

FDG PET/CT 검사 시 참고장기에서 측정된, 제지방체중으로 표준화한 표준화 섭취계수의 관찰자 사이 및 관찰자 내 재현성에 대한 연구

원광대학교병원 핵의학과

김성수 · 신용철 · 이선도 · 이남주 · 김종철 · 이춘호

Interobserver and Intraobserver Reproducibility of SUL Measurements in Reference Organs on FDG PET/CT

Seong Su Kim, Yong Cheol Shin, Sun Do Lee, Nam Ju Lee, Jong Cheol Kim and Chun Ho Lee
Department of Nuclear Medicine, Wonkwang University Hospital, Iksan, Korea

Purpose: The use of SUV which should be normalized by lean body mass (LBM) is recommended for PET response criteria in solid tumors. LBM which was determined by whole body CT was used for SUV normalization (SUL) in this study. The purpose of the present study was to assess interobserver and intraobserver reproducibility of SUL measurements in reference organs. **Materials and Methods:** F-18 FDG PET/CT was conducted on 52 subjects and LBMs were directly determined by whole body CT for normalization of SUV. The 3 cm diameter spherical VOI, 1×2 cm cylindrical VOI, 2 cm diameter spherical VOI were placed in the liver, descending aorta and spleen, respectively. Experienced two observers measured SULmax and SULmean in each organ. Repeated measurements were conducted two weeks apart by observer 1 blind to previous results. Similarly, measurements were conducted on the same patients by observer 2. For assessing reproducibility (or repeatability), the paired t-test, Pearson's correlation coefficients (CC), and technical error of measurement (TEM) were calculated. **Results:** For interobserver reproducibility in liver SULmax and SULmean, no significant differences were found between observers (paired t-test, $P=0.536$, 0.293 , respectively). CC and TEM for liver SULmean were 0.909 ($P=0.000$) and 0.067 SUL unit, respectively. Corresponding figures for liver SULmax were 0.882 ($P=0.000$) and 0.117 SUL unit, respectively. For intraobserver reproducibility in liver SULmax and SULmean, no significant differences were observed within observer 1 (paired t-test, $P=0.374$, 0.268 , respectively). CC and TEM for liver SULmean were 0.924 ($P=0.000$) and 0.061 SUL, respectively. Corresponding figures for liver SULmax were 0.908 ($P=0.000$) and 0.104 SUL, respectively. Similarly, no significant differences were found in SULmax and SULmean of the spleen and aorta between observers. **Conclusion:** The current study demonstrated that both SULmean and SULmax measurements in normal reference organs are highly reproducible. Reproducibility of SULmean in reference organs were slightly better than SULmax. Interobserver technical error of measurement was less than 0.10 SUL unit for liver SULmean, and 0.12 SUL unit for liver SULmax. Intraobserver technical error of measurement was less than 0.07 SUL unit for liver SULmean, and 0.11 SUL unit for liver SULmax. (*Korean J Nucl Med Technol 2013;17(1):11-17*)

Key Words : Standardized uptake value, Lean body mass, PET/CT, Reproducibility

서 론

- Received: January 31, 2013. Accepted: April 3, 2013.
- Corresponding author : **Chun Ho Lee**
Department of Nuclear Medicine, Wonkwang University Hospital,
344-2 Shinyong-dong, Iksan 570-711, Korea
Tel: +82-63-859-1826, Fax: +82-63-852-1310
E-mail: chlee1667@hanmail.net

과거 종양 치료 후 반응 평가는 전통적으로 해부학적 영상에서 종양 크기 변화를 측정하여 이루어져 왔다. 실제로 지난 2000년 국제적인 협력을 통해 Response Evaluation Criteria in Solid Tumors (RECIST) 기준이 마련되었으며

2009년에 이를 개정된 RECIST 1.1이 제안되어 임상에 사용되고 있다. 그러나 CT와 같은 종양의 크기 변화를 나타내는 해부학적 영상은 변화를 측정할 수 있을 정도로 크기가 변화하기까지 수주에서 수개월이 걸려 조기에 치료 반응을 예측하는데 한계가 있다. 해부학적 변화에 선행하는 기능적, 생화학적 변화를 측정하여 보다 조기에 그리고 정확하게 치료 반응을 평가하기 위해서 최근 FDG PET/CT가 임상에 널리 적용되고 있다. 2009년 Wahl 등이 이러한 PET 영상 데이터를 고형암의 치료반응 평가 기준으로 설정한 PET Response Criteria in Solid Tumors (PERCIST)를 제안하였다.¹⁾

