NEMA NU 4-2008 Comparison of Preclinical PET Imaging Systems

Andrew L. Goertzen[1], Qinan Bao[2], Mélanie Bergeron[3], Eric Blankemeyer[4], Stephan Blinder[5], Mario Cañadas[6], Arion F. Chatziioannou[2], Katherine Dinelle[5], Esmat Elhami[1,7], Hans-Sonke Jans[8], Eduardo Lage[9], Roger Lecomte[3], Vesna Sossi[5,10], Suleman Surti[4], Yuan-Chuan Tai[11], Juan José Vaquero[12], Esther Vicente[13], Darin A. Williams[2], and Richard Laforest[11]

[1]Department of Radiology, University of Manitoba, Winnipeg, Canada; [2]Crump Institute for Molecular Imaging, University of California, Los Angeles, California; [3]Sherbrooke Molecular Imaging Center, Université de Sherbrooke, Sherbrooke, Canada; [4]Department of Radiology, University of Pennsylvania, Philadelphia, Pennsylvania; [5]PET Imaging Group, University of British Columbia, Vancouver, Canada; [6]Unidad de Aplicaciones Médicas, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, Madrid, Spain; [7]Department of Physics, University of Winnipeg, Winnipeg, Canada; [8]Department of Oncology, University of Alberta, Edmonton, Canada; [9]Unidad de Medicina y Cirugía Experimental, Hospital General Universitario Gregorio Marañón, Madrid, Spain; [10]Department of Physics and Astronomy, University of British Columbia, Vancouver, Canada; [11]Department of Radiology, Washington University School of Medicine, St. Louis, Missouri; [12]Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid, Spain; and [13]Instituto de Estructura de la Materia, Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain

The National Electrical Manufacturers Association (NEMA) standard NU 4-2008 for performance measurements of small-animal tomographs was recently published. Before this standard, there were no standard testing procedures for preclinical PET systems, and manufacturers could not provide clear specifications similar to those available for clinical systems under NEMA NU 2-1994 and 2-2001. Consequently, performance evaluation papers used methods that were modified ad hoc from the clinical PET NEMA standard, thus making comparisons between systems difficult. **Methods:** We acquired NEMA NU 4-2008 performance data for a collection of commercial animal PET systems manufactured since 2000: microPET P4, microPET R4, microPET Focus 120, microPET Focus 220, Inveon, ClearPET, Mosaic HP, Argus (formerly eXplore Vista), VrPET, LabPET 8, and LabPET 12. The data included spatial resolution, counting-rate performance, scatter fraction, sensitivity, and image quality and were acquired using settings for routine PET. **Results:** The data showed a steady improvement in system performance for newer systems as compared with first-generation systems, with notable improvements in spatial resolution and sensitivity. **Conclusion:** Variation in system design makes direct comparisons between systems from different vendors difficult. When considering the results from NEMA testing, one must also consider the suitability of the PET system for the specific imaging task at hand.

The use of PET to study animal models of human disease has been expanding. By the late 1990s, several groups had constructed prototype PET systems (*1–9*) because it had been found that significant benefits in spatial resolution, sensitivity, image quality, and quantification were achievable using systems designed specifically for small laboratory animals. In 2000, commercial preclinical PET systems became available, and over the next 10 y the performance and capabilities of these systems evolved rapidly.

The maturing market for preclinical PET systems led to the need for standardized methods of performance evaluation. Such standardization facilitates acceptance testing and routine monitoring and allows comparison between systems from different vendors and of different designs. To address this need, the National Electrical Manufacturers Association (NEMA) NU 4 standard was published in 2008 (*10*). Before then, there was no agreed-upon method to evaluate the performance of preclinical PET systems, and manufacturers could not provide specifications as is done for clinical systems under the NEMA NU 2-1994 (*11*) and NU 2-2001 (*12*) standards. In addition, performance evaluation articles on many early-generation preclinical PET systems used methods that were modified ad hoc from the clinical NEMA standard. As no two early systems were evaluated in a consistent manner, it was

TABLE 1
Specifications of Systems Included in Test

Manufacturer	Model	Detector design	Scintillator	Ring diameter (mm)	Axial FOV (mm)	Crystal size (mm³)	Rotating ring	Reference
Concorde Microsystems/Siemens	microPET P4	Block, 8 × 8 array with PSPMT	LSO	261	78	2.2 × 2.2 × 10	No	(16)
Concorde Microsystems/Siemens	microPET R4	Block, 8 × 8 array with PSPMT	LSO	148	78	2.2 × 2.2 × 10	No	(17)
Concorde Microsystems/Siemens	microPET Focus 220	Block, 12 × 12 array with PSPMT	LSO	258	76	1.51 × 1.51 × 10	No	(18)
Concorde Microsystems/Siemens	microPET Focus 120	Block, 12 × 12 array with PSPMT	LSO	147	76	1.51 × 1.51 × 10	No	(19,20)
Siemens	Inveon	Block, 20 × 20 array with PSPMT	LSO	161	127	1.51 × 1.51 × 10	No	(15,21–24)
Philips	Mosaic HP	Pixelated Anger logic, 19 mm PMTs	LYSO	197	119	2 × 2 × 10	No	(25,26)
Raytest GmbH	ClearPET	Block, 8 × 8 dual-layer phoswich array with PSPMT	LYSO/LuYAP	135–225	110	2 × 2 × 10 + 2 × 2 × 10	Yes	(27)
Sedecal	Argus (formerly GE eXplore Vista)	Block, 13 × 13 dual-layer phoswich array with PSPMT	LYSO/GSO	118	48	1.45 × 1.45 × 7 (LYSO); 1.45 × 1.45 × 8 (GSO)	No	(28)
Sedecal	VrPET	Block, 30 × 30 array with PSPMT	LYSO	140	45.6	1.4 × 1.4 × 12	Yes	(29)
Gamma Medica	LabPET 8	Phoswich detector, 2 crystals per APD	LYSO/LGSO	162	75	2 × 2 × 11.9 (LYSO); 2 × 2 × 13.3 (LGSO)	No	(30,31)
Gamma Medica	LabPET 12	Phoswich detector, 2 crystals per APD	LYSO/LGSO	162	112.5	2 × 2 × 11.9 (LYSO); 2 × 2 × 13.3 (LGSO)	No	(32)

PSPMT = position-sensitive photomultiplier tube; LSO = lutetium oxyorthosilicate; LYSO = lutetium yttrium oxyorthosilicate; LuYAP= lutetium yttrium aluminum perovskite; GSO = germanium oxyorthosilicate; LGSO = lutetium gadolinium oxyorthosilicate.

