

SUV_{max} of 2.5 should not be embraced as a magic threshold for separating benign from malignant lesions

Thomas C. Kwee · Gang Cheng · Marnix G. E. H. Lam · Sandip Basu · Abass Alavi

Published online: 26 June 2013
© Springer-Verlag Berlin Heidelberg 2013

Lesions of an indeterminate nature are often encountered in images such as those from CT and MRI. Examples include solitary pulmonary nodules and various incidentalomas that may be found in any body region. FDG PET can be useful in differentiating benign from malignant lesions, given the fact that malignant lesions generally have a higher glycolytic rate and consequently higher FDG uptake. Both visual image interpretation and (semiquantitative) standardized uptake value (SUV) measurements may be used for this purpose. Some clinicians advocate the use of a maximum SUV (SUV_{max}) threshold of 2.5 to separate benign from malignant lesions, which is based on the results of several old studies that were published more than 10 years ago [1, 2]. However, this approach suffers from major shortcomings, which may lead to patient mismanagement. This communication aims to clarify several important issues on the validity of using an SUV_{max} threshold of 2.5 to differentiate benign from malignant lesions.

First, SUV measurements are affected by many parameters, including the equipment used, the physics, and biological factors. Partial volume and spillover effects, attenuation correction, the reconstruction method and parameters for scanner type, the count noise bias effect, radiotracer distribution time

(i.e. the time between radiotracer injection and imaging), competing transport effects, and body size all affect SUV measurements considerably [3]. Because of the variability in PET acquisition in different institutions [4], interinstitutional SUV measurements are likely to vary too. Consequently, thresholds such as the SUV_{max} of 2.5 that have been reported as diagnostically useful by research groups in some institutions [1, 2], may not be useful at all in other institutions if different FDG PET protocols are used. In this context, initiatives such as the “EANM procedure guidelines for tumour PET imaging”, which have been shown to significantly reduce variability in SUV across different centres [5], are of crucial importance.

One important factor that should not be overlooked is the effect of partial volume effects on SUV measurements [6, 7]. Because of the relatively low spatial resolution of PET, significant averaging of pixel intensities of lesions with the surrounding tissues occurs. Motion blurring (e.g. due to patient, cardiac and respiratory motion, and peristalsis) further leads to undesired averaging of pixel intensities of lesions with the surrounding tissues. If not corrected for, partial volume effects may lead to inaccurate (underestimated) measures of the true FDG activity, especially in small lesions [8–10]. Several methods can be used for partial volume correction, and these can be divided into methods applied at the regional level (e.g. use of recovery coefficients, geometric transfer matrix approach, and deconvolution) and methods applied at the pixel level (e.g. partition-based correction, multiresolution approach, fitting method, the so-called “maximum a posteriori” approach, and kinetic modelling) [9]. Partial volume correction has been shown to improve accuracy of SUV without decreasing (clinical) test–retest variability significantly, and it has a small but significant effect on observed tumour responses [10]. Unfortunately, no general, widely accepted solution to the partial volume effect problem has yet been found. There is an urgent need for a standard widely adopted method to deal with partial volume effects, because this will

T. C. Kwee · M. G. E. H. Lam
Department of Radiology and Nuclear Medicine, University
Medical Center Utrecht, Utrecht, The Netherlands

G. Cheng
Department of Radiology, Philadelphia VA Medical Center,
Philadelphia, PA, USA

G. Cheng · A. Alavi (✉)
Division of Nuclear Medicine, Department of Radiology, Hospital
of the University of Pennsylvania, 3400 Spruce St., 110 Donner
Bldg, Philadelphia, PA 19104, USA
e-mail: abass.alavi@uphs.upenn.edu

S. Basu
Radiation Medicine Center (Bhabha Atomic Research Center), Tata
Memorial Center Annexe, Bombay, India

accelerate the high potential of quantitative PET in oncology. Attempts are being made by both the industry and academia to develop robust partial volume correction methods that can be easily implemented in clinical practice [11, 12].

