

Original article

팬텀을 활용한 PET/CT 장비 별 영상 비교 평가

정무진, 함준철, 최용훈, 반영각, 임한상, 김재삼
연세의료원 세브란스병원 핵의학과

Image Comparative Evaluation by PET/CT Equipment Using Phantom

Moo-Jin Jeong, Jun-Chul Ham, Yong-Hoon Choi, Young-Kag Bahn, Han-Sang Lim, Jae-Sam Kim

Department of Nuclear Medicine, Severance Hospital, Yonsei University Health System, Seoul, 03722, Republic of Korea

Corresponding Author : Moo-Jin Jeong, Department of Nuclear Medicine, Severance Hospital, Yonsei University Health System, Seoul, 03722, Republic of Korea, Tel :+82-2-2228-6065, E-mail : moojin0217@naver.com

ABSTRACT

Purpose: This study aims to identify SUV, SNR, spatial resolution, and axial uniformity under the same reconstruction conditions and to find out the differences between equipment models. **Materials and Methods:** The equipment was GE's Discovery 600, 710, IQ, MI(GE Healthcare, USA), and the Phantom used ACR(American College of Radiology) Flangeless Esser Phantom and PET/SPECT Performance Phantom. The PET/SPECT Performance Phantom injected 18F-FDG at a concentration of 3.8 kBq/mL, and the ACR Flangeless Esser Phantom made the conditions for Hot Spot and Background activity for 4 : 1. Image evaluation was compared and evaluated for SUV, SNR, spatial resolution, and axial uniformity with the same reconstruction that added SharpIR of VPHD. **Results:** The SUV_{max} showed a difference up to 4.6% with an average of 2.71, 2.35, 1.89, and 1.43 from Hot Spot 1 to 4, and the SUV_{mean} showed a difference up to 4.7% with 2.06, 1.75, 1.49, and 1.27. There was a difference up to 5% between equipment, and there was no significant difference between both SUV_{max} and SUV_{mean}. SNR showed a difference up to 0.04 with an average of 0.37, 0.26, 0.18, and 0.11. FWHM showed a difference up to 0.27. Lastly, COV of axial uniformity was up to 0.018. **Conclusion:** SUV showed differences within 5% between equipment and showed no significant difference. This is considered to be used as basic data that can be used for the development and replacement of equipment because it has the advantage of being able to observe with a large number of equipment.

Key Words Phantom, VPHDs, SUV

서론

양전자방출단층촬영(Positron Emission Tomography, PET)은 각종 생화학적인 물질의 생체 내 분포를 영상화하는 기법으로 인체 내의 생리적 지표들을 정량적으로 측정할 수 있어 생화학 또는 병리 현상의 규명에 유용하다[1]. 이는 생체 기능을 평가하는 데 가장 적합하고 종양 분야에서의 활용도가 매우 높아 암의 진단과 치료계획, 추적관찰 등에 많은 검사가 진행되고 있으며, 지속적인 발전에 따라 그에 따른 수요가 증가하고 있다. 예후 판정을 위한 PET-CT 추적관찰 시 SNMMI(Society of Nuclear Medicine and Molecular Imaging)과

EANM(European Association of Nuclear Medicine)에서는 동일 장비에서의 검사를 권고하고 있으나 치료계획 변경이나 검사 일정 이동 등 한계점이 존재한다. SUV는 동일 제조회사이더라도 장비의 재원과 특성에 따른 차이가 발생하게 되며, 식품의약품안전평가원 PET/CT 안전성 및 성능평가 시험방법 가이드라인에서는 장비 간 SUV차이를 10% 이내, 장비 제조사에서는 5% 이내의 차이를 권고하고 있다[2]. 만일 타 장비에서의 검사 간에도 5% 이내의 SUV 차이를 보인다면, 장비 호환에 따른 교차검사가 가능할 것으로 예상되며 장비에 구매받지 않는 추적관찰이 가능할 것이다. 현재 본원에서는 추적관찰 시 동일 장비에서의 검사가 이루어지지 않고 있으며, 정밀 분석 시 동일 장비에서의 검사가 이루어지고 있다. 본 논문에서는 SUV 차이와 더불어 신호 대 잡음 비(Signal Noise Ratio, SNR), 공간 분해능(spatial resolution) 그리고 축 방향 균일도(axial uniformity)의 변동계수(Coefficient of Variation, COV)를 확인하여 장비 재원에 따라 발생할 수 있는 영상지표의 차이를 확인하고자 한다.