TABLE 2
Spatial Resolution Results

System energy window reconstruction	Radial offset (mm)	FWHM/FWTM (mm) at axial center of FOV			FWHM/FWTM (mm) at ¼ axial offset		
		Radial	Tangential	Axial	Radial	Tangential	Axial
microPET P4, 350–650 keV, Fourier rebinning + 2D FBP	5	2.29/4.03	2.18/3.81	2.20/4.52	2.34/4.22	2.14/3.77	1.75/4.22
	10	2.41/4.23	2.23/3.92	2.38/4.66	2.37/4.14	2.22/3.84	1.97/4.49
	15	2.42/4.19	2.28/3.83	2.42/4.68	2.39/4.16	2.27/3.87	2.04/4.53
	25	2.61/4.67	2.25/3.76	2.42/4.67	2.53/4.41	2.30/3.91	2.07/4.50
	50	3.27/6.40	2.40/4.10	2.58/5.09	3.20/6.08	2.45/4.29	2.30/4.74
	75	3.92/8.07	2.64/4.53	2.88/5.99	3.78/7.12	2.81/5.15	2.72/5.58
microPET R4, 350–650 keV, Fourier rebinning + 2D FBP	5	2.13/4.90	2.21/4.22	2.72/5.59	2.06/5.24	2.18/4.14	2.37/4.88
	10	2.30/4.60	2.31/4.36	3.02/6.54	2.30/4.61	2.29/4.39	2.66/5.31
	15	2.86/5.38	2.39/4.57	3.25/7.48	2.63/5.35	2.35/4.40	2.84/5.71
	25	3.30/6.32	2.51/4.66	3.27/7.57	3.31/6.23	2.53/4.80	3.09/6.31
microPET Focus 120, 350–650 keV, Fourier rebinning + 2D FBP	5	1.92/3.66	1.66/3.06	1.90/3.81	1.92/3.63	1.65/3.09	1.62/3.28
	10	1.88/3.95	1.74/3.22	1.94/3.91	1.83/3.69	1.76/3.28	1.66/3.34
	15	1.99/4.02	1.72/3.11	1.98/4.05	1.94/3.75	1.77/3.22	1.69/3.41
	25	2.53/4.84	1.73/3.01	2.05/4.34	2.45/4.49	1.80/3.20	1.81/3.67
microPET Focus 220, 250–750 keV, Fourier rebinning + 2D FBP	5	1.75	1.80	1.70	1.70	1.70	1.73
	10	1.68	1.78	1.73	1.66	1.79	1.75
	15	1.82	1.71	1.80	1.88	1.72	1.78
	25	2.07	1.69	1.84	2.09	1.74	1.87
	50	2.88	1.77	1.98	2.82	1.82	1.92
	75	4.08	1.90	2.16	3.92	1.90	2.11
Inveon, 350–625 keV, Fourier rebinning + 2D FBP	5	1.63/3.36	1.62/3.15	2.45/5.62	1.66/3.32	1.63/3.14	1.97/4.20
	10	1.80/3.84	1.58/2.91	2.40/5.51	1.72/3.40	1.64/3.18	2.12/4.44
	15	2.03/4.32	1.56/2.78	2.29/5.32	1.87/3.69	1.63/3.05	2.17/4.72
	25	2.49/5.17	1.61/2.86	2.09/4.67	2.38/4.76	1.65/2.97	2.06/4.54
ClearPET, 250–650 keV, 3D FBP	5	1.94/3.76	2.00/4.17	3.24/6.05	2.18/4.05	1.97/3.92	3.18/5.91
	10	1.85/3.47	2.27/5.97	3.19/5.97	1.87/3.68	2.14/4.86	3.20/5.88
	15	2.01/3.62	2.43/5.53	3.20/5.96	2.05/3.84	2.33/5.25	3.19/5.83
	25	2.55/4.28	2.42/5.69	3.21/5.97	2.50/4.18	2.43/6.59	3.19/5.85
Mosaic HP, 385–665 keV, 3D Fourier reprojection	5	2.32/5.30	2.32/4.97	2.64/6.07	2.33/5.32	2.40/4.88	2.48/5.32
	10	2.45/5.48	2.51/4.96	2.82/6.14	2.37/5.54	2.49/4.97	2.80/5.89
	15	2.43/5.44	2.65/5.24	2.79/6.28	2.48/5.62	2.63/5.25	2.80/5.92
	25	2.59/5.93	2.83/5.25	2.96/6.28	2.63/5.81	2.87/5.31	3.10/6.30
Argus, 250–700 keV, 2D FBP	5	1.63	1.65		1.65	1.70	
	10	1.71	1.70		1.74	1.75	
	15	1.85	1.70		1.85	1.75	
	25	2.25	1.73		2.15	1.85	
VrPET, 100–700 keV, SSRB + 2D FBP	5	1.52/2.76	1.62/2.99	2.66/4.81	1.62/2.95	1.68/2.86	2.57/5.21
	10	1.58/2.85	1.68/3.02	3.03/5.45	1.54/2.89	1.68/3.07	3.11/5.45
	15	1.78/3.25	1.51/2.79	3.11/5.50	1.69/3.09	1.73/3.14	3.22/5.72
	25	2.03/3.69	2.12/3.72	3.32/5.92	1.79/3.29	1.98/3.60	3.60/6.87
LabPET 8, 250–650 keV, SSRB + 2D FBP	5	1.65/3.40	1.70/3.30	*	1.57/3.30	1.65/3.50	*
	10	1.91/3.60	1.82/3.67	*	1.92/3.40	1.74/3.45	*
	15	2.01/4.10	1.83/3.70	*	1.92/3.77	1.86/3.90	*
	25	2.56/4.65	1.90/4.28	*	2.55/4.70	1.93/4.30	*

*Axial intrinsic resolution (FWHM/FWTM) = 1.4/4.3 mm.

TABLE 3
Comparison of Effective Transaxial FWHM Resolution with Crystal Size

System	Crystal size (mm)	Effective transaxial FWHM resolution at 5 mm (mm)	Resolution/crystal size
microPET P4	2.2	2.24	1.02
microPET R4	2.2	2.20	1.00
microPET Focus 120	1.51	1.78	1.18
microPET Focus 220	1.51	1.74	1.15
Inveon	1.51	1.64	1.08
ClearPET	2	2.02	1.01
Mosaic HP	2	2.34	1.17
Argus	1.45	1.66	1.14
VrPET	1.4	1.61	1.15
LabPET 8	2	1.64	0.82

difficult to compare performance between early and newer camera designs.

