slice of the IQ phantom, featuring reconstructed images utilizing various configurations: (a) VPHD, (b) Q.Clear with a β parameter of 350 and processing via the PDL network at the 'low' setting, (c) a β parameter of 900 with PDL set to 'low,' (d) a β parameter of 350 without PDL processing, (e) a β parameter of 350 with PDL set to 'medium,' and (f) a β parameter of 100 with PDL set to 'high.' All images employ a grey linear color scale with a maximum cut-off of 350,000 Bq ml^{-1}. These settings were thoughtfully selected to represent optimal values across the entire set, as evidenced in the averaged optimal value presented in table 6, with (d) exemplifying the impact of the absence of PDL processing.

Table 8. Comparison.

System	Omni Legend	Discovery MI 6 ring	Discovery IQ
Assessment	Measured NEMA NU2–2018	Zeimpekis *et al* [18] NEMA NU2–2018	Reyn´es-Llompart *et al* [17] NEMA NU2–2012
Parameter ↓			
Axial FOV (cm)	32	30	26
Detector (type)	SiPM	SiPM	PMT
Scintillator type	BGO	LYSO	BGO
Scintillator size (mm)	.4.1 × 4.1 × 30	3.95 × 5.3 × 25	6.3 × 6.3 × 30
Sensitivity (cps/kBq) (center)	44.36	32.64	22.8
Sensitivity (cps/kBq) (10 cm)	44.63	32.88	20.43
Peak NECR (kcps)	501	434.3	123.6
Peak NECR conc. (KBq/ml)	17.8	23.6	9.1
SF % at Peak NECR	35.48	40.21	36.2
Max error at peak NECR (%)	5.5	3.95	3.9
Spatial Resolution	VPHD 6 iterations, 22 subsets, 2 mm Gaussian cut-off	VPHD 4 iterations, 34 subsets, 2 mm Gaussian cut-off	VPHD 12 iterations; 12 subsets; 2 mm Gaussian cut-off [a]
Radial @1 cm	3.73	3.72	4.2
Tangential @1 cm	3.76	3.87	4.7
Transverse @1 cm	3.75	3.8	4.45
Axial @1 cm	4.25	4.26	4.8
Radial @10 cm	5.11	4.8	5.6
Tangential @10 cm	3.85	3.79	5.1
Transverse @10 cm	4.48	4.3	5.4
Axial @10 cm	4.22	4.55	4.8
Radial @20 cm	7.67	7.63	8.5
Tangential @20 cm	4.19	4.21	5.5
Transverse @20 cm	5.93	5.95	7
Axial @20 cm	4.21	4.50	4.8
Image Quality	VPHD 6 iterations, 22 subsets, 2 mm Gaussian cut-off	VPFX 4 iterations, 34 subsets, 2 mm Gaussian cut-off	VPHD 8 iterations, 12 subsets, 2.mm Gaussian cut-off
Diameter Hot Sphere			
10 mm % Contrast	52.7 ± 3.2	54.5 ± 5.5	27
13 mm % Contrast	63.7 ± 5.3	63.2 ± 3.2	53
17 mm % Contrast	76.3 ± 2.1	68 ± 1.5	75
22 mm % Contrast	78.9 ± 1.1	76.9 ± 4.0	83
28 mm % Contrast	84.9 ± 2.5	82.4 ± 1.25	70
37 mm % Contrast	92.5 ± 2.3	85.8 ± 1.6	75
Lung error			
50 mm % Contrast	11.4 ± 0.1	3.16 ± 0.1	17.3
Background Variability			
10 mm %	8.7 ± 0.5	6.8 ± 0.5	9.8
13 mm %	6.8 ± 0.2	5.0 ± 0.6	7.9
17 mm %	5.3 ± 0.2	4.0 ± 0.6	6.0
22 mm %	4.2 ± 0.4	3.2 ± 0.3	4.8
28 mm %	3.1 ± 0.2	2.5 ± 0.3	4.3
37 mm %	2.5 ± 0.1	1.9 ± 0.2	3.9

[a] Please note reported values for the contrast recovery and background variability are reported as mean ± SD % where, in the case of Reyn´es-Llompart *et al* only the mean is available and reported.