실험재료 및 방법

1. 실험대상

실험을 위한 장비로는 GE Discovery 600(D600), Discovery 710(D710), Discovery IQ(DIQ), Discovery MI(DMI)(General Electric Healthcare, USA)로 하였으며, 각 장비의 재원은 다음과 같다(Fig. 1, Table. 1). 영상 평가를 위한 팬텀은 ACR(American College of Radiology) Flangeless Esser Phantom, PET/SPECT Performance Phantom으로 하였다(Fig. 2). 최대 표준 섭취계수(Standard Uptake Value maximum, SUV_{max})와 평균 표준 섭취 계수(Standard Uptake Value mean, SUV_{mean}), SNR은 ACR Flangeless Esser Phantom을 이용하여, spatial resolution과 axial uniformity는 PET/SPECT Performance Phantom을 이용하여 비교 평가하였다.



Fig. 1. The equipment used in this study is GE Discovery (General Electric Healthcare, USA).

Table 1. PET specifications of each PET/CT in this study.

	D600	D710	DIQ	DMI
Scintillator (Detector rings)	BGO(24)	Lutetium based(24)	BGLS(24)	Lutetium based(36)
PM tube	1024	1024	720	544

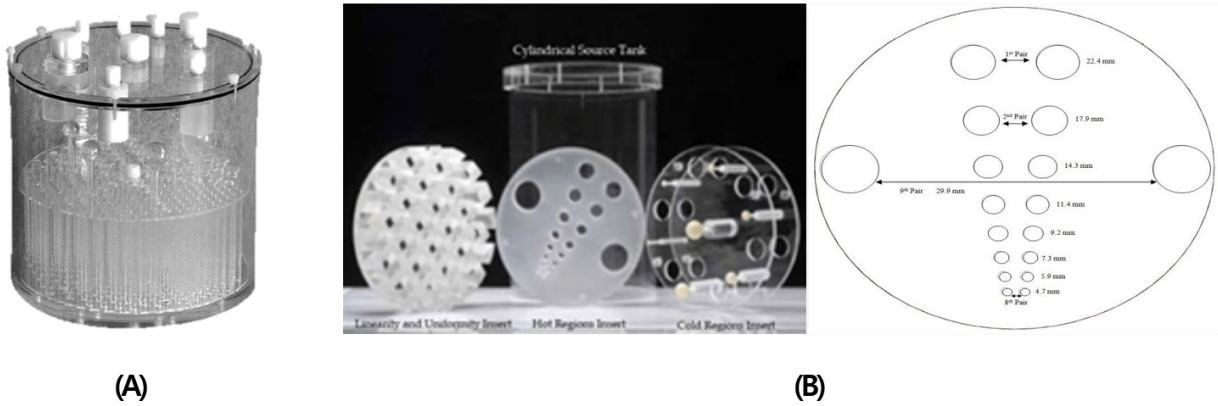


Fig. 2. These figures show ACR Flangeless Esser Phantom(A) and PET/SPECT Performance Phantom(B) using this study.

2. 실험방법

ACR Flangeless Esser Phantom은 ACR Phantom Study Guideline 조건에 따라 열소(Hot Spot)와 배후방사능(background)의 방사능 농도 비를 4:1로 하였다[3]. background에 18.5 MBq를 주입하여 교반하였고, Hot Spot은 1 L의 증류수에 7.78 MBq를 섞은 후 Hot Spot 1, 2, 3, 4에 주입하였다. PET/SPECT Performance Phantom은 3.8 kBq/mL의 방사능 농도로 주입 후 10분 교반하였다. 촬영은 1 bed 당 3 분 촬영조건으로 2 bed scan 하였다. 재구성은 동일한 조건에서의 실험을 위해 VUE Point HD(VPHD)를 사용하였으며, SharpIR을 추가하여 재구성하였다. scan parameter는 subsets은 D600, D710, DMI 장비에서 16으로 설정하였고 DIQ 장비에서 12로 설정하였다. cutoff는 5.0 mm, iterations은 2로 설정하였다(Table 2). VPHD +SharpIR(VPHDs)의 reconstruction loop는 다음과 같다(Fig. 3).

Table 2. Scan parameters of PET reconstruction.