PERCIST에서는 기존에 널리 사용되었던 표준화섭취계수(standardized uptake value, SUV) 대신 제지방체중(Lean Body Mass, LBM)으로 표준화한 SUV, 즉 SUL (SUV normalized by LBM)을 사용할 것을 권고하고 있다. 이는, 전체 체중으로 표준화한 SUV가 체중에 비례해 증가하는 체중의존성이 강한 반면, 제지방체중으로 표준화하면 체중의존성을 줄여줄 수 있기 때문이다.^{2,3)}

제지방체중이란 전체 체중에서 지방(fat) 또는 지방조직(adipose tissue)을 제외한 체중을 말한다. 제지방체중을 측정하는 여러 가지 방법이 있는데 기존의 프로그램들에서는 체중과 키, 성별, 나이 등의 변수를 이용하는 예측식(predictive equation)을 이용해 제지방체중을 계산하여 SUL을 측정하였다. 그 이외에 dual energy X-ray absorptiometry, bio-impedance assay, conventional reference methods, CT 또는 MRI 등이 있다. 이 중 'gold standard'로 여겨지는 conventional reference methods에는 hydrodensitometry (underwater weighing), total body water measurement, total body potassium counting 등이 있으나 특별한 시설과 장비가 갖추어져 있어야 가능하다.⁴⁾ 하지만 전신 CT를 이용하는 방법은 임상에서 쉽게 적용할 수 있을 뿐 아니라 환자 개인의 지방 조직을 정확하게 측정하는 것으로 알려져 있다.⁵⁾ 따라서 PET/CT 검사 시 전신 CT를 시행하여 계산한 제지방체중을 SUV의 표준화에 활용할 수 있을 것이다.⁶⁾

모든 측정방법에는 측정오차가 따르기 마련이다. 새로운 측정방법을 임상에 적용하기 위해서는 재현성이 확립되어야 하며 실제 신호변화(true metabolic change)와 구별하기 위해서 그 측정오차의 범위를 대략 알아야 한다. 종양 병변에서 두 번 PET/CT를 시행하여 SUV 측정의 재현성을 보고한 연구는 수 편 있지만 SUL 측정의 재현성에 대한 연구는 저자들이 알기로는 보고된 바 없다.⁷⁻⁹⁾

이 연구의 목적은 FDG PET/CT 검사 시 참고장치에서

측정한 SUL의 관찰자 사이 및 관찰자 내 재현성을 평가하는 데 있다.

대상 및 방법

2010년 8월부터 2012년 4월까지 원광대학교병원에서 PET/CT를 시행한 환자 및 지원자를 대상으로 연구를 시행하였다. 이 중 참고장치인 간, 복부 대동맥, 비장에 질환이 있는 사람과 조영증강 CT를 시행한 사람은 제외하였다. 남자 27명, 여자 25명 총 52명을 대상으로 전신 CT를 포함한 FDG PET/CT (Biograph 16, SIEMENS, Knoxville, TN, USA)를 시행하였다. PET은 통상적인 방법으로 스캔하였다. CT 스캔 길이가 최대 185cm이어서 환자의 팔을 몸에 붙인 자세로 전신이 field of view에 포함되도록 스캔하였다. CT acquisition 파라미터는 다음과 같다: 최대 관전압 120 kVp, 관전류 60 mA, 관 회전속도 0.5 s, 테이블 이동속도 36 mm/s, 선속 조준 (beam collimation) 24 mm/16채널, matrix 크기 512 × 512, 절편두께 5 mm로 하였다.

CT number 범위가 -140 HU에서 -30 HU 인 voxel을 지방조직으로 판단하였고¹⁰⁾ 위크스테이션에 내장된 프로그램(Volumetric analysis, Siemens Leonardo system)을 이용하여 그 용적을 계산하였다. 이 용적에 이미 알려진 지방조직의 밀도(0.95 kg/l)를 곱하여 개별 환자의 지방조직의 무게를 계산하였고, 다시 전체 체중에서 빼 주어 제지방체중을 구하였다.

PERCIST에서는 간, 하행 대동맥을 참고장치로 정하고 있으며 우리는 이외에도 비장을 참고장치로 추가하여 측정해 간에 병변이 있는 환자들에서 참고장치로 사용할 수 있을지 알아보았다. VOI (Volume of interest)는 PERCIST에서 제안한 대로 간과 하행 대동맥에서는 각각 지름 3 cm의 구형 VOI, 지름 1 cm × 높이 2 cm의 원통형 VOI를 그렸고, 비장에서는 지름 2 cm의 구형 VOI를 그렸다. 간에서는 우엽의 비교적 균일한 FDG 섭취를 보이는 위치에 VOI를 그렸고, 하행 대동맥에서는 동맥벽과 같습 침착된 곳을 피하여 채장이 관찰되는 복부 대동맥에 그렸으며, 비장에서도 FDG 섭취가 비교적 균일한 부위에 VOI를 그렸다(Fig. 1).