3.5. PET-CT Co-registration

The Maximum PET-CT Co-registration error was measured to be 2.4 mm. Table 7 displays the result together with the GE specification.

4. Discussion

In this work the PET performance of the Omni Legend system with a 32-cm axial field-of-view was evaluated.

The Omni Legend provides the first digital BGO based PET-CT system on the market. The emergence of digital PET-CT has seen a leap forward in PET image quality and lesion detectability, particularly for smaller lesions [10]. Increasing the scanner geometry, with an extended axial field-of-view, inherently increases the system sensitivity, albeit at increased cost. The BGO crystal, at a lower cost than L(Y)SO boasts higher detection efficiency and plays an important role in selecting a suitable scintillator for extended FOV's.

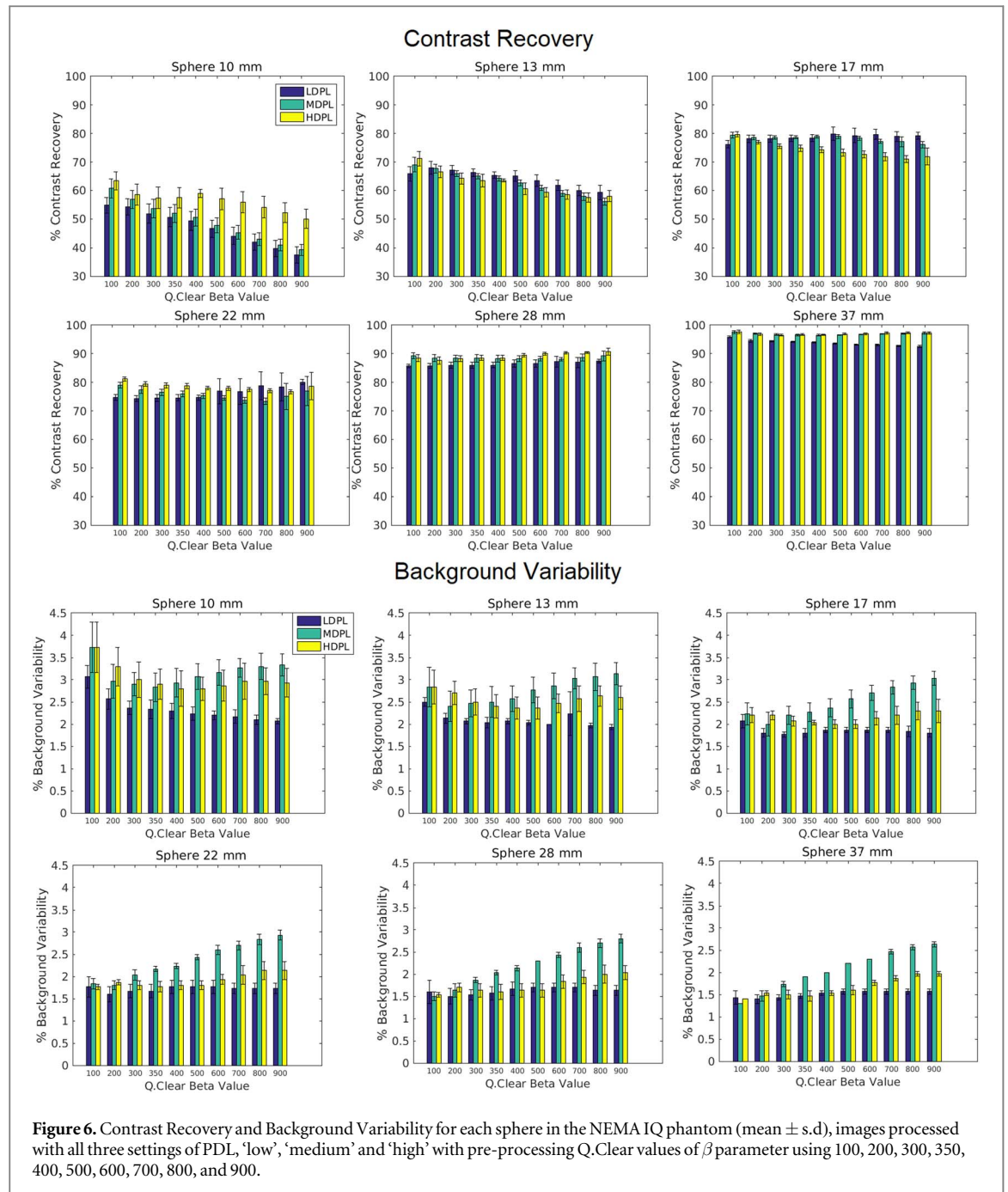**Table 9.** Comparison omni legend and discovery IQ.