Reconstruction type	Subsets	Cutoff	Iterations
VPHDs	16(12)	5.0	2

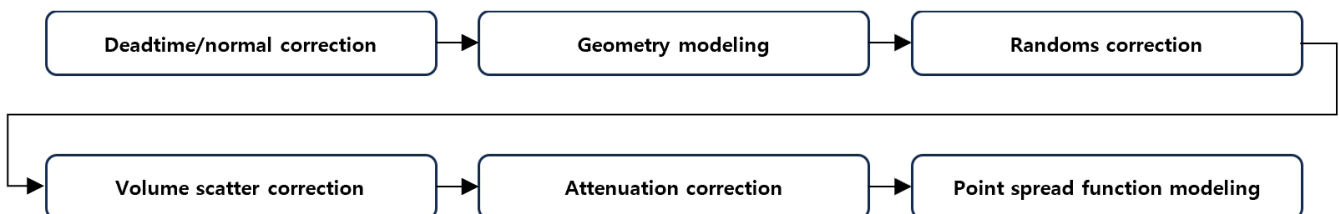


Fig. 3. The process of the reconstruction method used in this study is described in this way.

1. 영상분석

획득된 ACR Flangeless Esser Phantom 영상의 Hot Spot 별 25 mm의 관심 영역(Region of Interest, ROI)을 설정하여 SUV를 측정하였다. ROI 설정은 manual으로 하였다. 관계식은 다음과 같다(Eq. 1). 장비 간 비교를 위해 측정된 SUV_{max} 와 SUV_{mean} 는 1 : 1로 비교 분석하였다(Fig. 4). 만일 비교 값이 5% 이내의 차이를 보인다면 추적관찰을 위한 장비 간 교차검사가 가능할 것이라 가정하였다. SNR은 측정된 SUV와 background의 값의 관계를 통해 분석하였다[4]. 관계식은 다음과 같다(Eq. 2). spatial resolution은 반치폭(Full Width at Half Maximum, FWHM)을 통해 확인하였으며, axial uniformity는 ROI를 그려 측정된 background의 평균과 COV를 비교하였다. COV는 분산 정도를 나타내는 지표이다(Eq. 3). 집단이 같은 비교군에서는 표준편차를 이용하여 분산 정도를 평가하지만, 본 연구에서의 각 장비 간 균일도의 평균이 모두 다르기 때문에 COV를 이용하였다.

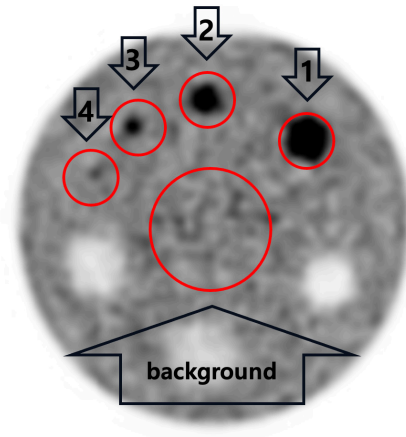


Fig. 4. We set ROI from phantom images obtained from PET/CT equipment.

$$SUV = \frac{\text{(Concentration of Tissue}(mCi/g))}{\text{F-18 FDG injected Dose}(mCi)/\text{Patient's weight}(g)} \quad \text{Eq. 1}$$

$$SNR = \frac{ROI\ SUV - Background\ SUV}{ROI\ SUV + Background\ SUV} \quad \text{Eq. 2}$$

$$COV = \frac{ROI\ Standard\ Deviation}{ROI_{average}} \quad \text{Eq. 3}$$

2. 통계 분석

영상 통계분석은 SPSS Statistics 29(IBM, Corp., Armonk, NY, USA)를 이용하였으며, 집단 간 SUV_{max} 와 SUV_{mean} 의 유의한 차이를 확인하였다. 자료 분석은 Mann-Whitney U test을 통해 실시하였다.

3. 임상 영상

대상은 본원에서 검사한 환자들 중 D600, D710, DIQ, DMI 장비에서 추적관찰을 목적으로 검사한 환자들을 대상으로 하였으며, 2년 이내에 추적관찰을 목적으로 검사한 환자들을 대상으로 선정하여 분석하였다. 분석을 위해 muscle, lung, liver에 ROI를 설정하여 SUV_{max} 와 SUV_{mean} 을 비교분석하였다. ROI는 lung과 muscle에서 25 mm의 ROI를 설정하였고, liver에서는 60 mm의 ROI를 설정하였다(Fig. 5). 팬텀 영상과 마찬가지로 ROI 설정은 manual으로 하였다.