경험 있는 두 명의 관찰자가 각 장치에서 SUV_{max}와 SUV_{mean}을 측정하였다. SUL은 전체 체중 대신 제지방체중으로 표준화한 수치이므로 Eq 1에 의해 SUL_{max}와 SUL_{mean}을 다시 계산하였다.

$$\text{Eq 1. } SUL = SUV \times \frac{LBM}{BW}$$

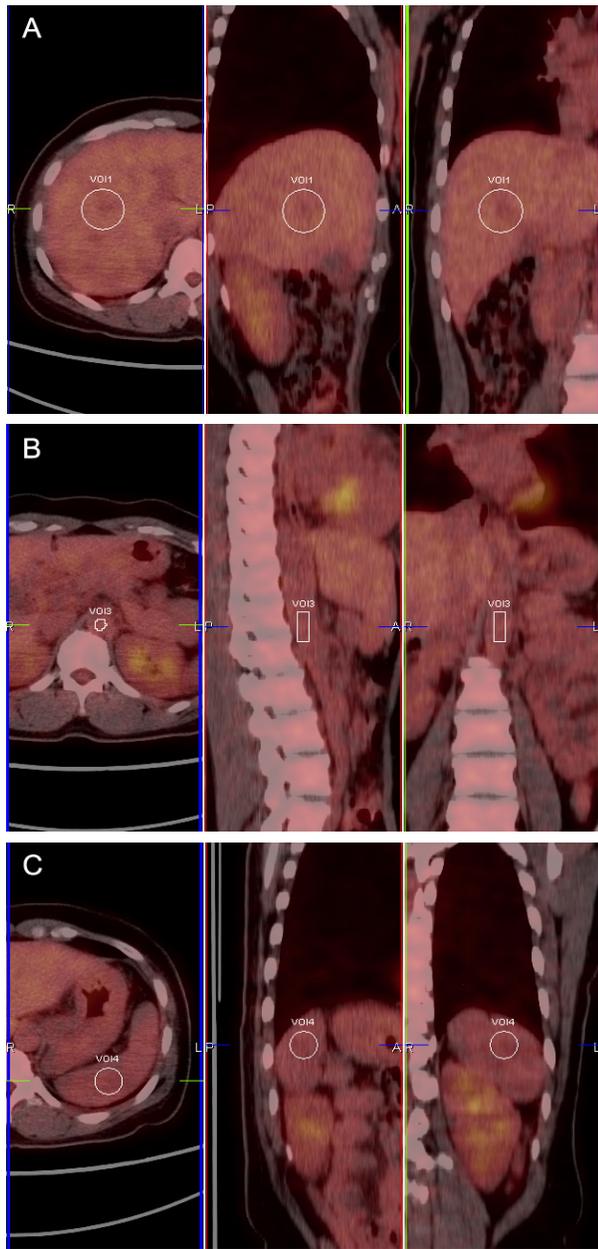


Fig. 1. The 3 cm diameter spherical VOI (A), 1×2 cm cylindrical VOI (B), 2 cm diameter spherical VOI (C) were placed in the liver, abdominal aorta and spleen, respectively.

관찰자 사이 재현성을 평가하기 위해 관찰자 1과 2가 동일 환자의 간, 복부 대동맥, 비장에서 SUL을 각각 독립적으로 측정하였다. 또한 관찰자 내의 재현성을 평가하기 위해 관찰자 1은 2주 간격으로 이전 측정결과를 알지 못하는 상태에서 간에서 SUL을 반복 측정하였다.

재현성(반복성)을 평가하기 위해서 대응표본 *t*-test를 시행하고, Pearson의 상관계수 (correlation coefficient, CC)와 technical error of measurement (TEM)을 계산하였다.

