Original Article

Molecular Imaging and Radionuclide Therapy 2014;23(1): 16-20 DOI: 10.4274/Mirt.76376

Repeatability of the Maximum Standard Uptake Value (SUV_{max}) in FDG PET

FDG PET'te Maximum Standard Uptake Değerinin (SUV_{max}) Tekrarlanabilirliği

Henry Lindholm[1], Johan Staaf[2], Hans Jacobsson[1], Fredrik Brolin[2], Robert Hatherly[2], Alejandro Sânchez-Crespo[2]

[1]*Karolinska University Hospital, Department of Radiology, Stockholm, Sweden*

[2]*Karolinska University Hospital, Department of Hospital Physics, Stockholm, Sweden*

Abstract

Objective: SUV_{max} is often calculated at FDG PET examinations in systematic studies as well as at clinical examinations. Since SUV_{max} represents a very small portion of a lesion it may be questioned how statistically reliable the figure is. This was studied by assessing the repeatability of SUV_{max} between two FDG acquisitions acquired immediately upon each other in patients with chest lesions.

Methods: In 100 clinical patients with a known chest lesion, two identical 3 min PET registrations (PET1 and PET2, respectively) were initiated within 224±31 sec of each other. The difference in SUV_{max} between the lesion for the two PET scans (ΔSUV_{max}) was calculated and the uncertainty expressed as the coefficient of variation, CV (%). The correlation between ΔSUV_{max} and the lowest SUV_{max} from PET1 or PET2, the approximate metabolic lesion volume, the time from FDG injection to PET1 and the time between PET1 and PET2, respectively, was also assessed.

Results: In 56 patients SUV_{max} increased at the second acquisition and in 44 patients it decreased. Mean of SUV_{max} was 7.8±6.1 and 7.8±6.2 for PET1 and PET2, respectively. The mean percentage difference was 0.9±7.8. The difference was not significant (p=0.20). CV gave an uncertainty of 4.3% between the two measurements which is a strong indicator of equivalence. There was no correlation between ΔSUV_{max} and any of the assessed four parameters. The difference between the acquisitions, 0.9%, was much lower compared to the 3 previous published similar, but more restricted studies where the difference was 2.5-8.2%.

Conclusion: From camera and computational perspectives, SUV_{max} is a stable parameter

Key words: Examinations and diagnoses, positron-emission tomography , fluorodeoxyglucose F18, reproducibility of results

Özet

Amaç: Sistematik çalışmalarda ya da klinikte yapılan muayenelerde yapılan FDG PET çalışmalarında SUV_{max} sıklıkla hesaplanmaktadır. SUV_{max} bir lezyonun küçük bir parçasını temsil etttiği için, bu rakamın istatstiksel olarak güvenilirliği sorgulanabilir. Bu amaçla toraks lezyonu olan hastalarda derhal ardarda yapılan iki FDG PET çekimlerinde SUV_{max}'ın tekrarlanabilirliği araştırılmıştır.

Yöntem: Bilinen toraks lezyonu olan 100 hastada, iki eşdeğer 3 dakikalık PET çekimi (sırasıyla, PET1 ve PET2) aralarında 224±31 saniye olacak şekilde gerçekleştirildi. İki PET taraması arasında lezyondaki SUV_{max} farkı (ΔSUV_{max}) hesaplandı ve belirsizlik varyasyon katsayısı (CV (%)) olarak ifade edildi. ΔSUV_{max} ile; PET1 veya PET2 deki en düşük SUV_{max} değeri, yaklaşık metabolik lezyon hacmi, FDG enjeksiyonundan PET1 e kadar geçen zaman ve PET1 ve PET2 arasındaki süre arasındaki korelasyon hesaplandı.

Bulgular: İkinci çekimde, 56 hastada SUV_{max} arttı, 44 hastada azaldı. Ortalama SUV_{max} PET1'de 7.8±6.1, PET2'de 7.8±6.2 idi. Ortalama yüzde fark 0.9±7.8 idi. Fark önemli bulunmadı (p=0.20). İki ölçüm arasında CV'nin verdiği belirsizlik %4.3 idi ve eşdeğerliğin kuvvetli bir göstergesiydi. ΔSUV_{max} ile diğer araştırılan dört parametre arasında korelasyon görülmedi. %0.9 olarak bulunan iki çekim arası fark değeri, daha önce yayınlanan 3 benzer ancak daha kısıtlı çalışmalarda bulunan farklara (%2.5-8.2) göre daha düşüktü.