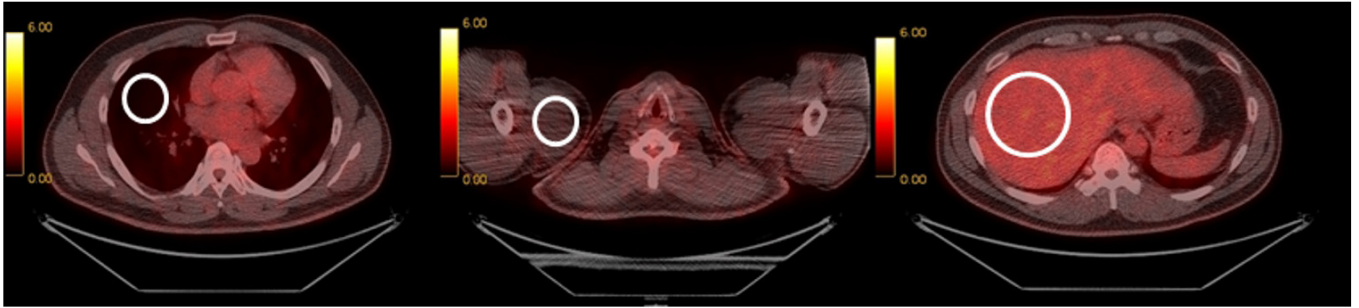


Fig. 5. We set ROI and analyzed the images of follow-up patients by four equipment.

결과

1. SUV 평가

ACR Flangeless Esser Phantom 영상을 통해 SUV_{max} , SUV_{mean} 를 측정하였다. 5회 팬텀 실험 후 평균값으로 비교 분석하였으며 결과는 다음과 같다(Table 3). 장비 간 SUV를 비교하였을 때 SUV_{max} 는 Hot Spot 1부터 4까지 최대 3.0%, 4.6%, 3.1%, 2.8%의 차이를 보였고, SUV_{mean} 은 최대 4.7%, 4.6%, 3.3%, 4.1%의 차이를 보여 장비 간 차이가 최대 5% 이내임을 알 수 있다(Fig. 6). 또한, SUV_{max} 와 SUV_{mean} 모두 장비 간 유의확률 $p > 0.05$ 를 보여 유의한 차이를 보이지 않았다.

Table 3. Experiment of results SUV.

		D600	D710	DIQ	DMI	Diff _{max}
SUV_{max}	Hot Spot 1	2.70	2.76	2.68	2.72	3.0%
	Hot Spot 2	2.30	2.39	2.29	2.4	4.6%
	Hot Spot 3	1.87	1.83	1.88	1.90	3.1%
	Hot Spot 4	1.41	1.45	1.41	1.45	2.8%
SUV_{mean}	Hot Spot 1	2.01	2.07	2.05	2.11	4.7%
	Hot Spot 2	1.72	1.80	1.73	1.76	4.6%
	Hot Spot 3	1.46	1.51	1.50	1.47	3.3%
	Hot Spot 4	1.26	1.27	1.26	1.30	4.1%