System		Omni Legend Measured NEMA	Discovery IQ Reyn´es-Llompart
Assessment		NU2–2018	*et al* (9). NEMA NU2–2012
Recon Type		VPHD	VPHD
Iteration		4	4
Subset		12	12
Smoothing		4.8 mm Gaussian	4.8 mm Gaussian
Image Quality			
Diameter Hot Sphere			
10 mm	% Contrast	37.5 ± 3.1	25
13 mm	% Contrast	52.2 ± 2.8	40
17 mm	% Contrast	64.9 ± 2.1	61
22 mm	% Contrast	69.0 ± 2.1	68
28 mm	% Contrast	76.5 ± 2.5	64
37 mm	% Contrast	85.5 ± 2.3	68
Lung error			
50 mm	% Contrast	22.5 ± 0.1	22.2
Background Variability			
10 mm	%	3.9 ± 0.3	5.5
13 mm	%	3.4 ± 0.2	4.9
17 mm	%	2.9 ± 0.2	4.2
22 mm	%	2.3 ± 0.1	3.6
28 mm	%	1.9 ± 0.2	3.4
37 mm	%	1.5 ± 0.1	3.3

Further advantages may be obtained by using a BGO coupled to digital detector technology. BGO does however lack the timing resolution of Lu-based scintillators. However, despite their advantages such as higher light output and faster decay times ,intrinsic radioactivity in scintillator materials like L(Y)SO, can contribute to background noise in PET images, potentially affecting the clarity and diagnostic accuracy of the scans. The time of flight capabilities of L(Y)SO do show demonstrable improvements in Signal to Noise ratio (SNR) and hence greater contrast to noise ratio [11]. Promising work does however exist in using BGO Cherenkov based emission for improved coincidence timing. [12] As the Omni Legend is a digital BGO system with extended FOV, for comparative purposes, we assess its measured performance against two commercially available GE systems. Firstly, the 6 ring Discovery MI, which has an axial FOV of 30 cm consisting of lutetium-yttrium-oxyorthosilicate (LYSO) scintillator crystals (3.95 mm × 5.3 mm × 25 mm) backed onto SiPM arrays. Secondly, at the time of writing, the Discovery IQ, which from the GE BGO range has the highest reported axial FOV of 26 cm (other than the Omni Legend) [13]. The Discovery IQ crystals (6.3 mm × 6.3 mm × 30 mm) are coupled with traditional photo-multiplier tube PMT technology. The summary comparison is described in table 8. It is clear to observe that the sensitivity of the Omni, far surpasses the Discovery MI