Sonuç: Kamera ve hesaplama açısından SUV_{max} stabil bir parametredir.

Anahtar kelimeler: Muayeneler ve tanılar, pozitron emisyon tomografi, fluorodeoksiglukoz F18, sonuçların yeniden üretilebilirliği

Address for Correspondence:
Henry Lindholm MD, Stockholm 17176 Stockholm - Sweden
Phone: +46 85 177 92 66 E-mail: henry.lindholm@karolinska.se **Received:** 20.11.2013 **Accepted:** 23.12.2013

Introduction

In medical imaging there is a tendency to accord numbers special value, irrespective of how robust they are (1). Positron Emission Tomography (PET) represents such a quantitative process, as the raw signal can be transformed into absolute concentrations of the radiotracer after a series of transitions and corrections. Hence, PET is often used for quantification of different molecular processes. This is usually made by calculating the Standardized Uptake Value (SUV). In a given image volume, the SUV is defined as the tracer concentration, normalised to the administered activity and e. g. the body weight. The mean or the maximum SUV (SUV_{mean} and SUV_{max}, respectively) within a somehow defined Volume-Of-Interest (VOI), are commonly used figures for description of the tracer uptake.

SUV based PET image quantifications are hampered by technical, physical and biological processes. Despite these well known limitations, SUV calculations are commonly used at clinical examinations as well as at scientific studies. In this respect, SUV quantifications are often ascribed a value which is overrated in relation to their actual precision (1,2,3,4).

SUV_{max} is regarded as a figure of merit for prognosis and therapy evaluation (5,6,7,8). While SUV_{mean} heavily depends on the definition of the VOI, SUV_{max} does not, thereby being almost reader independent. Furthermore, SUV_{max} is less affected by partial volume effects while it, on the other hand, is affected by the image noise (9,10). This may be considerable as the administered activity and the scanning time is restricted in clinical practice. In addition, SUV_{max} represents a very small portion of the VOI, why it may be questioned how well it reflects the biology of an entire lesion (11). In the current report, the statistical reliability of SUV_{max} has been studied by assessing its repeatability in an uptake of a chest lesion between two identical acquisitions obtained immediately upon each other in clinical patients.

Materials and Methods

Patients

The study is based on 100 clinical patients (mean age 64 years; 53 males and 47 females) with suspected tumours of the chest, referred for a clinical PET/CT examination with [18F]-2-fluoro-2-deoxy-D-glucose (FDG). The Regional Research Ethical Committee approved the study.

Examination

Approximately one hour after i. v. administration of 4 MBq/kg bw of FDG, the examination was initiated. This was accomplished using a Biograph 64 True Point (Siemens Medical Solutions, Erlangen, Germany) with an axial PET Field-Of-View (FOV) of 21.5 cm. First a low dose CT without contrast medium, from the middle skull to the proximal thigh, for photon attenuation and scatter correction of the PET images was performed. Directly following this, the clinical PET-examination was performed with a 3 min acquisition time for each FOV position and normal tidal breathing. Immediately after this, the two additional study-specific identical PET registrations were sequentially acquired (PET1 and PET2, respectively) with an acquisition time of 3 min and one single FOV including the known lesion. There was a mean of 224±31 sec between the beginnings of the two acquisitions. Thereafter, a full-dose CT, with or without administration of i.v. contrast medium, was performed at breath-holding at a mean inspiratory level. The patient did not change position versus the camera during the series of acquisitions.

All PET images were reconstructed using the manufacturer ordered subset expectation maximization algorithm (OSEM) with 4 iterations and a matrix image size of 168x168 pixels, a nominal slice thickness of 5 mm and a voxel volume of 0.08 cm^3. Corrections for photon attenuation, random coincidences and photon scatter were made.

Evaluation

Only patients showing FDG-uptake of the lesion were included. In some patients several lesions were identified, but for all patients only one lesion was studied. The lesions assessed were based on the possibility of including a wide spread of lesion sizes. There were 3 lesions of the chest wall, 16 lesions of the mediastinum/lung hili, and 81 pulmonary lesions. 70 were suspected lung/oesofageal tumours, 16 were metastases or lymphoma/leukaemia, and 14 were considered to have benign explanations (inflammation/infection, radiation pneumonitis, Wegeners granulomatosis or Hamptons hump).

Evaluation of PET1 and PET2 was made using the commercial software allowing simultaneous assessments of both studies at an identical position. SUV_{max} was calculated by allocating a VOI enclosing the uptake with some margin and using this for evaluation of both examinations.