SUL 값이 체중의존성을 얼마나 개선하는지 알아보기 위해 전체 체중과 각 참고장치에서 측정된 SUV 및 SUL 측정치들과의 관계를 단순회귀분석 하였다. 대응표본 *t*-test, Pearson의 상관계수, 단순 회귀분석은 PASW Statistics version 18.0을 사용하였으며, TEM은 Eq 2에 의해 계산하였다.¹¹⁾ 유의확률 $P < 0.05$ 인 경우를 통계적으로 유의하다고 판단하였고, CC는 0.8 이상일 때 강한 상관관계를 갖는다고 판단하였다.

$$\text{Eq 2. } TEM = \sqrt{\frac{\sum_{i=1}^N d_i^2}{2N}} \quad (d = \text{두 측정치의 차이, } N = \text{표본수})$$

결 과

1. 관찰자 사이 재현성

관찰자 1과 관찰자 2가 측정된 간, 복부 대동맥, 비장의 SUL_{max} 사이의 대응표본 검정 결과 통계적으로 유의한 차이가 없었으며 (각각 $P=0.536, 0.593, 0.320$), 이는 SUL_{mean} 사이에서도 마찬가지였다 (각각 $P=0.293, 0.777, 0.506$). 상관계수 또한 복부 대동맥을 제외하면 0.8 이상으로 서로 강한 양의 상관관계를 보였다. 또한 SUL_{max} 의 TEM은 간, 복부 대동맥, 비장에서 각각 0.117, 0.128, 0.091 SUL unit 이었고, SUL_{mean} 의 TEM은 각각 0.067, 0.081, 0.040 SUL unit 이었다(Table 1, Fig. 2).

Table 1. Interobserver repeatability of SUL measurements in reference organs

	Observer 1 (mean±S.D.)	Observer 2 (mean±S.D.)	Paired <i>t</i> -test (<i>P</i> value)	CC [†]	TEM [‡] (SUL [*])
Liver SUL_{max}	2.053±0.350	2.067±0.332	0.536	0.882	0.117
Liver SUL_{mean}	1.536±0.226	1.549±0.217	0.293	0.909	0.067
Aorta SUL_{max}	1.574±0.293	1.561±0.267	0.593	0.791	0.128
Aorta SUL_{mean}	1.156±0.156	1.152±0.142	0.777	0.694	0.081
Spleen SUL_{max}	1.508±0.240	1.527±0.252	0.320	0.865	0.091
Spleen SUL_{mean}	1.126±0.142	1.131±0.155	0.506	0.928	0.040

^{*}SUL = standardized uptake value normalized by lean body mass, [†]CC = correlation coefficient, [‡]TEM = technical error of measurement