and Discovery IQ by ∼30% and ∼65% respectively. The primary factor contributing to the enhanced sensitivity of these systems is the increase in Axial Field of View (AFOV). A distinction that is further underscored by the improved contrast recovery observed in the Omni as compared to the Discovery IQ. Time of flight reconstructions have negligible impact on spatial resolution of reconstructed images at current timing resolutions [14]. Spatial resolution is compared with the clinically used VPHD algorithm for all three systems. With the larger voxel size, the Discovery IQ's spatial resolution as expected is inferior. Taking the average spatial resolution (Radial, tangential, axial) at all offsets (1 cm, 10 cm and 20 cm), the Omni resolution is 4.55 mm versus 4.59 mm of the Discovery MI.
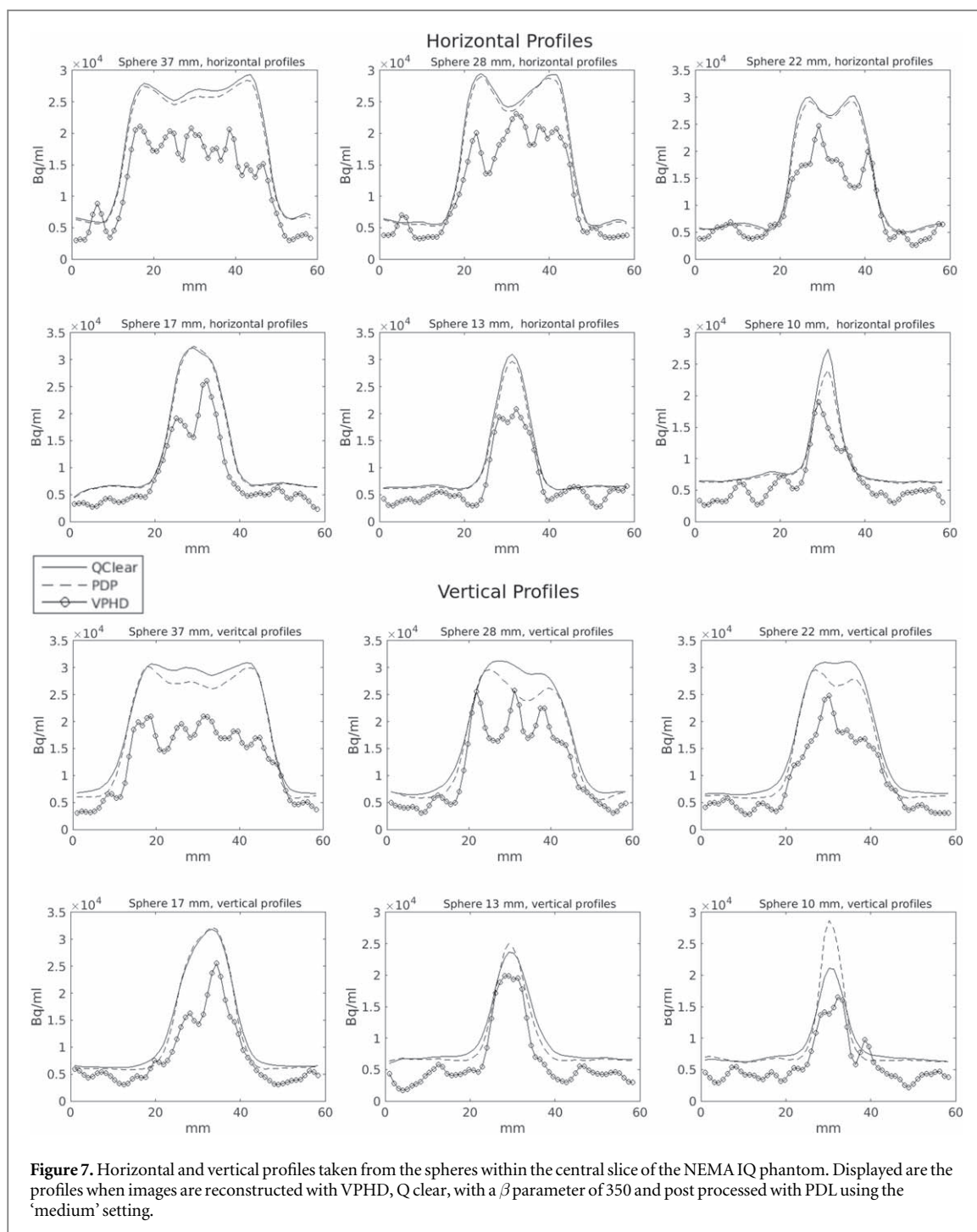
Considering the image quality results, a comparison using manufacturer recommended reconstruction settings of VPFX (TOF-OSEM) from the Discovery MI 6-ring to the VPHD of the Omni Legend is made. We observe an average improvement (average across all spheres) in contrast recovery of 4% for the Omni. The Omni does, however, demonstrate an average increase in background variability by 27% in comparison to the Discovery MI. The background variability reflects the noise in the image and demonstrates the improved performance of VPFX versus VPHD. The Omni does however exhibit an 18% decrease in background variability in comparison to the Discovery IQ (PMT based detectors) and a 16% increase in contrast recovery. Increased background variability in SiPM versus PMT systems has previously been reported [15] and suggested that the benefits of digital systems over analogue mainly being the utilization of advanced reconstruction algorithms [16]. This increase in peformance of the Omni versus the Discovery IQ could also be predominantly attributed to the extended field of view of the Omni and hence greater sensitivty. As Image quality metrics can vary with image reconstruction settings a comparison is also made of the Omni VPHD and Discovery IQ VPHD reconstructions using the clinical parameters as published in the Discovery IQ paper (table 9). This demonstrates the OMNI has a 33% decrease in BV compared to the Discovery IQ and a 15% increase in contrast recovery.