Another important, but frequently ignored, physical/biological factor that affects FDG uptake (and SUV measurements) is the radiotracer distribution time (i.e. the time between FDG injection and imaging). This is clearly demonstrated in the liver, which shows a 25–30 % decrease in SUV from 1 hour to 2–3 hours [13]. Thus, while focal hepatic activity with an SUV_{max} of 3.5–4.0 is common on 1-hour PET imaging (and regarded benign), it may be suspicious for malignancy if PET image acquisition was done 2–3 hours after FDG injection. Along with several background tissues (including blood pool, liver, spleen, lungs, pancreas, lymph nodes and skeletal muscle) [13], inflammatory lesions also tend to exhibit decreased FDG activity, while cancers (particularly aggressive, rapidly proliferating cancers) often show increasing FDG activity with increasing radiotracer distribution time [6, 14]. These phenomena form the basis for performing dual time-point or delayed PET imaging, which may improve lesion detection and characterization. However, it should also be noted that some granulomatous lesions show higher SUVs on delayed imaging [14, 15]. Nevertheless, the point here is that FDG uptake and washout from cells (which is regulated by the amount of glucose membrane receptors and the ratio of hexokinase to glucose-6-phosphatase activity within cells) is a dynamic process that varies among malignant, benign and background tissues [7, 13, 14]. Therefore, diagnostic performance is heavily dependent on the radiotracer distribution time that is applied. The use of a static SUV_{max} threshold of 2.5 has important limitations in this context.

Second, various nonmalignant lesions may cause increased FDG uptake that is much higher than the SUV_{max} of 2.5, most commonly inflammation and infection [16]. While infectious lesions in their acute phase are more easily diagnosed if they display typical CT or MRI features, chronic/subacute inflammatory lesions, such as granulomatous lesions, may cause considerable diagnostic difficulties. For example, it has been reported that FDG PET cannot distinguish malignant solitary pulmonary nodules from tuberculoma [15]. Conversely, several low-grade cancers exhibit an SUV_{max} that is lower than 2.5, as is explained in the next section. Thus, both the positive and negative predictive values for discriminating benign from malignant lesions will be far from optimal when using an SUV_{max} threshold of 2.5.

Third, the use of an FDG PET threshold dichotomizes tumours into two categories (either benign or malignant), but this simplification does not reflect the differences in biology of many cancers, and it is unlikely that such a dichotomization is useful for appropriate treatment planning and assessment of prognosis. Cancers of the same type can behave very differently in terms of histology and clinical behaviour (i.e.

degree of invasiveness, growth, metastatic potential, response to therapy and associated survival). Importantly, FDG PET can noninvasively assess the biological behaviour of a cancer, with aggressive cancers generally exhibiting higher FDG uptake than less aggressive cancers, and this relationship between FDG uptake and cancer aggressiveness can be observed over a continuous spectrum [17]. This concept applies to various cancers, including breast cancer, thyroid cancer, prostate cancer, lymphoma, neuroendocrine tumours, and many others [17]. Several low-grade cancers (typical examples include lung adenocarcinoma in situ, indolent lymphoma and carcinoid) may exhibit an SUV_{max} that is lower than 2.5 [17], and will erroneously be classified as benign when using this FDG PET threshold. On the other hand, it would also be wrong to consider such an FDG PET result as “false-negative” from the view point of tumour biology. FDG PET is true-negative in such cases, given the fact that lower levels of FDG uptake often correspond to histologically and clinically less-aggressive tumour behaviour [17]. Thus, instead of using a simplified and erroneous dichotomization, FDG PET results should be regarded as a spectrum with a very different meaning for each observation [17]. This FDG PET-based assessment of tumour biology will play an important role in future risk stratification models in many cancers.

In conclusion, SUV measurements are affected by many parameters that should be accounted for when using certain FDG PET thresholds to improved characterization of disease. Standardization can considerably reduce SUV variability and improve the utility of certain SUV thresholds among different institutions. Importantly, however, even when accurate and reproducible FDG quantification may be achieved, one can consider the use of FDG PET thresholds for tumour characterization as conceptually wrong because such a dichotomization completely ignores the clinically important information on tumour biology that is available from an FDG PET examination. Thus, SUV_{max} of 2.5 should not be embraced as a magic threshold for separating benign from malignant lesions.