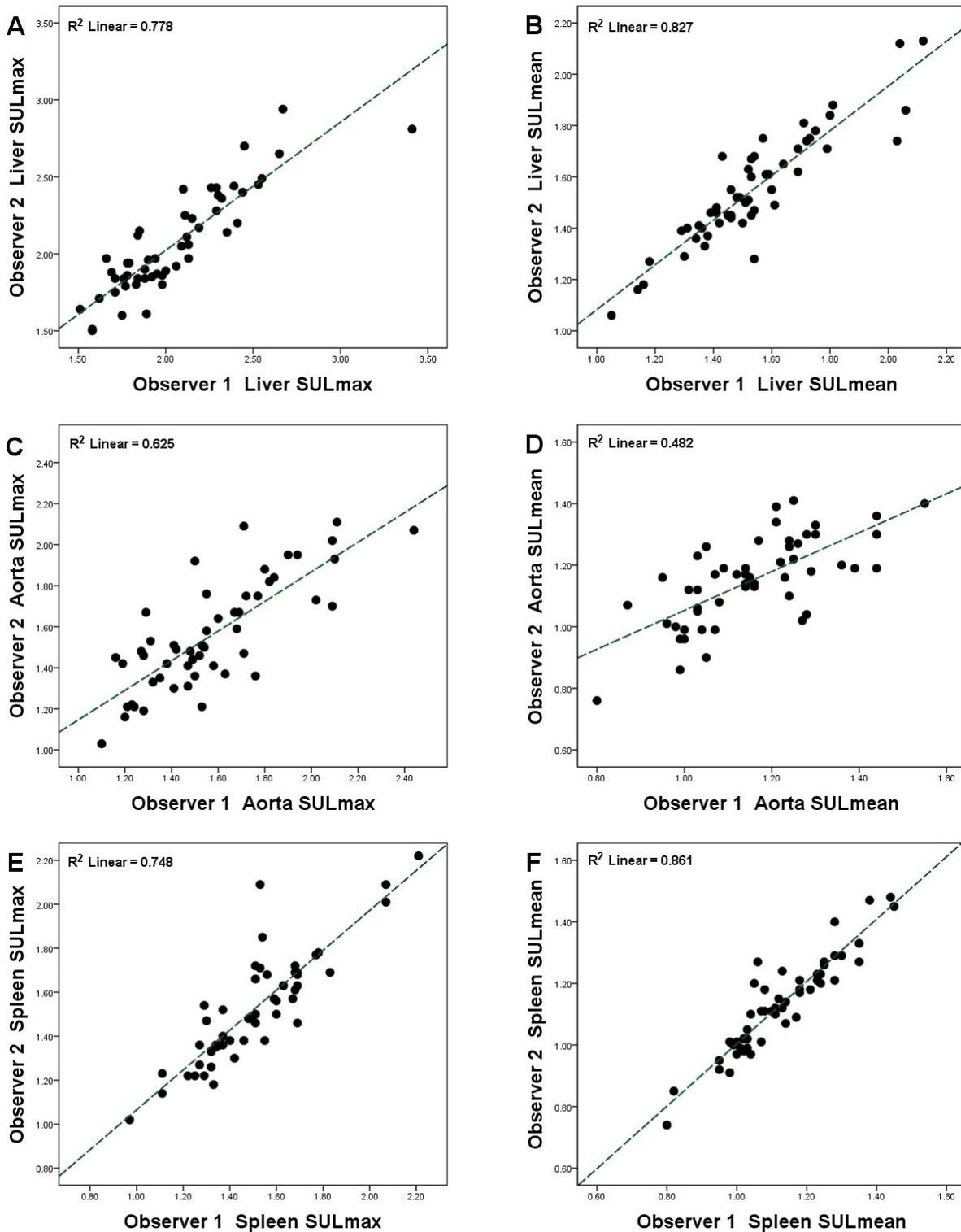


Fig. 2. The scatter diagrams show correlation between observer 1 and observer 2 in measuring liver SUL_{max} (A), liver SUL_{mean} (B), aorta SUL_{max} (C), aorta SUL_{mean} (D), spleen SUL_{max} (E) and spleen SUL_{mean} (F). Dashed lines in each scatter diagrams represent regression line. The coefficients of determination (R²) for each regression analysis are 0.778, 0.827, 0.625, 0.482, 0.748 and 0.861, respectively.

Table 2. Intraobserver repeatability of SUL measurements in reference organ

	Observer 1A (mean±S.D.)	Observer 1B (mean±S.D.)	Paired <i>t</i> -test (<i>P</i> value)	CC	TEM (SUL)
Liver SUL _{max}	2.053±0.350	2.072±0.310	0.374	0.908	0.104
Liver SUL _{mean}	1.536±0.226	1.522±0.210	0.268	0.924	0.061

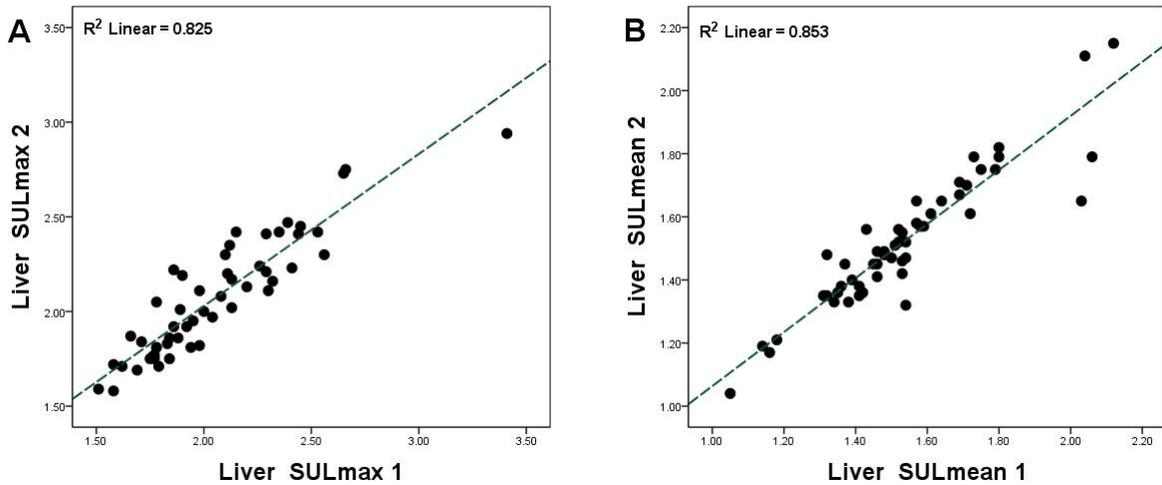


Fig. 3. The scatter diagrams show correlation between replicative measurements by observer 1 in measuring liver SUL_{max} (A) and liver SUL_{mean} (B). Dashed lines in each scatter diagrams represent regression line. The coefficients of determination (R^2) for each regression analysis are 0.825 and 0.853, respectively.