To provide the image quality performance benefits most associated with hardware-based Time-of-Flight the Omni provides the PDL algorithm, to map non-ToF images to ToF images. The PDL algorithm takes as input the Bayesian Penalizaed Likelihood Reconsuction images, termed Q.Clear, which themselves have a parameter β which determines the noise supression. Increasing values of β reduce both contrast recovery and background variability. No work has yet explored the effect of contrast recovery and background variability of Q.clear images, processed with the new PDL algorithm. Furtermore, no work has explored the interplay between the β parameter and PDL setting and the optimization of these

Figure 6. Contrast Recovery and Background Variability for each sphere in the NEMA IQ phantom (mean ± s.d), images processed with all three settings of PDL, 'low', 'medium' and 'high' with pre-processing Q.Clear values of β parameter using 100, 200, 300, 350, 400, 500, 600, 700, 800, and 900.

hyperparameters. Figure 6 demonstrates, as expected, the reduction in contrast recovery with increasing β which is more emphasized at small sphere sizes. Background variability, as expected, also reduces with increasing β when using PDL at the 'low' setting. This is more predominant at smaller sphere sizes. With increasing the PDL setting, at a small sphere size of 10 mm, we observe less of a reduction in contrast recovery with increasing β, thus circumnavigating the traditional reduction in contrast that is observed with increasing β. However, this effect is not consistent with increasing sphere size. Furthermore, the background variability tends to increases with increasing β using the 'medium' and 'high' PDL setting. This is more marked with the 'medium' setting which results in the highest background variability. This interplay

demonstrates the importance of optimizing hyperparameters whilst also taking into account object size. Table 6 summarizes the optimal parameters in terms of PDL setting and Q.clear β value. We found that a low ($\beta = 100$) and high PDL setting optimizes contrast at small sphere sizes (10 mm, 13 mm and 17 mm), whilst the low PDL setting (with average $\beta = 900$) minimizes background variability. For larger sphere sizes (22 mm, 28 mm and 37 mm), the medium PDL setting with a $\beta = 100$ minimizes background variability, whereas a higher Q.Clear β value ($\geqslant 700$) combined with varying PDL settings, achieves optimal contrast recovery. This illustrates the dynamic interplay between Q.Clear β values and PDL settings across different sphere sizes. The novelty and significance of our work therefore lies in the application of a genetic

Figure 7. Horizontal and vertical profiles taken from the spheres within the central slice of the NEMA IQ phantom. Displayed are the profiles when images are reconstructed with VPHD, Q clear, with a β parameter of 350 and post processed with PDL using the 'medium' setting.

algorithm to optimize these hyperparameters. This approach is pivotal, especially in situations characterized by diverse object sizes and complex interactions among image quality metrics. The algorithm thoroughly explores parameter space and enhances understanding of parameter interplay without needing a precise differentiable mathematical description of the cost function. Moreover, it underscores the imperative need for more advanced optimization strategies when seeking to determine optimal settings. For the purposes of optimized image quality in terms of NEMA testing a joint cost function that maximizes contrast recovery whilst minimizing background variability determined the optimal parameters for the Omni,

when peforming NEMA IQ testing to be Q.Clear ($\beta = 350$) and PDL = 'medium'. It can be observed in figure 7 that count recovery from spheres is more pronounced with Q.Clear ($\beta = 350$) and PDL ('medium') in comparison to VPHD. For the 10 mm lesion, Q. Clear in comparison to VPHD, increases counts, when averaged along the horizontal and vertical profiles by ∼19%; for PDL, this value is ∼58%.