To estimate the metabolic volume of the lesions, the FDG-uptake was approximated as an ellipsoid/sphere and calculated by manually allocating three orthogonal diameters with a precision of a ½ cm. The distribution of the approximately calculated volumes is shown in Figure 1. The mean volume was 50 cm^3, and the median volume was 7 cm^3.

Data Analysis and Statistics

The difference between the SUV_{max} of the two PET scans (ΔSUV_{max}) was calculated as:

$\Delta SUV_{max} = SUV_{max}(PET1) - SUV_{max}(PET2)$

Since the Jarque-Bera test showed that ΔSUV_{max} was not normally distributed, the Wilcoxon matched-pairs signed-ranks test was used to test whether the two measurements were significantly different. The uncertainty of ΔSUV_{max} (the measurement error) was

evaluated according to Dahlberg's formula and presented as the coefficient of variation, CV (%) (12).

Correlations were assessed between ΔSUV_{max} and the four parameters: the lowest (minimum) of the measured SUV_{max} from either PET1 or PET2, the approximate metabolic lesion volume, the time from FDG injection to PET1, and the time between PET1 and PET2. Analysis showed that this could be made by calculation of Pearson correlation coefficient, which can be used as a measure of strength of linear correlations. Of these, there was a skewed distribution (>1) for the lowest of the measured SUV_{max}, the approximate lesion volume and the time from FDG injection to PET1. Thus, a reciprocal transformation was made for these data prior to the analysis. A rule of thumb is that a Pearson correlation coefficient of 0-0.25 indicates little or no relationship (13). The coefficient of determination (R2) was also calculated. This shows the proportion of the total variation explained by the variable studied.

Results

In 56 patients SUV_{max} increased at PET2 and in 44 patients it decreased. Mean of ΔSUV_{max} was 7.8±6.1 and 7.8±6.2 for PET1 and PET2, respectively. The mean percentage difference was 0.9±7.8. The difference was not statistically significant (p=0.20). The distribution of ΔSUV_{max} is shown in Figure 2. Corresponding CV gave a relative uncertainty of 4.3% between the two measurements. The voxel containing the SUV_{max} in the patient with the lowest number of counts after the various transitions and corrections was 28218 which corresponds to a CV of 0.6% (assuming Poisson counting statistics). There was no correlation between ΔSUV_{max} and any of the studied parameters: the lowest of the measured SUV_{max} from either PET1 or PET2, the approximate metabolic lesion volume, the time from FDG injection to PET1, or the time between PET1 and PET2 (Table 1).

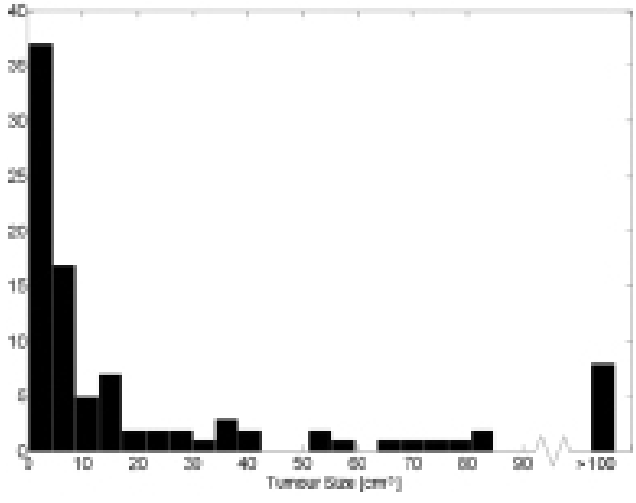

Figure 1. Histogram distribution of the approximately metabolic tumour volume. All tumours larger than 100 cm^3 are presented as one column.

Discussion

The reproducibility of FDG-PET examinations, i. e. the variation between two different examinations carried out at standardised conditions has been studied by several authors and subjected to a metaanalysis (14). In contrast to this, the repeatability of SUV_{max}, i. e. the variability of between two identical consecutive PET scans of the same patient was evaluated in the current study. The aim was to test this at a clinical setting using regularly applied acquisition and reconstruction parameters, thereby also including effects by true image noise. The latter being important as the accuracy of SUV_{max} is limited by a sensitivity to the noise (9,10).

As the lesions were located in the chest, the tracer uptake is affected by respiratory movements. This should not influence the final results, since the movements must have the same impact at the two acquisitions. The metabolic PET volume of the lesion is relevant for the assessments of a possible size influence, why the anatomical (CT) volume was not assessed. The latter would have been much more precise, but of limited value, as the metabolic and anatomic volumes do not always correspond to each other (15). Edge definitions in nuclear medicine examinations, however made, are not precise why the volume assessments are subjected to uncertainties, which represents a limitation of the study. The data presented is dominated by small lesions, this merely reflecting the clinical situation. There was no correlation between the lesion size and ΔSUV_{max} which may have been expected.