2. 관찰자 내 재현성

간에서 관찰자 1이 2주 간격으로 두 번 측정된 SUL_{max} 사이의 대응표본 검정 결과 통계적으로 유의한 차이가 없었고($P=0.374$), 이는 SUL_{mean} 사이에서도 마찬가지였다($P=0.268$). SUL_{max}와 SUL_{mean} 모두 두 번 측정된 값들의 상관관계수가 0.9 이상으로 강한 양의 상관관계를 보였다. 관찰자 내 SUL의 TEM은 SUL_{mean}과 SUL_{max}에서 각각 0.061, 0.104 SUL unit 이었다(Table 2, Fig. 3).

3. 체중 의존성 비교

각 환자의 전체 체중과 간의 SUV_{max} 사이의 관계, 그리고 각 환자의 전체 체중과 간의 SUL_{max} 사이의 관계, 각 환자의 전체 체중과 간의 SUV_{mean} 사이의 관계, 그리고 각 환자의 전체 체중과 간의 SUL_{mean} 사이의 관계를 단순회귀분석하여 SUV와 SUL의 체중의존성을 비교하였다. 회귀분석에서 결정계수(R -Square : R^2)란 종속변수의 전체 변동 중에서 회귀모형에 의하여 설명할 수 있는 변동의 정도를 말한다. 즉, 결정계수가 크면 클수록 두 변수 간의 회귀관계는 직선적 관계에 있다는 뜻이 된다. SUV_{max}의 결정계수 0.579가

SUL_{max}의 결정계수 0.378로 감소하였으며, SUV_{mean}의 결정계수 0.293이 SUL_{mean}의 결정계수 0.146로 역시 감소하였다. 따라서 SUL값의 체중의존성이 SUV보다 현저히 감소되었음을 시사하나 SUL_{max}는 여전히 약간의 체중의존성을 보인다(Fig. 4).

고찰 및 결론

관찰자 사이 또는 관찰자 내의 측정오차는 여러 요소가 관여할 것이다. 우선 각 참고 장치에서 VOI를 동일한 위치에 설정하기 어려우며, 복부 대동맥의 경우 주변 척추의 퇴행성 병변, 대동맥 자체의 칼슘침착병변으로 인한 영향이 있을 수 있고, 비장은 그 크기와 모양의 개인차가 커 지름 2 cm의 구형 VOI를 그리기 어려운 경우도 있다는 점 등이 있었다. 실제로 복부 대동맥의 경우 두 측정치의 상관관계가 다른 두 장기에 비해 다소 떨어지는 경향을 보였다.

그럼에도 불구하고 관찰자 사이 또는 관찰자 내의 측정치들은 통계적으로 유의한 차이가 없어 우수한 재현성을 보여주고 있다. 따라서 PET/CT 검사 시 전신 CT를 시행하여 계측한 제지방체중으로 표준화한 SUL의 측정법은 임상에서의 적용이 충분히 가능할 것으로 생각된다.