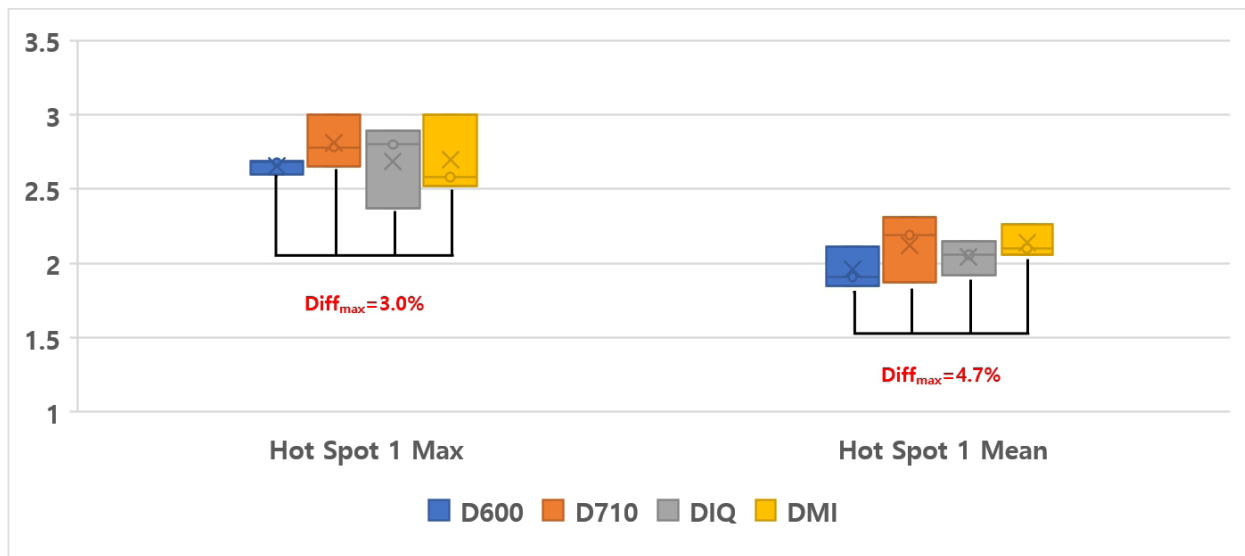




Fig. 6. We compared the values of SUVs for Hot Spot of each equipment.

2. SNR 평가

SNR은 ACR Flangeless Esser Phantom을 이용하여 확인한 SUV와 background의 값을 통해 비교 평가하였으며, Hot Spot 1, 4에서는 최대 0.02의 차이를 보였고, Hot Spot 2에서는 최대 0.04의 차이를 보였으며, Hot Spot 3에서는 최대 0.03의 차이를 보였다(Table 4).

Table 4. Experiment of results SNR.

SNR	D600	D710	DIQ	DMI	Diff _{max}
Hot Spot 1	0.36	0.38	0.37	0.37	0.02
Hot Spot 2	0.26	0.28	0.24	0.28	0.04
Hot Spot 3	0.18	0.20	0.17	0.19	0.03
Hot Spot 4	0.11	0.10	0.10	0.12	0.02

3. Spatial Resolution 평가

spatial resolution 평가를 위해 PET/SPECT Performance Phantom의 22.4mm, 17.9mm, 14.3mm의 hot lesion에서 horizontal FWHM(Full Width at Half Maximum)을 구하였다. profile 분석은 Origin Pro 2024 Node-Lock(Origin Lab, USA)를 이용하였다. hot lesion 별 horizontal FWHM 결과값은 다음과 같다(Fig. 7, Table 5).

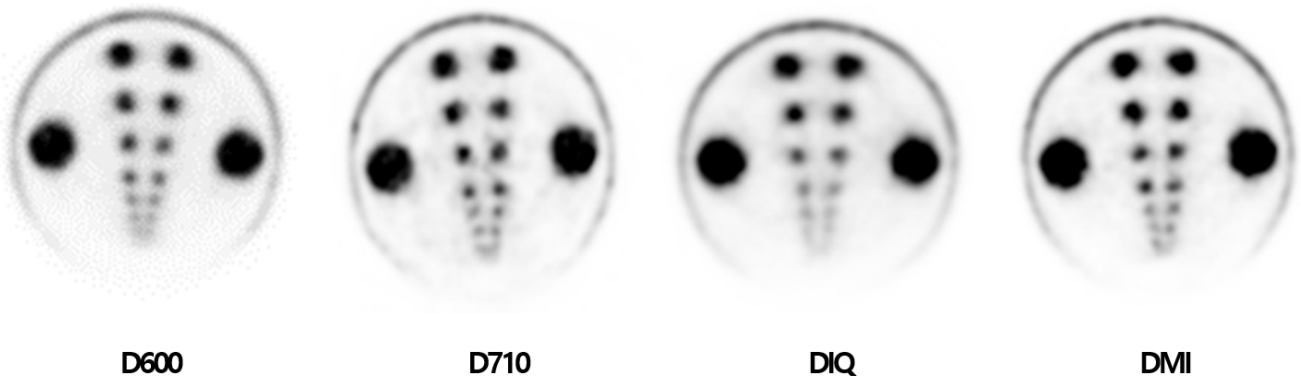


Fig. 7. These figures show Phantom Results tested on four equipment.

Table 5. Experiment of results horizontal FWHM.