In this work, we present NEMA NU 4-2008 performance measurements for a collection of preclinical PET systems that span the first 10 y of commercial availability. Our intent is to provide an objective source to which future systems can be compared, understand how the different design decisions of preclinical PET systems affect performance data, and examine whether the NU 4-2008 tests are adequate to characterize performance. We avoid making qualitative statements about whether one system is better than another (except when comparing systems from a single manufacturer). We also do not consider the performance of add-on features such as CT, animal-handling equipment, or data analysis software, all of which may factor into the choice of the optimal system for a given research program.

MATERIALS AND METHODS

Systems

To be included in this work, a PET system needed to have been commercially manufactured since 2000 and be in good working order. Prototype research systems were specifically excluded. The 11 systems included are summarized in Table 1.

Testing

All testing followed the NEMA NU 4-2008 standard (*10*) as closely as possible. We refer the reader to the NU 4-2008 standard for details. Data were collected and analyzed at each contributing site. For each system, the settings used, such as energy and timing windows and coincidence acceptance angle, were those typically applied in routine imaging. We mention these settings, as appropriate, when we list results. Reasonable effort was made to ensure the completeness of the data; however, in some cases complete results could not be obtained. All testing was performed independent of the system manufacturer and represents the performance of a single system of each model.
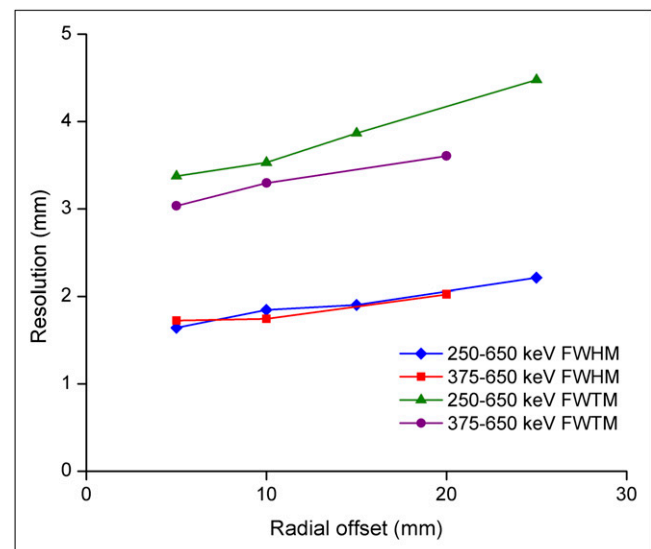
Spatial Resolution. Spatial resolution was measured using a ^{22}Na point source embedded in a 1-cm^3 acrylic cube. For each dataset, the full width at half maximum (FWHM) and full width at tenth maximum (FWTM) are reported for the axial center and ¼-axial-offset positions.

Sensitivity. The NU 4-2008 sensitivity measurement uses the same ^{22}Na point source as used for spatial resolution measure-

ment. We report values only for absolute sensitivity, a unitless percentage corrected for the 0.9060 branching fraction of ^{22}Na. We report the absolute system sensitivity for the mouse length ($s_{MA,tot}$), calculated as the average absolute sensitivity over the central 7 cm of the axial field of view (FOV), and total absolute system sensitivity ($s_{A,tot}$), calculated as the average absolute sensitivity over the entire axial FOV. We do not report the absolute system sensitivity for the rat length ($s_{RA,tot}$), since it is equivalent to total absolute sensitivity for all systems because the axial FOV is less than 15 cm.

For 2 systems (Inveon and Argus), a centered line source filled with ^{18}F and surrounded by an aluminum cylindric shell was used instead of the ^{22}Na source. In these cases, the sensitivity measured with the ^{18}F line source was calibrated by benchmarking the peak sensitivity in the central slice of the tomograph to a measurement with the ^{22}Na at the center of the FOV. The line source measurement gives values for absolute sensitivity that differ from the point source measurement by less than 1% using the microPET P4 (*13*).

Scatter Fraction, Count Losses, and Random Coincidence Measurements. The NU 4-2008 methodology for the counting-rate test closely follows the NU 2-2001 methodology for clinical PET, in which a line source filled with ^{18}F is inserted along the

FIGURE 1. Plot of FWHM and FWTM spatial resolution for LabPET 8 system for 2 energy windows.
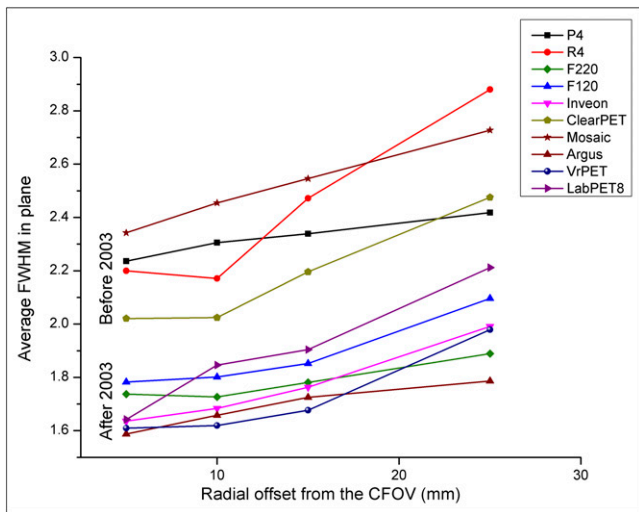
FIGURE 2. Comparison of spatial resolution of first-generation animal PET systems with later-generation systems.

length of a high-density polyethylene cylinder. In the NU 4-2008 standard, 3 phantom sizes are used: a mouse phantom (70 mm long, 25 mm diameter [Ø]), a rat phantom (150 mm long, 50 mm Ø), and a monkey phantom (400 mm long, 100 mm Ø).

For each system and phantom tested, we report the peak noise-effective counting rate (NECR), the activity at which peak NECR occurs, and the low-counting-rate scatter fraction. In addition, we report NECR at 3.7 MBq for the mouse phantom and 10 MBq for the rat phantom, as these activity levels correspond to values that are often encountered in routine imaging.