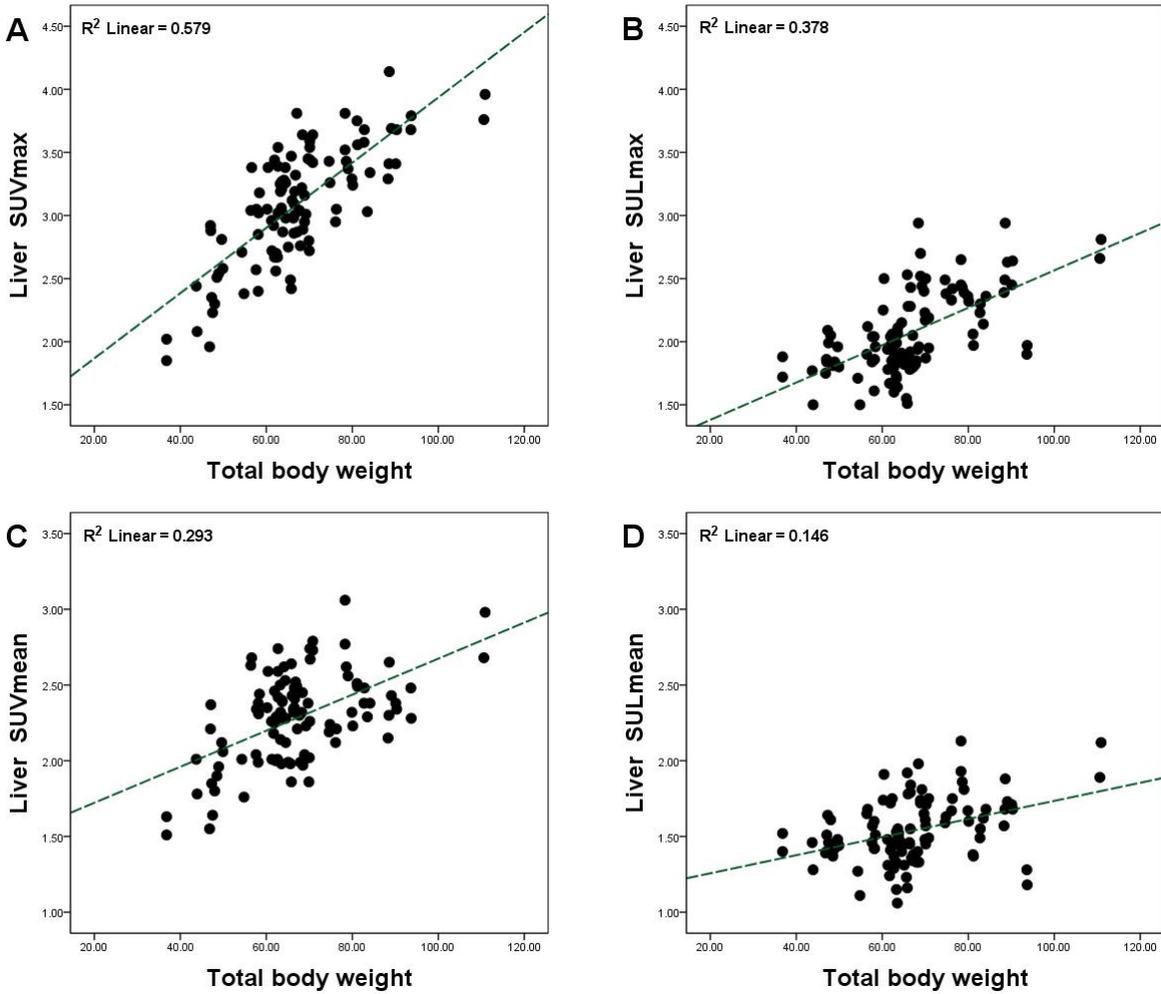


Fig. 4. The scatter diagrams show relationship between liver SUV_{max} and total body weight (A), between liver SUL_{max} and total body weight (B), between liver SUV_{mean} and total body weight (C), between liver SUL_{mean} and total body weight (D). Dashed lines in each scatter diagrams represent regression line. The coefficients of determination (R^2) for each regression analysis are 0.579, 0.378, 0.293 and 0.146, respectively. In all cases(A, B, C and D), R^2 values are smaller in SUL measurements than that of SUV, indicating reduced body weight dependency in SUL measurements.

복부 대동맥의 경우 대응 표본 검정에서 유의한 차이는 없었지만 상관계수는 간과 비장에 비해 상대적으로 낮았고, 이에 반해 비장은 간과 마찬가지로 우수한 재현성을 보이고 있어 앞으로 정상 참고 장기로 간을 사용할 수 없을 경우 대체 참고장기로 사용할 수 있음을 시사한다.

본 연구에서는 한 번의 검사 결과를 가지고 관찰자 사이 또는 관찰자 내의 재현성을 평가하였지만, 본원에서 같은 환자에서 짧은 시간 간격을 두고 두 번 시행한 검사에서의 재현성을 평가하는 연구가 진행되고 있으며, 나아가 정상 참고장기가 아닌 병변에서의 SUL 측정의 재현성도 연구해야 할 과제이다.

최근 PET 영상 분야에서는 기존의 ‘SUV_{max}’ 기반으로 이

루어진 종양의 치료반응 평가에서 벗어나 일정 역치 이상의 ‘metabolic volume’ 기반의 평가에 대한 연구가 대두되고 있다.¹²⁾ 본 연구에서 사용한 전신 CT를 시행하여 측정된 체지방체중으로 표준화한 SUL의 새로운 측정법은 개별 환자의 체중에 따른 의존성을 개선할 수 있을 뿐 아니라, 향후 ‘metabolic volume’ 기반의 연구에 활용한다면 더욱 기능적이며 생리적인 PET 척도를 개발하는 데에도 도움이 될 것이다.