Hot Lesion	D600	D710	DIQ	DMI	Diff _{max}
22.4 mm	10.17	10.19	10.31	10.04	0.27
17.9 mm	7.46	7.45	7.46	7.43	0.03
14.3 mm	6.43	6.45	6.52	6.43	0.09

4. Axial Uniformity 평가

균일도 평가는 ROI를 5회 그린 후 평균으로 비교 분석하였다. 장비 간 비교로는 COV를 이용하였으며, 변이계수인 COV는 평균을 표준편차로 나눈 값으로 평균이 다른 대조군과의 비교를 위해 사용한다. background와 COV는 D600에서 1.016과 0.005를, D710에서 1.028과 0.004를 DIQ에서 1.004, 0.018을 보였으며, DMI에서 1.03, 0.006으로 나타났다(Fig. 8, Table 6). ACR Flangeless Esser Phantom Guideline에서 권고하는 axial uniformity 권장 오차 범위는 0.9-1.1으로 장비 모두 이에 부합하였다. COV는 D710에서 가장 낮았으며 DIQ에서 가장 높았다.

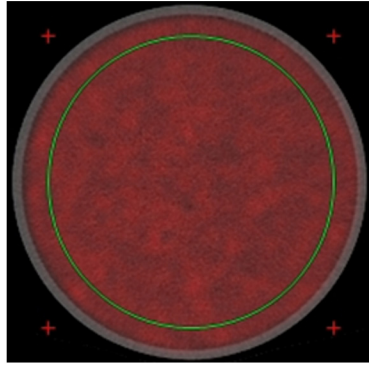


Fig. 8. This figure shows axial uniformity results tested on four equipment models.

Table 6. Experiment of results background and COV.

	D600	D710	DIQ	DMI
Background	1.016	1.028	1.004	1.03
COV	0.005	0.004	0.018	0.006

5. 임상 영상 평가

본원에서 검사한 환자들 중 2년 이내에 추적관찰을 목적으로 검사한 임상 영상으로 SUV_{max} 와 SUV_{mean} 을 평가하였으며 팬텀 실험과 같은 방법으로 장비 모델 간의 차이를 확인 후 최대 차이를 비교하였다(Table 7). SUV_{max} 와 SUV_{mean} 은 Muscle에서 4.5%, 2.9%의 차이를 보였고, Lung에서는 4.4%, 4.1%의 차이를 보였으며, Liver에서는 2.4%, 4.0%의 차이를 보였다. 스캔 영상은 다음과 같다(Fig. 9).

Table 7. Experiment of results SUV in patient image.

SUV	D600	D710	DIQ	DMI	Diff _{max}
Muscle _{max}	0.92	0.93	0.89	0.90	4.5%
Muscle _{mean}	0.67	0.67	0.66	0.68	2.9%
Lung _{max}	0.68	0.71	0.68	0.70	4.4%
Lung _{mean}	0.47	0.49	0.48	0.49	4.1%
Liver _{max}	2.87	2.92	2.94	2.93	2.4%
Liver _{mean}	2.42	2.52	2.44	2.50	4.0%

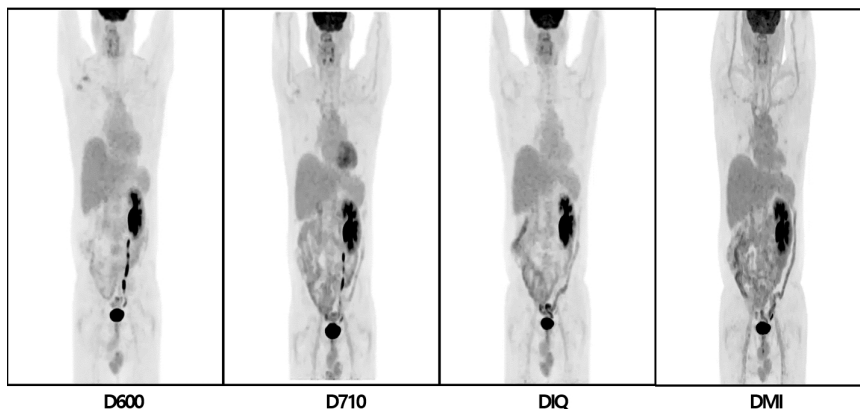


Fig. 9. The figures are an image of follow-up observation with four equipment in the process of treatment.