Stable conditions between PET1 and PET2 are a prerequisite for the study. This could be influenced by a continuing FDG incorporation still after 60 min, since this has been shown to peak later in tumours

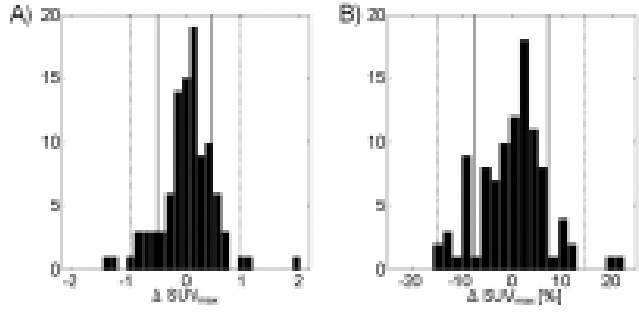

Figure 2. Histogram distribution of the difference ΔSUV_{max}=SUV_{max} (PET1)–SUV_{max}(PET2). Full and dashed lines represent ±1 and 2 SDs respectively. A, Absolute difference. B, ΔSUV_{max} expressed as a percentage of SUV_{max} of PET1.

Table 1. Correlations between ΔSUV_{max} and the lowest of the measured SUV_{max} from either PET1 or PET2, the approximate metabolic lesion volume, the time from FDG injection to PET1 and the time between PET1 and PET2, respectively. ns = not significant

	Lowest SUV_{max}	Lesion volume	Time between FDG injection and PET1	Time between PET1 and PET2
Correlation coefficient	0.08	0.11	-0.12	-0.03
Significance	ns	ns	ns	ns
R2	0.01	0.01	0.01	0.00

Table 2. Reports on the repeatability of SUV_{max}

Author (reference)	ΔSUV_{max} ±SD (%)	Number of patients	Time to first acquisition (min)	Time of acquisition (min)	Administered activity (MBq)	PET/CT-camera
Schwartz et al (18)	2.5±1.7	8	60	3	444/patient	Discovery STE GE Healthcare
Lodge et al (19)	6.1±5.2* 6±6.2	20	60	3	8.4/kg bw	Discovery VCT (RX) GE Healthcare
Burger et al (20)	8.2±11.2	18	35	5	350/patient	Discovery ST 16 or Discovery VCT64 GE Healthcare
Current study	0.9±7.8	100	60	3	4.0/kg bw	Biograph 64 True Point, Siemens Medical Solutions

*Two different image matrices were used.

(16,17). As there was no difference of SUV_{max} between PET1 and PET2, the (short) interval between the two acquisitions allows for the analyses to be made. This is further supported by the lack of correlation of the time between FDG injection and PET1 versus ΔSUV_{max}, as well as the lack of correlation of the time between PET1 and PET2 versus ΔSUV_{max}.

The CV corresponding to the lowest number of counts in the study, 0.6%, is much lower than any other uncertainty of the study. Together with the lack of correlation with the lowest of the measured SUV_{max}, our findings are hardly influenced by an insufficient number of counts.

There are three previous similar studies (18,19,20). They are not as extensive as in the current report and do not include possible effects on the repeatability by other mechanisms as also studied by us. In the previous studies, the variation of SUV_{max} is much larger than in the current study. In this, there was difference of 0.9% between the two acquisitions, while at the previous studies this figure varied between 2.5-8.2% (Table 2). The activity administered at our examinations, 4 MBq/kg bw, is the lowest compared to the previous studies, why this does not explain our lower value. It should rather have an opposite effect. All previous studies are based on a small number of observations, 8-20 individuals, while we studied a much large cohort to cover differences between patients and lesions sizes. In one of the previous studies, the two comparative acquisitions were initiated as early as 35 min after administration of the radiotracer, why a still ongoing strong increasing tracer uptake may explain the higher uptake at the second acquisition (20). Another study was restricted to normal liver tissue uptake of FDG (18). In contrast to this, the current study is based on pathological lesions located in the chest, which compared to the normal liver uptake are subjected to a lower influence from the surrounding activity, thereby reducing any errors caused by partial volume effects. The divergence towards the previous studies cannot, however, be completely explained other than the current study is based on a larger number of observations and very strictly controlled.