Image Quality and Accuracy of Attenuation and Scatter Corrections. The NEMA NU 4-2008 method uses a fillable phantom (66 mm long, 33.5 mm Ø) for the image quality test. For the uniform cylinder region, we report the maximum and minimum values as ratios, with the mean value and the SD of the pixel values as a percentage of the mean. For the cold cylinder regions, we report the spillover ratio. For the hot rod region, we report the recovery coefficients. On each system, the phantom was imaged for 20 min with an activity level of 3.7 MBq of ^{18}F. All available corrections were applied to the data. All systems had corrections for normalization, dead time, and randoms, but not all systems had corrections available for scatter and attenuation at the time of testing.

RESULTS

Spatial Resolution

Table 2 lists the FWHM and FWTM spatial resolution for each system. Spatial resolution was generally better at the ¼-axial-offset position than at the center, particularly for axial resolution, because of the more oblique lines of response used by the central position than by the ¼-axial-offset position. This effect was particularly noticeable for systems with long axial FOVs.

The NU 4-2008 standard requires a filtered backprojection (FBP) algorithm to reconstruct the point source data. This requirement is problematic for system designs that have irregular crystal spacing in the azimuthal and axial directions, such as the LabPET systems. Systems of this design do not typically use FBP algorithms because of the degradation of resolution and artifacts introduced by the interpolation and rebinning of measured data onto projections with regular spacing. Such artifacts, which make the resolution unstable across the FOV, are apparent in the FWHM resolution data measured for the LabPET 8, in which the FWHM radial resolution increases from approximately 1.6 mm at a 5-mm radial offset to 1.9 mm at a 10-mm radial offset. Consequently, the results for the LabPET axial resolution are the intrinsic resolution obtained by finely stepping the point source through the axial direction. Because these results will be the same for the LabPET 8 and the LabPET 12, only LabPET 8 results are presented in Tables 2 and 3. Vendors of most preclinical PET systems offer iterative 3-dimensional reconstruction algorithms, many of which include spatially variant models of system response. As a result, it may be unusual to use FBP reconstruction in routine preclinical PET. For these systems, the FBP algorithm requirement in the NU 4-2008 standard leads to situations never realized in routine imaging situations. This limitation of the NU 4-2008 spatial resolution test does not presently have a practical solution.

Table 3 compares the system crystal size with the effective transaxial FWHM resolution at the 5-mm-offset radial

TABLE 4
Sensitivity over Central 7 cm of Axial FOV (Mouse Sensitivity), Complete Axial FOV (Total Sensitivity), and Peak Detection Efficiency for Each System Tested

System	Energy window (keV)	Axial length (cm)	Mouse sensitivity (%)	Total sensitivity (%)	Peak detection efficiency (%)
microPET P4	350–650	7.8	0.67	0.61	1.19
microPET R4	350–650	7.8	1.19	1.10	2.06
microPET Focus 220	350–650	7.6	1.26	1.18	2.28
microPET Focus 120	350–650	7.6	1.98	1.82	3.42
Inveon	350–625	12.7	4.0	2.8	6.72
ClearPET	250–650	11.0	2.32	1.87	3.03
Mosaic HP	385–665	11.9	2.43	1.77	2.83
Argus	250–700	4.8			4.32
VrPET	100–700	4.56	1.09	1.09	2.22
LabPET 8	250–650	7.5	1.45	1.42	2.36
LabPET 12	250–650	11.25	3.6	2.74	5.4

TABLE 5
Summary of Counting-Rate Test Results for Mouse and Rat Phantoms

System	Energy window (keV)	Timing window (ns)	Randoms correction	Mouse phantom				Rat phantom			
				Peak NECR (kcps)	Activity* (MBq)	NECR at 3.7 MBq (kcps)	Scatter fraction (%)	Peak NECR (kcps)	Activity* (MBq)	NECR at 10 MBq (kcps)	Scatter fraction (%)
microPET P4	350–650	6	Calculated	>601†	>174†	22.1	5.2	173	254	19.2	16.7
microPET R4	350–650	6	Calculated	618	156	37.2	9.3	164	137	30.5	22.2
microPET Focus 120	350–650	6	Calculated	897	103	66.5	5.6	267	129	50.9	20.3
microPET Focus 220	250–700	6	Calculated	>763†	>89†	47.3	7.2	359	162	51.8	19.3
Inveon	350–625	3.4	Calculated	1670	131	129.0	7.8	592	110	137.8	17.2
ClearPET	250–650	12	Calculated	73	18	29.3	31.0				
Mosaic HP	385–665	7	Measured	555	92	59.6	5.4	244	87	65.2	12.7
Argus	250–700	7	Calculated	117	50	18.7	21.0	40	41	20.4	34.4
VrPET	100–700	3.8	Calculated	74	22		11.5	31	34		23.3
LabPET 8	250–650	20	Calculated	279	82	23.5	15.6	94	91	19.4	29.5
LabPET 12	250–650	20	Calculated	362	81	38.9	16.0	156	83	40.5	29.3

*Activity at which peak NECR occurs.
†Peak value not reached because of insufficient activity at start of scan.

position. Effective transaxial FWHM resolution is calculated as a geometric mean according to

$$\text{Effective transaxial FWHM resolution} = \sqrt{\left(\frac{\text{FWHM}_{\text{rad., center}} + \text{FWHM}_{\text{rad., 1/4 offset}}}{2}\right)\left(\frac{\text{FWHM}_{\text{tan., center}} + \text{FWHM}_{\text{tan., 1/4 offset}}}{2}\right)}$$

where $\text{FWHM}_{\text{rad., center}}$ and $\text{FWHM}_{\text{rad., 1/4 offset}}$ are the FWHM radial resolutions at the center and ¼-offset axial position, respectively, and $\text{FWHM}_{\text{tan., center}}$ and $\text{FWHM}_{\text{tan., 1/4 offset}}$ are the corresponding tangential resolutions. For every system except the LabPET 8, the FWHM spatial resolution is greater than the crystal size. This characteristic reflects the unique acquisition architecture of LabPET scanners, which allow individual crystal readout, as compared with systems that use signal multiplexing and light sharing at the detector level. For the LabPET 8 system, spatial resolution was measured at 2 energy windows, 250–650 keV and 375–650 keV. Figure 1 shows effective transaxial FWHM resolution and effective transaxial FWTM for these energy windows. The FWHM shows no dependence on energy window, whereas the FWTM shows a significant degradation with the wider energy window, indicating that more scatter is accepted. This effect is not unique to the LabPET 8 and is also apparent with a wider energy window on other systems. However, there may be differences in how the FWTM changes with energy window between the LabPET 8 system and block detector–based systems because of the relative contributions of intercrystal scatter and gantry scatter components to the FWTM broadening.