Using the optimized settings and comparing contrast recovery for VPHD, Q.Clear ($\beta = 350$) and PDL ('Medium' with Q.Clear ($\beta = 350$) against VPFX of the Discovery MI; the Omni demonstrates when averaged across all spheres, a 4% increase, 1% decrease and 6% increase respectively. The background variability

demonstrates a 27% increase, 79% reduction and a 56% reduction in comparison to the Discovery MI (VPFX). The 'medium' PDL algorithm at this setting ($\beta = 350$) when averaged across all spheres, in comparison to Q.Clear alone, increases contrast recovery by 7%, however we also observe an increase in background variability of 26%. In summary, when using optimized advanced reconstruction settings (Q.Clear $\beta = 350$ and PDL = 'Medium'), the background variability and contrast recovery of the Omni demonstrate superior performance in comparison to the VPFX of the Discovery MI with on average a 56% reduction and 6% increase respectively This demonstrates, when evaluated on an overall basis, the Omni's ability to recover image quality performance comparable to a ToF system. It is also important to note that the PDL reconstruction algorithm was trained using clinical PET-CT scans from a variety of PET-CT scanners and not phantoms. Therefore, using standard phantoms to measure PDL performance should only be considered as a guide to performance. Variabilities inherently exist in the optimal value of β across lesion volumes when utilizing Q.Clear [9] which serves as the input to PDL. We optimized the hyperparameters using the novel application of a genetic algorithm. Future work in validating performance with varying hyperparameters in the clinical setting is required. This includes the interplay between both the Q.Clear β parameter and PDL setting. With optimization this will likely demonstrate improvements in both contrast recovery and background variability in the clinical setting coupled with reduced patient dose and / or clinical scan imaging time.

Furthermore, the application of our genetic algorithm for image reconstruction parameter optimization holds significant implications for clinical practice. It enables clinicians to fine-tune reconstruction parameters, tailoring them to specific diagnostic challenges. For example, while the optimized parameters that work well for general whole body imaging may provide good overall image quality, different settings might be more effective for small lesion detection or for images requiring specific views. This ability to adapt the imaging parameters according to the clinical scenario—such as enhancing detail in small lesions or adjusting the view for clearer visualization of certain anatomical areas—can lead to more accurate diagnoses, particularly in complex cases. This optimization process, therefore, has a direct clinical utility: by applying these tailored settings, clinicians can achieve greater precision in lesion detection and characterization, leading to improved patient outcomes. It underscores the need for an intelligent, adaptable approach to PET-CT imaging, where the 'one-size-fits-all' model is replaced by a more nuanced, patient-specific strategy.

5. Conclusion

A NEMA performance evaluation is made of the new digital Omni Legend. A comparison was made to the Discovery MI (6 ring) and Discovery IQ. The Omni demonstrates increased sensitivity in comparison to its counterparts, allowing the possibility of shorter scan times and/or less patient dose. Omni image quality in terms of spatial resolution and contrast recovery is competitive even in comparison to an LYSO system (Discovery MI). Using traditional reconstruction methods, notably VPHD with the manufacturer's recommended settings, the Omni does show increased background variability (noise) compared to the Discovery MI using the VPFX. However, this is effectively mitigated by implementing software corrections like Q.Clear and PDL. When these parameters are optimized with our proposed genetic algorithm, they compensate for the Omni's lack of Time-of-Flight (ToF) capabilities, leading to improved phantom image quality. This results in increased contrast and reduced background variability for the Omni, surpassing the performance seen with the VPFX of the Discovery MI. Moreover, our research demonstrates the effectiveness of the genetic algorithm in optimizing image reconstruction settings, applicable even in situations where cost functions lack a clear derivative. This work underscores the potential of using advanced algorithms for personalized and precise adjustment of PET-CT parameters in clinical settings. The ability to tailor imaging protocols to specific patient scenarios, such as optimizing for small lesion detection or varying geometries like lesion size and attenuating medium, can significantly enhance diagnostic accuracy and patient outcomes. Future work will further explore these applications, particularly assessing the impact of hyperparameters in the clinical setting. Our findings pave the way for more nuanced imaging strategies that could lead to improvements in both image quality and patient care, including optimized patient dose and imaging time. The flexibility and adaptability offered by this algorithm could be particularly beneficial in optimizing image quality for diverse clinical scenarios, affirming the crucial role of optimization in the evolving landscape of medical imaging.