고찰

본 연구에서는 정량지표인 SUV와 SNR, spatial resolution, axial uniformity와 같은 영상지표로 재원에 따른 장비 모델 간의 차이를 확인하였다. CT 기종의 차이에 따른 SUV 변화가 5% 이내로 크지 않음을 유추할 수 있는 선행 논문에서 scintillator, PM tube와 같이 PET/CT를 구성하는 장비 재원의 차이에 의한 SUV의 차이를 확인하였다[5]. 본 실험에서는 4개의 동일한 제조 회사의 장비와 재구성 방법을 설정하여 5%이내의 SUV 차이를 보였으나, 서로 다른 제조사의 장비의 SUV를 비교하였을 때의 차이가 발생할 수 있는 것으로 보아 장비 재원과 재구성 방법의 차이로 인해 SUV값이 달라지는 것을 확인하였다[6]. 본 연구에서는 SUV_{max} 와 SUV_{mean} 의 장비 간 비교에서 최대 5% 이내의 차이를 보였지만, 실험에서 사용한 D710과 DMI 장비에서 SUV가 상대적으로 높게 측정되는 결과를 보였다. 이는 D600과 DIQ와 다른 scintillator의 차이(BGO, Lutetium based)로 예상된다. 또한, 실험에 사용된 4개의 장비 중 DMI 장비에서 가장 높은 SUV를 보였는데 반도체 타입의 PM tube를 적용한 결과로 보인다. 본 실험을 진행하며 조건 설정 한계로 재구성 parameter 중 D600, D710, DMI장비에서 subsets을 16으로 설정하였고, DIQ장비에서는 12로 설정하여 실험을 진행하였다. 임상영상 간 비교평가를 통해 실제 추적관찰 간 병변의 치료, 재발 등 다양한 변수에 따라 SUV가 변화한다는 점에서 본 논문에서는 다른 시술 및 치료에 대한 변수가 고려되지 않아 lung, liver와 background의 역할이 되는 muscle의 섭취정도를 비교 평가하였다. 추가적으로, 본 연구에서 사용한 VPHDs의 경우 TOF(Time of Flight)를 고려하지 않은 재구성 방법으로 추후 연구 시 다양한 재구성에 따른 연구가 필요할 것으로 보인다.

결론

PET/CT는 암의 진단과 치료계획, 추적관찰 등에 유용해 매년 검사의 수요가 증가하고 있으며, 장비의 발전에 따라 다양한 장비에서의 검사와 추적관찰이 이루어지고 있다. 하지만 검사 장비 재원의 차이와 재구성 방법에 따라 측정되는 SUV가 달라질 수 있으며, 이에 따라 본 연구는 장비 모델 별 영상 비교평가를 통해 장비 재원 간에 따른 차이를 확인하고자 하였다. 이는 다양한 장비로 추적관찰이 가능할 것으로 예상되며, 미세한 차이라도 영상 지표에 따른 차이는 판독 시 참고사항이 될 수 있을 것으로 예상된다. 추후, 장비 노후에 따른 교체, 장비 발전 등에 따라 장비 모델 간의 비교가 필요할 경우 활용가능한 기초자료로 사용될 것으로 사료된다.

REFERENCES

1. Kang KW, KOH's NUCLEAR MEDICINE. 4th edition. Seoul: Korea Medical Book Publishing Company; 2019. p. 3.
2. The Korea Food and Drug Administration. Guideline for testing PET/CT stability and performance evaluation. The Korea Food and Drug Administration. 2015. Korea.
3. American College of Radiology. ACR Nuclear Medicine Accreditation Program Phantom Testing. 2023. USA.
4. Moon AR, Lee H, Kwak IS, Choi SW, Suk JD. Evaluation of Reasonable 18F-FDG Injected Dose for Maintaining the Image Quality in 3D WB PET/CT. J Nucl Med Technol. 2011;15(2):36-40.
5. Kim WH, Go HS, Lee JE, Kim HS, Ryu JK et al. The Comparison Evaluation of SUV Using Different CT Devices in PET/CT Scans. J Nucl Med Technol. 2014;18(1):10-18.
6. Kim TY, Lim JJ, Lee HJ, Kim HJ, Kim JH et al. Assessment and Comparison of SUVs of Three Different PET/CT Scanners. J Nucl Med Technol. 2011;15(1):34-38.