Figure 2 shows the values of effective transaxial FWHM resolution for each system plotted against radial offset position. The plots can clearly be grouped into 2 families of systems: those manufactured before 2003, that is, first-generation commercial systems, and those manufactured after 2003, that is, second-generation systems.

Sensitivity

Table 4 shows the values of mouse sensitivity and total sensitivity for each system. As expected, the largest factor affecting detection efficiency is the solid-angle coverage of the detector ring, with higher values for long-axial-FOV and small-ring-diameter systems.

Scatter Fraction, Count Losses, and Random Coincidence Measurements

Table 5 summarizes the results of the counting-rate test for mouse- and rat-sized phantoms. In 2 cases, the peak NECR values were not reached because of limited starting activity. This problem is due to the low volume of the line source and is more likely to occur for the mouse phantom. As expected, the scatter fraction was generally lower for larger-ring systems and for narrower energy windows. The
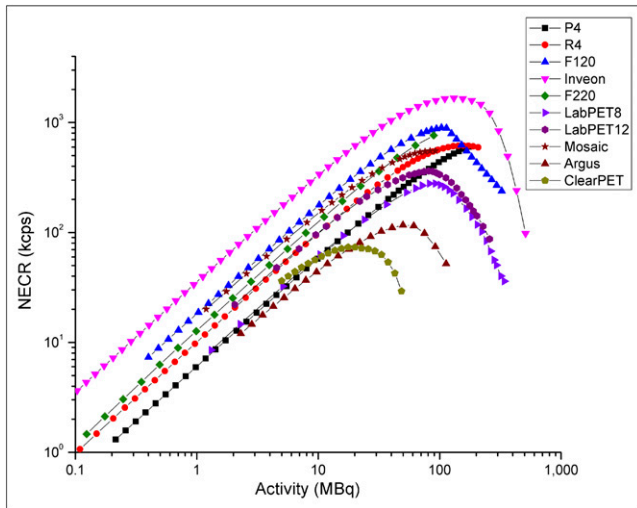
FIGURE 3. Plot of NECR vs. activity for mouse-sized phantom.
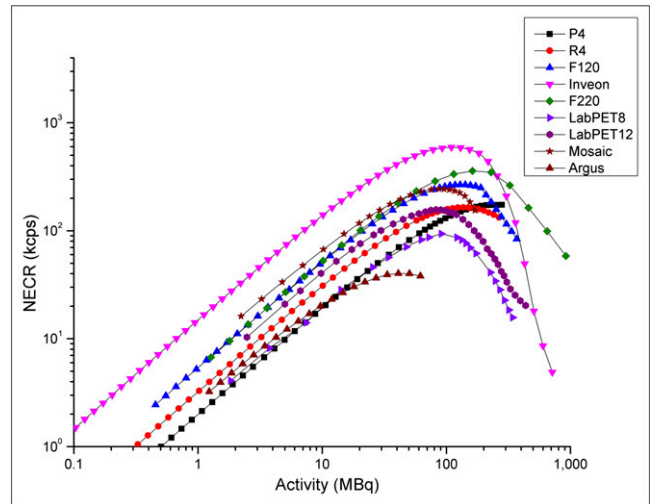
FIGURE 4. Plot of NECR vs. activity for rat-sized phantom.

exception was the VrPET system, which used an energy window of 100–700 keV yet had a scatter fraction lower than that of any system using a 250- to 750-keV window. The most likely reason is that the VrPET is a partial-ring system and thus has less scatter from gantry materials. This explanation is consistent with the work of Yang and Cherry (*14*), who showed that for mouse-sized phantoms imaged in the microPET II, the dominant source of scatter was the gantry. The scatter fraction was lowest for systems that use conventional single-layer block detector designs, such as the Siemens family of systems, or the pixelated Anger logic approach of the Mosaic HP. The highest observed scatter fractions were for the 2 dual-layer systems, with the ClearPET having a scatter fraction of 31% and the Argus having a scatter fraction of 21% for the mouse phantom. It is not clear whether the increased scatter fraction is due to event mispositioning in the block, high levels of gantry scatter events, or the effect of using scintillators with lower photofractions (germanium oxyorthosilicate for the Argus, lutetium yttrium aluminum perovskite for the ClearPET). The ClearPET and the Argus were the 2 systems with the smallest ring diameters, which likely has an effect on the amount of gantry scatter. The LabPET systems, with their individual crystal readout design, have scatter fractions between these 2 extremes. It is believed that the higher scatter fraction measured for the LabPET systems is due to increased gantry scatter from the Kovar (Carpenter Technology Corp.) packages surrounding the detector modules.

The peak NECR value and activity level at which it occurs represent a complex interplay between system design factors. Compared with clinical PET systems, preclinical systems from different vendors have a much wider variation in design, making it difficult to directly use NECR for comparing systems. It is, however, instructive to compare a few systems directly to understand the effects of system differences on the counting-rate results. When the microPET R4 is compared with the microPET Focus 120, the improved sensitivity of the microPET Focus 120 results in significantly higher NECR values at lower activity levels than for the microPET R4. The effects of an extended axial FOV can be seen by comparing the LabPET 8 with the LabPET 12. For the mouse phantom, the peak NECR increased from 279 kcps for the LabPET 8 to 362 kcps for the LabPET 12, with only a 1 MBq change in the activity at which peak NECR occurs. The effects of system ring diameter can be seen by comparing the microPET P4 and microPET R4 results for the rat phantom. The peak NECR for the 2 systems was similar; however, the activity at which peak NECR occurs is larger by nearly a factor of 2 for the microPET P4. The Inveon system had the highest values of peak NECR for the mouse and rat phantoms. A key reason for these high NECR values is the minimal block dead time due to the Quicksilver processing electronics (*15*), which allow minimal pulse shaping before digitization with 100-MHz analog-to-digital converters and a timing window of 3.4 ns.

TABLE 6
Summary of Counting-Rate Test Results for Monkey-Sized Phantom

	Energy window (keV)	Timing window (ns)	Peak NECR (kcps)	Activity* (MBq)	Scatter fraction (%)
microPET P4	350–650	6	32.8	276	35.5% at 6.5 MBq
microPET Focus 220	250–700	6	60.0	183.6	46.6% at 2.9 MBq

*Activity at which peak NECR occurs.