Data availability statement

All data is contained within the publication. The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Rhodri Lyn Smith ● https://orcid.org/0000-0001-9359-4730

References

[1] Townsend D W 2008 Combined positron emission tomography- computed tomography: the historical perspective *Seminars in Ultrasound, CT and MRI* **29** 232–5

[2] Aide N, Lasnon C, Desmonts C, Armstrong I S, Walker M D and McGowan D 2022 Advances in PET/CT technology: an update *Semin. Nucl. Med.* **52** 286–301

[3] NEMA 2018 NU 2-2018-performance measurements of positron emission tomographs *National Electrical Manufacturers Association. Rosslyn, USA*

[4] Abouhawwash M and Alessio A M 2021 Multi-objective evolutionary algorithm for PET image reconstruction: Concept *IEEE Trans. Med. Imaging* **40** 2142–51

[5] Marinelli Y, Positano M, Tucci V, Neglia D F and Landini L 2012 Automatic PET-CT image registration method based on mutual information and genetic algorithms *The Scientific World Journal* **2012** 567067

[6] Mehranian A *et al* 2022 Deep learning-based time-of-flight (ToF) image enhancement of non-ToF PET scans *Eur. J. Nucl. Med. Mol. Imaging* **49** 3740–9

[7] Hudson H M and Larkin R S 1994 Accelerated image reconstruction using ordered subsets of projection data *IEEE Trans. Med. Imaging* **13** 601–9

[8] Asma E, Ahn S, Ross S G, Chen A and Manjeshwar R M 2012 Accurate and consistent lesion quantitation with clinically acceptable penalized likelihood images *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC)* 4062–6

[9] Andersen T and Hoilund-Carlsen P F 2016 The Q. Clear PET reconstruction algorithm: evaluation using the NEMA IQ phantom *J. Nucl. Med.* **57** 1973

[10] de Jong T L *et al* 2022 Performance of digital PET/CT compared with conventional PET/CT in oncologic patients: a prospective comparison study *Annals of Nuclear Medicine.* **36** 756–64

[11] Conti M and Bendriem B 2019 The new opportunities for high time resolution clinical TOF PET *Clinical and Translational Imaging.* 2019 **7** 139–47

[12] Efthimiou N *et al* 2020 TOF-PET image reconstruction with multiple timing kernels applied on Cherenkov radiation in BGO *IEEE Transactions on Radiation and Plasma Medical Sciences* **5** 703–11

[13] Nadig V, Herrmann K, Mottaghy F M and Schulz V 2022 Hybrid total-body pet scanners–current status and future perspectives. *Eur. J. Nucl. Med. Mol. Imaging* **49** 445–59

[14] Toussaint M, Lecomte R and Dussault J P 2020 Improvement of spatial resolution with iterative PET reconstruction using ultrafast TOF *IEEE Transactions on Radiation and Plasma Medical Sciences* **5** 729–37

[15] Wagatsuma K *et al* 2017 Comparison between new-generation SiPM-based and conventional PMT-based TOF-PET/CT *Physica Med.* **42** 203–10

[16] Oddstig J, Brolin G, Trägårdh E and Minarik D 2021 Head-to-head comparison of a Si-photomultiplier based and a conventional photomultiplier based PET-CT system *EJNMMI Physics* **8** 1–12

[17] Reynés-Llompart G, Gámez-Cenzano C, Romero-Zayas I, Rodríguez-Bel L, Vercher-Conejero J L and Martí-Climent J M 2017 Performance characteristics of the whole-body discovery IQ PET/CT system *J. Nucl. Med.* **58** 1155–61

[18] Zeimpekis K G, Kotasidis F A, Huellner M, Nemirovsky A, Kaufmann P A and Treyer V 2022 NEMA NU 2-2018 performance evaluation of a new generation 30-cm axial field-of-view discovery MI PET/CT *Eur. J. Nucl. Med. Mol. Imaging* **49** 3023–32