요 약

고형암의 치료반응 평가 기준을 제시한 PERCIST에서는

제지방체중으로 표준화한 표준화 섭취계수 즉, SUL을 사용하도록 권고하고 있다. 이 연구에서는 전신 CT에서 측정된 제지방체중을 이용해 SUV를 표준화하여 얻은 SUL의 새로운 측정법을 제시하고 관찰자 사이 및 관찰자 내 재현성 여부를 평가하였다. 52명의 대상에서 전신 CT를 포함한 PET/CT를 시행하였고 -140~-30HU를 가지는 지방조직의 용적을 정확히 측정할 수 있었다. 이를 이용하여 개별 환자에서의 제지방체중을 계산하였고, 이를 이용해 SUV를 표준화하여 SUL을 측정하였다. 정상 참고 장기인 간, 복부 대동맥, 비장에 각각 지름 3 cm의 구형 VOI, 1×2 cm의 원통형 VOI, 그리고 지름 2 cm의 구형 VOI를 그렸다. 경험 있는 두 명의 관찰자가 각 장기의 VOI에서 SUL_{max} 와 SUL_{mean} 을 측정하였다. 관찰자 1이 2주 간격으로 이전 측정결과를 알지 못하는 상태에서 반복 측정하였고 같은 방법으로 관찰자 2가 동일 환자에서 SUL을 측정하였다. 재현성을 평가하기 위해서 대응표본 *t*-test를 시행하고 Pearson의 상관계수 (CC)를 계산하였고, 측정오차를 알아보기 위해 TEM을 계산하였다. 관찰자 사이 재현성 즉, 관찰자 1과 관찰자 2가 측정된 간, 복부 대동맥, 비장의 SUL_{max} 사이에는 통계적으로 유의한 차이가 없었고 이는 SUL_{mean} 에서도 마찬가지였다. 관찰자 내 재현성 즉, 간에서 관찰자 1이 2주 간격으로 두 번 측정된 SUL 측정치 사이에도 통계적으로 유의한 차이가 없었다. 이 연구는 정상 참고 장기에서 SUL 측정이 매우 재현성이 있음을 보여준다. SUL_{mean} 측정의 재현성이 SUL_{max} 보다 약간 우수하였으며 SUL 측정치는 체중 의존성이 개선되었다. 전신 CT를 시행하여 측정된 제지방체중으로 표준화한 SUL의 새로운 측정법은 향후 여러 연구에서 유용하게 이용될 것으로 기대된다.

REFERENCES

1. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: Evolving Considerations for PET response criteria in solid tumors. *J Nucl Med* 2009;50:122S-150S.
2. Sugawara Y, Zasadny KR, Neuhoff AW, Wahl RL. Reevaluation of the standardized uptake value for FDG: variations with body weight and methods for correction. *Radiology* 1999;213:521-525.
3. Zasadny KR, Wahl RL. Standardized uptake values of normal tissues at PET with 2-[fluorine-18]-fluoro-2-deoxy-D-glucose: variations with body weight and a method for correction. *Radiology* 1993;189:847-850.
4. Mattsson S, Thomas BJ. Development of methods for body composition studies. *Phys Med Biol* 2006;51:R203-R208.
5. Thaete FL, Colberg SR, Burke T, Kelley DE. Reproducibility of computed tomography measurement of visceral adipose tissue area. *Int J Obes* 1995;19:464-467.
6. Kim WH, Kim CG, Kim DW. Comparison of SUVs normalized by lean body mass determined by CT with those normalized by lean body mass estimated by predictive equations in normal tissues. *Nucl Med Mol Imaging* 2012;46:182-188.
7. Langen AJ, Vincent A, Velasquez LM, Tinteren H, Boellaard R, Shankar LK, Boers M, Smit EF, Stroobants S, Weber WA, Hoekstra OS. Repeatability of 18F-FDG uptake measurements in tumors: a metaanalysis. *J Nucl Med* 2012;53:701-708.
8. Langen AJ, Klabbbers B, Lubberink M, Boellaard R, Spreuuenberg MD, Slotman BJ, Bree R, Smit EF, Hoekstra OS, Lammertsma AA. Reproducibility of quantitative 18F-3'-deoxy-3'-fluorothymidine measurements using positron emission tomography. *Eur J Nucl Med Mol Imaging* 2009;36: 389-395.
9. Shields AF, Lawhorn-Crews JM, Briston DA, Zalzal S, Gadgeel S, Douglas KA, Mangner TJ, Heilbrun LK, Muzik. Analysis and reproducibility of 3'-Deoxy-3'-[18F]fluorothymidine positron emission tomography imaging in patients with non-small cell lung cancer.
10. Kim WH, Kim CG, Kim DW. Optimal CT number range for adipose tissue when determining lean body mass in whole-body F-18 FDG PET/CT studies. *Nucl Med Mol Imaging* 2012;46: 294-299.
11. Lohman TG, Roche AF, Martorelle R. Anthropometric standardization reference manual. Champaign, IL: Human Kinetics Books; 1988. p.83-86.
12. Moon SH, Hyun SH, Choi JY. Prognostic significance of volume-based PET parameters in cancer patients. *Korean J Radiol* 2013;14:1-12.