TABLE 7
Summary of Results from Image-Quality Phantom

| System | Energy window (keV) | Reconstruction algorithm | Attenuation/ scatter correction | Uniform region | | | Recovery coefficients | | | | | Spillover ratios | |
				% SD	Ratio maximum/ mean	Ratio minimum/ mean	1 mm	2 mm	3 mm	4 mm	5 mm	Water	Air
microPET P4	350–650	FORE + 2D FBP	Yes/yes	5.2	1.20	0.81	0.11	0.37	0.60	0.77	0.86	4.9%	4.0%
microPET Focus 220	250–700	FORE + 2D FBP	Yes/yes	6.8	1.27	0.71	0.15	0.41	0.63	0.74	0.86	1.2%	4.1%
microPET R4	350–650	FORE + 2D FBP	Yes/no	4.5	1.14	0.80	0.14	0.35	0.60	0.79	0.87	6.2%	4.6%
microPET Focus 120	350–650	FORE + 2D FBP	Yes/yes	6.0	1.25	0.74	0.15	0.48	0.75	0.86	0.93	1.8%	−0.3%
Inveon	350–625	FORE + 2D FBP	Yes/yes	5.3	1.18	0.80	0.17	0.48	0.72	0.84	0.93	1.7%	−0.6%
ClearPET	250–650	3D OSEM	No/no	10.9	1.43	0.58	0.11	0.21	0.42	0.73	0.90	36.9%	26.7%
Mosaic HP	385–665	3D RAMLA	Yes/yes	5.1	1.19	0.80	0.16	0.36	0.56	0.70	0.84	6.3%	2.7%
Argus	250–700	3D OSEM	Yes/yes	6.0	1.23	0.81	0.27	0.65	0.93	0.95	0.97	15.0%	13.0%
VrPET	100–700	3D OSEM	No/no	15.4	1.75	0.47	0.22	0.62	0.72	0.75	0.75	9.3%	8.5%
LabPET 8	250–650	2D MLEM	No/no	6.0	1.24	0.76	0.19	0.78	0.97	1.00	1.02	24.4%	13.7%
LabPET 12	250–650	2D MLEM	No/no	7.9	1.29	0.73	0.24	0.77	0.92	0.93	0.97	25.6%	16.0%

FORE = Fourier rebinning; 2D = 2-dimensional; OSEM = ordered-subsets expectation maximization; RAMLA = row-action maximum-likelihood algorithm; MLEM = maximum-likelihood expectation maximization.

Figures 3 and 4 show the NECR counting-rate curves for the systems for the mouse and rat phantoms, respectively. The general shape of the curves is similar for all systems, with an extended linear range of the NECR-versus-activity level below the peak NECR value. For all systems, this linear range extends at least up to 10 MBq, which is sufficient for performing most imaging studies in rodents.

Table 6 summarizes the results of the counting-rate test for the monkey-sized phantom. Data on this phantom were acquired using only the 2 systems with the largest ring diameters, the microPET P4 and microPET Focus 220. For this test, the microPET P4 used an energy window of 350–650 keV and the microPET Focus 220 used an energy window of 250–700 keV. This wider energy window results in a significant increase in the scatter fraction from 35.5% for the microPET P4 to 46.6% for the microPET Focus 220.

Image Quality and Accuracy of Attenuation and Scatter Corrections

Table 7 summarizes the results from the image-quality phantom for each system tested. The results of the image-quality test are highly dependent on the reconstruction algorithm and the corrections applied. This point is illustrated in Figure 5 for recovery coefficients measured for the microPET P4 system for data reconstructed using 5 different methods: Fourier rebinning followed by 2-dimensional FBP, Fourier rebinning followed by 2-dimensional ordered-subsets expectation maximization, maximum a posteriori (MAP) with $\beta = 0.1$, maximum a posteriori with $\beta = 0.447$, and 3-dimensional reprojection. Recovery coefficients greater than 1 are measured with MAP reconstructions, likely caused by a combination of using an iterative algorithm to reconstruct pointlike objects in a region that does not have background activity and using a recovery coefficient based on a single-pixel measurement from the average image created by summing a 10-mm axial region.

The variability in the results from the microPET P4 system makes it difficult to compare the results from the image-quality phantom across systems from different manufacturers. In general, systems with lower scatter fractions in the mouse phantom counting-rate test had lower spillover ratios regardless of whether corrections were applied. As discussed by Yang and Cherry (14), this observation may reflect the fact that more scatter in preclinical PET originates from sources other than the object being imaged and that scatter correction methods assume scatter originates in the object. For systems that did not use either scatter or attenuation correction in the image-quality test, the spillover ratio for the water compartment correlated with the scatter fraction measured in the counting-rate test using the mouse-sized phantom, which is similar in size to the image-quality phantom (Fig. 6).

Figure 7 shows transverse images through the cold compartment region and coronal images through the 5-mm hot rod for a selection of systems tested. The higher spillover ratios in the uncorrected images from the ClearPET and LabPET 12 can be seen as an increase in apparent activity in the cold compartments.
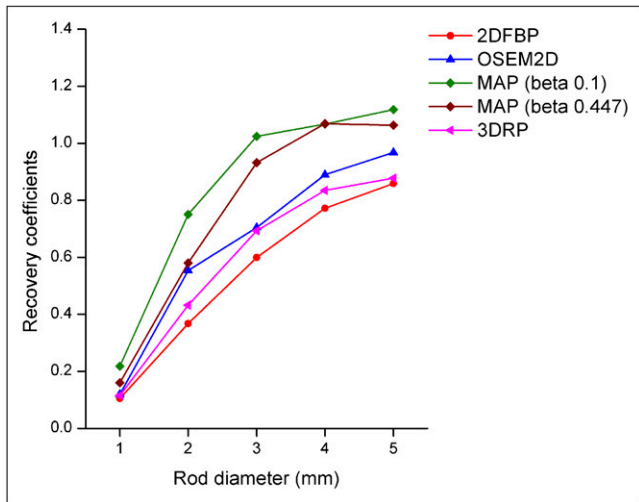
FIGURE 5. Plots of recovery coefficient for the microPET P4 system for 5 reconstruction algorithms. MAP = maximum a posteriori; OSEM2D = 2-dimensional ordered-subsets expectation maximization; 3DRP = 3-dimensional reprojection; 2DFBP = 2-dimensional filtered backprojection.

**FIGURE 6.** Plot of spillover ratio in water compartment vs. scatter fraction measured in counting-rate test using mouse-sized phantom. Line has slope of 1.