References

1. Hain SF, Curran KM, Beggs AD, Fogelman I, O’Doherty MJ, Maisey MN. FDG-PET as a “metabolic biopsy” tool in thoracic lesions with indeterminate biopsy. *Eur J Nucl Med.* 2001;28:1336–40.
2. Beggs AD, Hain SF, Curran KM, O’Doherty MJ. FDG-PET as a “metabolic biopsy” tool in non-lung lesions with indeterminate biopsy. *Eur J Nucl Med Mol Imaging.* 2002;29:542–6.
3. Thie JA. Understanding the standardized uptake value, its methods, and implications for usage. *J Nucl Med.* 2004;45:1431–4.
4. Graham MM, Badawi RD, Wahl RL. Variations in PET/CT methodology for oncologic imaging at U.S. academic medical centers: an

- imaging response assessment team survey. *J Nucl Med.* 2011;52:311–7.
5. Boellaard R, O'Doherty MJ, Weber WA, Mottaghy FM, Lonsdale MN, Stroobants SG, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur J Nucl Med Mol Imaging.* 2010;37:181–200.
 6. Basu S, Alavi A. Partial volume correction of standardized uptake values and the dual time point in FDG-PET imaging: should these be routinely employed in assessing patients with cancer? *Eur J Nucl Med Mol Imaging.* 2007;34:1527–9.
 7. Basu S, Kwee TC, Torigian D, Saboury B, Alavi A. Suboptimal and inadequate quantification: an alarming crisis in medical applications of PET. *Eur J Nucl Med Mol Imaging.* 2011;38:1381–2.
 8. Hickeson M, Yun M, Matthies A, Zhuang H, Adam LE, Lacorte L, et al. Use of a corrected standardized uptake value based on the lesion size on CT permits accurate characterization of lung nodules on FDG-PET. *Eur J Nucl Med Mol Imaging.* 2002;29:1639–47.
 9. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumor imaging. *J Nucl Med.* 2007;48:932–45.
 10. Hoetjes NJ, van Velden FH, Hoekstra OS, Hoekstra CJ, Krak NC, Lammertsma AA, et al. Partial volume correction strategies for quantitative FDG PET in oncology. *Eur J Nucl Med Mol Imaging.* 2010;37:1679–87.
 11. Basu S, Saboury B, Torigian DA, Alavi A. Current evidence base of FDG-PET/CT imaging in the clinical management of malignant pleural mesothelioma: emerging significance of image segmentation and global disease assessment. *Mol Imaging Biol.* 2011;13:801–11.
 12. Musiek ES, Saboury B, Mishra S, Chen Y, Reddin JS, Newberg AB, et al. Feasibility of estimation of brain volume and 2-deoxy-2-(18)F-fluoro-D-glucose metabolism using a novel automated image analysis method: application in Alzheimer's disease. *Hell J Nucl Med.* 2012;15:190–6.
 13. Cheng G, Alavi A, Lim E, Werner TJ, Del Bello CV, Akers SR. Dynamic changes of FDG uptake and clearance in normal tissues. *Mol Imaging Biol.* 2013;15:345–52.
 14. Cheng G, Torigian DA, Zhuang H, Alavi A. When should we recommend use of dual time-point and delayed time-point imaging techniques in FDG PET? *Eur J Nucl Med Mol Imaging.* 2013;40:779–87.
 15. Sathekge MM, Maes A, Pottel H, Stoltz A, van de Wiele C. Dual time-point FDG PET-CT for differentiating benign from malignant solitary pulmonary nodules in a TB endemic area. *S Afr Med J.* 2010;100:598–601.
 16. Shreve PD, Anzai Y, Wahl RL. Pitfalls in oncologic diagnosis with FDG PET imaging: physiologic and benign variants. *Radiographics.* 1999;19:61–77.
 17. Kwee TC, Basu S, Saboury B, Ambrosini V, Torigian DA, Alavi A. A new dimension of FDG-PET interpretation: assessment of tumor biology. *Eur J Nucl Med Mol Imaging.* 2011;38:1158–70.