The calculated coefficient of variation (CV), of 4.3% between the examinations is very small. Results of < 5% are considered as a strong indicator of equivalence (12). Consequently, the observed relative measurement error is most likely only due to quantum statistical fluctuations in the disintegration and detection of the positrons. Any odd SUV_{max} results cannot be blamed on a random error in image reconstruction or hardware, but that the value is stable in regards to those factors.

Conclusion

From camera and computational perspectives, the SUV_{max} is a stable parameter. Any fluctuation can be explained by physiological variations in the radiopharmaceutical uptake, which is what it is meant to describe in the end.

Acknowledgment: The authors wish to thank Elisabeth Berg, B. Sc., Karolinska institutet, for the professional statistical analysis.

Conflicts of interest: None.

References

1. Keyes JW Jr. SUV: standard uptake or silly useless value? J Nucl Med 1995;36:1836-1839.
2. Vriens D, Visser EP, de Geus-Oei L-F, Oyen WJG. Methodological considerations in quantification of oncological FDG PET studies. Eur J Nucl Med Mol Imaging 2010;37:1048-1025.
3. Boellaard R, Oyen WJG, Hoekstra CJ, Boellaard R, Oyen WJ, Hoekstra CJ, Hoekstra OS, Visser EP, Willemsen AT, Arends B, Verzijlbergen FJ, Zijlstra J, Paans AM, Comans EF, Pruim J. The Netherlands protocol for standardisation and quantification of FDG whole body PET studies in multi-centre trials. Eur J Nucl Med Mol Imaging 2008;35:2320-2333.
4. Huang S-C. Anatomy of SUV. Nucl Med Biol 2000;27:643-646.
5. de Geus-Oei L-F, Oyen WJG. Predictive and prognostic value of FDG-PET. Cancer Imaging 2008;8:70-80.
6. Herrmann K, Krause BJ, Bundschuh RA, Dechow T, Schwaiger M. Monitoring response to therapeutic interventions in patients with cancer. Semin Nucl Med 2009;39:210-232.
7. Weber WA, ed. Personalizing cancer therapy with FDG PET: from RECIST to PERCIST. J Nucl Med 2009;50(Suppl 1):1-150.
8. Kramer-Marek G, Capala J. Can PET imaging facilitate optimization of cancer therapies? Current Pharmaceutical Design 2012;18:2657-2669.
9. Boellaard R, Krak NC, Hoekstra OS, Lammertsma AA. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. J Nucl Med 2004;45:1519-1527.
10. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. J Nucl Med 2007;48:932-945.
11. Harrison MR, George DJ. Better late than early: FDG-PET imaging in metastatic renal cell carcinoma. Clin Cancer Res 2011;17:5841-5843.
12. Dahlberg G. Statistical methods for medical and biological students. New York Interscience Publications; 1940.
13. Colton T. Statistics in Medicine. Little, Brown and Company. Boston 1974.
14. de Langen AJ, Vincent A, Velazquez LM, van 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.
15. Hatt M, Cheze-le Rest C, van Baardwijk A, Lambin P, Pradier O, Visvikis D. Impact of tumor size and tracer uptake heterogeneity in 18F-FDG PET and CT non-small cell lung cancer delineation. J Nucl Med 2011;52:1690-1697.
16. Hustinx R, Smith RJ, Benard F, Rosenthal DI, Machtay M, Farber LA, Alavi A. Dual time point fluorine-18 fluorodeoxyglucose positron emission tomography: a potential method to differentiate malignancy from inflammation and normal tissue in the head and neck. Eur J Nucl Med 1999;26:1345-1348.
17. Zhuang H, Pourdehnad M, Lambright ES, Yamamoto AJ, Lanuti M, Li P, Mozley PD, Rossman MD, Albelda SM, Alavi A. Dual time point 18F-FDG PET imaging for differentiating malignant from inflammatory processes. J Nucl Med 2001;42:1412-1417.
18. Schwartz J, Humm JL, Gonen M, Kalaigian H, Schoder H, Larson SM, Nehmeh SA. Repeatability of SUV measurements in serial PET. Med Phys 2011;38:2629-2638.
19. Lodge MA, Chaudhry MA, Wahl RL. Noise considerations for PET quantification using maximum and peak standardized uptake value. J Nucl Med 2012;53:1041-1047.
20. Burger IA, Huser DM, Burger C, von Schulthess GK, Buck A. Repeatability of FDG quantification in tumor imaging: averaged SUVs are superior to SUV_{max}. Nucl Med Biol 2012;39:666-670.