DISCUSSION

The data show a steady improvement in system performance for newer systems as compared with first-generation systems. This trend is most clearly seen in the improvement in spatial resolution for systems produced after 2003, with all newer systems having an average in-plane FWHM resolution of better than 2 mm over the central 30-mm diameter of the FOV. Similarly, there has been a steady improvement in system sensitivity, driven largely by the extended axial coverage of newer systems.

Several observations were made about the NEMA NU 4-2008 standard. The spatial resolution test requires FBP reconstruction. This test therefore is favorable for systems using a ring geometry and limited axial extent and unfavorable for systems with unconventional geometries. It is therefore possible that a system with poor measured spatial resolution may produce images of exceptional resolution and quality when reconstructed with an iterative algorithm. The sensitivity measurement, performed by stepping a ^{22}Na point source through the axial FOV of the system, requires a large number of repetitive measurements and is time-consuming. We suggest that a line source measurement with the peak sensitivity at the center of the FOV benchmarked to a measurement made with the ^{22}Na source at the center of the FOV can be a rapid and accurate alternative. The counting-rate tests are useful for specifying the range of activities that are suitable for use in the system for various animal sizes. However, in the NU 4 test, there is no requirement to evaluate the ability of the PET system to form an image at activity levels other than the 3.7 MBq used in the image quality phantom. This is a distinct difference from the NU 2 standard for clinical PET in which the counting-rate data are reconstructed to determine up to what activity level
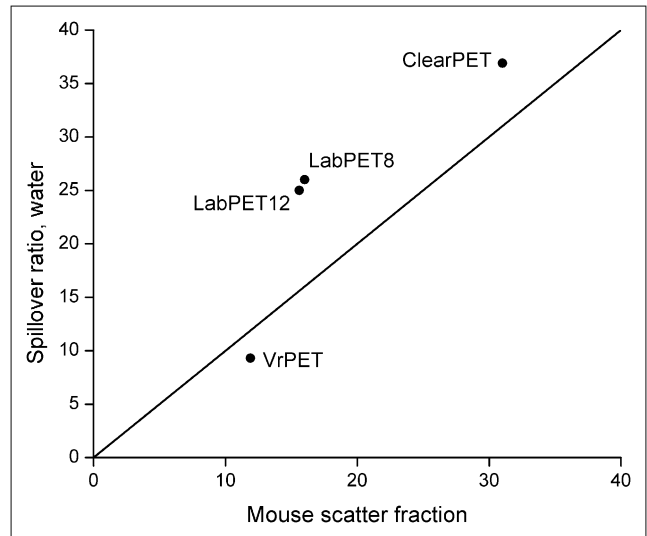
the system dead-time correction is functioning accurately and the image is free from pileup artifacts. We suggest that the counting-rate test be modified to include a requirement for reconstructing the counting-rate data and for analyzing
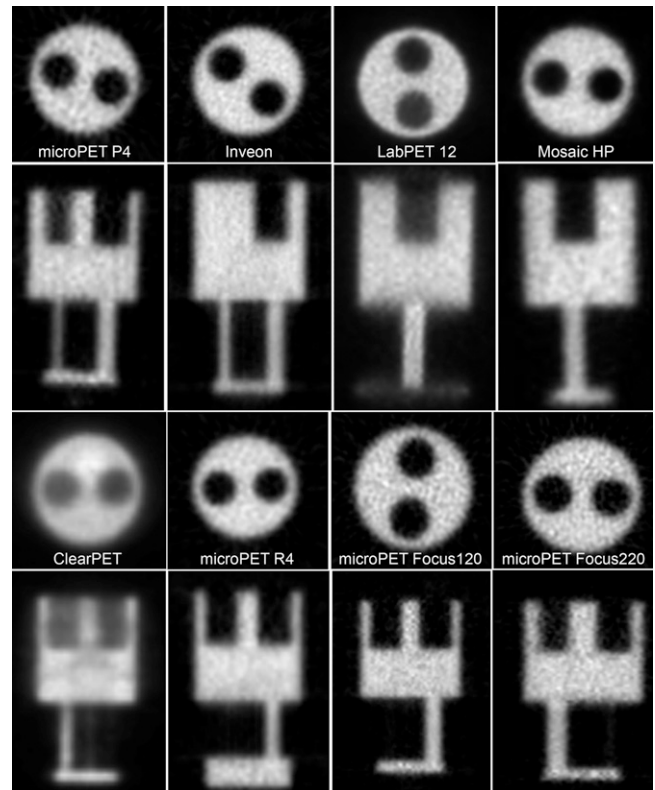
FIGURE 7. Image-quality phantom images. Intensity scale of each image is set so that minimum value is 0 and maximum value is 1.25 times mean value of uniform cylinder region.

the resultant images to determine the quantitative accuracy of the system as a function of counting rate.

When comparing the NEMA test results from preclinical PET systems from different manufacturers, it is important to remember that beyond enabling comparison of PET systems, the additional purpose of the NU 4-2008 standard is to provide a standard set of tests and methods for manufacturers to specify the performance of their imaging systems and for customers to perform acceptance testing and long-term monitoring. Therefore, the number of tests suggested is purposefully limited so that the data can be acquired in a timely manner. However, additional metrics of performance can be evaluated to provide valuable information. For example, we can mention measurement of spatial resolution at low and high counting rates.

CONCLUSION

In this work we have collected and presented NEMA NU 4-2008 performance data for 11 preclinical PET systems commercially manufactured since 2000. Their performance over the range of tests reflects the unique design attributes of each system, the settings at which it is operated, and the manner in which the data are handled and reconstructed. In general, there is a much wider variation in the design of preclinical PET systems than in clinical PET systems because of the different approaches implemented to push the limits of resolution and sensitivity. This variation in system design makes direct comparisons between systems from different vendors difficult since one must also consider the suitability of the PET system for the specific imaging task at hand when considering the results from NEMA testing. The data show a steady improvement in system performance for newer systems as compared with first-generation systems, with notable improvements in spatial resolution and sensitivity.

REFERENCES

1. Cutler PD, Cherry SR, Hoffman EJ, Digby WM, Phelps ME. Design features and performance of a PET system for animal research. *J Nucl Med.* 1992;33:595–604.
2. Watanabe M, Uchida H, Okada H, et al. A high resolution PET for animal studies. *IEEE Trans Med Imaging.* 1992;11:577–580.
3. Bloomfield PM, Rajeswaran S, Spinks TJ, et al. The design and physical characteristics of a small animal positron emission tomography. *Phys Med Biol.* 1995;40:1105–1126.
4. Lecomte R, Cadorette J, Rodrigue S, et al. Initial results from the Sherbrooke avalanche photodiode positron tomograph. *IEEE Trans Nucl Sci.* 1996;43:1952-1957.
5. Watanabe M, Okada H, Shimizu K, et al. A high resolution animal PET scanner using compact PS-PMT detectors. *IEEE Trans Nucl Sci.* 1997;44:1277–1282.
6. Cherry SR, Shao Y, Silverman RW, et al. MicroPET: a high resolution PET scanner for imaging small animals. *IEEE Trans Nucl Sci.* 1997;44:1161–1166.
7. Bruyndonckx P, Liu XA, Tavernier S, Zhang SP. Performance study of a 3D small animal PET scanner based on BaF2 crystals and a photo sensitive wire chamber. *Nucl Instrum Methods Phys Res. Sect A.* 1997;392:407–413.
8. Pichler B, Boning G, Lorenz E, et al. Studies with a prototype high resolution PET scanner based on LSO-APD modules. *IEEE Trans Nucl Sci.* 1998;45:1298–1302.
9. Jeavons AP, Chandler RA, Dettmar CAR. 3D HIDAC-PET camera with sub-millimetre resolution for imaging small animals. *IEEE Trans Nucl Sci.* 1999;46:468–473.
10. National Electrical Manufacturers Association. *NEMA Standard Publication NU 4-2008: Performance Measurements of Small Animal Positron Emission Tomographs.* Rosslyn, VA: National Electrical Manufacturers Association; 2008.
11. National Electrical Manufacturers Association. *NEMA Standards Publication NU 2-1994: Performance Measurements of Positron Emission Tomographs.* Rosslyn, VA: National Electrical Manufacturers Association; 1994.
12. National Electrical Manufacturers Association. *NEMA Standards Publication NU 2-2001: Performance Measurements of Positron Emission Tomographs.* Rosslyn, VA: National Electrical Manufacturers Association; 2001.
13. Elhami E, Boulding J, Goertzen AL. Variations on the NEMA NU4-2008 testing procedures and effect on the performance measurement results. In: *2011 IEEE Nuclear Science Symposium Conference Record.* Piscataway, NJ: IEEE; 2011:3244–3246.
14. Yang Y, Cherry SR. Observations regarding scatter fraction and NEC measurements for small animal PET. *IEEE Trans Nucl Sci.* 2006;53:127–132.
15. McFarland AR, Siegel S, Newport DF, Mintzer R, Atkins B, Lenox M. Continuously sampled digital pulse processing for Inveon small animal PET scanner. In: *2007 IEEE Nuclear Science Symposium Conference Record.* Piscataway, NJ: IEEE; 2007:4262–4265.
16. Tai C, Chatziioannou A, Siegel S, et al. Performance evaluation of the microPET P4: a PET system dedicated to animal imaging. *Phys Med Biol.* 2001;46:1845–1862.
17. Knoess C, Siegel S, Smith A, et al. Performance evaluation of the microPET R4 PET scanner for rodents. *Eur J Nucl Med Mol Imaging.* 2003;30:737–747.
18. Tai YC, Ruangma A, Rowland D, et al. Performance evaluation of the microPET Focus: a third-generation microPET scanner dedicated to animal imaging. *J Nucl Med.* 2005;46:455–463.
19. Kim JS, Lee JS, Im KC, et al. Performance measurement of the microPET Focus 120 scanner. *J Nucl Med.* 2007;48:1527–1535.
20. Laforest R, Longford D, Siegel S, Newport DF, Yap J. Performance evaluation of the microPET Focus-F120. *IEEE Trans Nucl Sci.* 2007;54:42–49.
21. Bao Q, Newport D, Chen M, Stout DB, Chatziioannou AF. Performance evaluation of the Inveon dedicated PET preclinical tomograph based on the NEMA NU-4 standards. *J Nucl Med.* 2009;50:401–408.
22. Kemp BJ, Hruska CB, McFarland AR, Lenox MW, Lowe VJ. NEMA NU 2-2007 performance measurements of the Siemens Inveon (TM) preclinical small animal PET system. *Phys Med Biol.* 2009;54:2359–2376.
23. Visser EP, Disselhorst JA, Brom M, et al. Spatial resolution and sensitivity of the Inveon small-animal PET scanner. *J Nucl Med.* 2009;50:139–147.
24. Mintzer RA, Siegel SB. Design and performance of a new pixelated-LSO/PSPMT gamma-ray detector for high resolution PET imaging. In: *2007 IEEE Nuclear Science Symposium Conference Record.* Piscataway, NJ: IEEE; 2007:3418–3422.
25. Surti S, Karp JS, Perkins AE, et al. Imaging performance of A-PET: a small animal PET camera. *IEEE Trans Med Imaging.* 2005;24:844–852.
26. Surti S, Karp JS, Perkins AE, Freifelder R, Muehllehner G. Design evaluation of A-PET: a high sensitivity animal PET camera. *IEEE Trans Nucl Sci.* 2003;50:1357–1363.
27. Ziemons K, Auffray E, Barbier R, et al. The ClearPET(TM) project: development of a 2nd generation high-performance small animal PET scanner. *Nucl Instrum Methods Phys Res Sect A.* 2005;537:307–311.
28. Wang Y, Seidel J, Tsui BMW, Vaquero JJ, Pomper MG. Performance evaluation of the GE Healthcare eXplore VISTA dual-ring small-animal PET scanner. *J Nucl Med.* 2006;47:1891–1900.
29. Lage E, Vaquero JJ, Sisniega A, et al. Design and performance evaluation of a coplanar multimodality scanner for rodent imaging. *Phys Med Biol.* 2009;54:5427–5441.
30. Bergeron M, Cadorette J, Beaudoin JF, et al. Performance evaluation of the LabPET APD-based digital PET scanner. *IEEE Trans Nucl Sci.* 2009;56:10–16.
31. Fontaine R, Belanger F, Viscogliosi N, et al. The hardware and signal processing architecture of LabPET (TM), a small animal APD-based digital PET scanner. *IEEE Trans Nucl Sci.* 2009;56:3–9.
32. Bergeron M, Cadorette J, Bureau-Oxton C, et al. Performance evaluation of the LabPET12, a large axial FOV APD-based digital PET scanner. In: *2009 IEEE Nuclear Science Symposium Conference Record.* Piscataway, NJ: IEEE; 2009: 4017–4021.