Receiver operating characteristic (ROC) curve for classification of ¹⁸F-NaF uptake on PET/CT*

Curva de característica de operação do receptor (ROC) para classificação da captação de ¹⁸F-NaF em PET/CT

Agnes Araujo Valadares¹, Paulo Schiavom Duarte², Giovanna Carvalho³, Carla Rachel Ono², George Barberio Coura-Filho², Heitor Naoki Sado², Marcelo Tatit Sapienza⁴, Carlos Alberto Buchpiguel⁵

Valadares AA, Duarte PS, Carvalho G, Ono CR, Coura-Filho GB, Sado HN, Sapienza MT, Buchpiguel CA. Receiver operating characteristic (ROC) curve for classification of ¹⁸F-NaF uptake on PET/CT. Radiol Bras. 2016 Jan/Fev;49(1):12–16.

Abstract Objective: To assess the cutoff values established by ROC curves to classify ¹⁸F-NaF uptake as normal or malignant.

Materials and Methods: PET/CT images were acquired 1 hour after administration of 185 MBq of ¹⁸F-NaF. Volumes of interest (VOIs) were drawn on three regions of the skeleton as follows: proximal right humerus diaphysis (HD), proximal right femoral diaphysis (FD) and first vertebral body (VB1), in a total of 254 patients, totalling 762 VOIs. The uptake in the VOIs was classified as normal or malignant on the basis of the radiopharmaceutical distribution pattern and of the CT images. A total of 675 volumes were classified as normal and 52 were classified as malignant. Thirty-five VOIs classified as indeterminate or nonmalignant lesions were excluded from analysis. The standardized uptake value (SUV) measured on the VOIs were plotted on an ROC curve for each one of the three regions. The area under the ROC (AUC) as well as the best cutoff SUVs to classify the VOIs were calculated. The best cutoff values were established as the ones with higher result of the sum of sensitivity and specificity.

Results: The AUCs were 0.933, 0.889 and 0.975 for UD, FD and VB1, respectively. The best SUV cutoffs were 9.0 (sensitivity: 73%; specificity: 99%), 8.4 (sensitivity: 79%; specificity: 94%) and 21.0 (sensitivity: 93%; specificity: 95%) for UD, FD and VB1, respectively. **Conclusion:** The best cutoff value varies according to bone region of analysis and it is not possible to establish one value for the whole body.

Keywords: ¹⁸F-NaF PET/CT; ROC curve; Cutoff values; Normal uptake; Malignant uptake.

Resumo Objetivo: Acessar valores de corte estabelecidos pela curva ROC para classificar a captação de ¹⁸F-NaF como normal ou maligna. Materiais e Métodos: Imagens de PET/CT foram realizadas 1 hora após a administração de 185 MBq de ¹⁸F-NaF e volumes de interesse (VOIs) foram desenhados em três regiões do esqueleto: diáfise umeral proximal direita (UD), diáfise femoral proximal direita (FD) e corpo da primeira vértebra lombar (VB1), em 254 pacientes, totalizando 762 VOIs. A captação nos VOIs foi classificada como normal ou maligna baseada no padrão de distribuição do radiofármaco e nas imagens de CT. Um total de 675 volumes foi classificado como normais e 52 como malignos. Trinta e cinco VOIs classificados como indeterminados ou lesões não malignas foram excluídos da análise. Os valores de captação (SUVs) medidos nos VOIs foram plotados em uma curva ROC para cada uma das três regiões. Foi calculada a área sob a curva (AUC), bem como os valores de SUV mais adequados para a classificação dos VOIs (maior resultado da soma da sensibilidade e especificidade).

Resultados: As AUCs foram 0,933, 0,889 e 0,975 para UD, FD e VB1, respectivamente. Os valores de corte mais adequados de SUV foram 9,0 (sensibilidade: 73%; especificidade: 99%), 8,4 (sensibilidade: 79%; especificidade: 94%) e 21,0 (sensibilidade: 93%; especificidade: 95%)para UD, FD e VB1, respectivamente.

Conclusão: O valor de corte de SUV mais adequado varia de acordo com a região óssea em análise e não é possível estabelecer um valor adequado para todo o esqueleto.

Unitermos: ¹⁸F-NaF PET/CT; Curva ROC; Valores de corte; Captação normal; Captação maligna.

* Study developed at Center of Nuclear Medicine – Instituto de Radiologia do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (InRad/HC-FMUSP), and at Service of Nuclear Medicine – Instituto do Câncer do Estado de São Paulo Octavio Frias de Oliveira (Icesp), São Paulo, SP, Brazil.

1. Nuclear Physician, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HC-FMUSP), São Paulo, SP, Brazil.

2. PhD, Physician Assistant, Service of Nuclear Medicine – Instituto do Câncer do Estado de São Paulo Octavio Frias de Oliveira (Icesp), São Paulo, SP, Brazil.

3. Physician Assistant, Service of Nuclear Medicine – Instituto do Câncer do Estado de São Paulo Octavio Frias de Oliveira (Icesp), São Paulo, SP, Brazil.

 Private Docent, Professor, Department of Radiology and Oncology – Faculdade de Medicina da Universidade de São Paulo (FMUSP), São Paulo, SP, Brazil.

INTRODUCTION

Sodium fluoride (¹⁸F-NaF) is a highly sensitive boneseeking PET tracer used to detect skeletal abnormalities. The uptake mechanism of ¹⁸F-fluoride resembles that of ^{99m}Tc-

Private Docent, Full Professor, Department of Radiology and Oncology – Faculdade de Medicina da Universidade de São Paulo (FMUSP), São Paulo, SP, Brazil.

Mailing Address: Dra. Agnes Araujo Valadares. Avenida Dr. Arnaldo, 251, 4º Subsolo (Medicina Nuclear), Cerqueira César. São Paulo, SP, Brazil, 01246-000. E-mail: agnesvaladares@me.com.

Received November 10, 2014. Accepted after revision May 7, 2015.

MDP with better pharmacokinetic characteristics including faster blood clearance and two-fold higher uptake in bone⁽¹⁾.

Over the last years there has been a renewed clinical interest in the use of ¹⁸F-NaF as a bone scanning agent⁽²⁾. Reasons for this resurgence include periodic worldwide shortages of ^{99m}Tc needed for conventional bone scanning agents⁽³⁾, and the improved sensitivity⁽⁴⁻⁶⁾ and quantitative potential of ¹⁸F-NaF PET/CT^(7,8) as compared with ^{99m}Tc-based conventional bone scans. In a pilot study, ¹⁸F-NaF was also demonstrated to be superior to ¹⁸F-FDG PET/CT and magnetic resonance imaging (MRI)⁽⁹⁾.

Standardized uptake value (SUV), which averages tracer uptake with respect to the injected dose and body weight, is the most widely used PET index for assessment of tracer uptake in the routine clinical practice because it does not require blood sampling and is obtained by static PET acquisition^(10–12). Although SUV has been used predominantly for ¹⁸F-FDG PET imaging quantification, research reports demonstrate that SUVs can detect significant metabolic change in individual metastatic lesions at ¹⁸F-NaF PET images, even in cases where visual evaluation reveals little if any difference⁽⁷⁾. Additionally, in the field of oncology, an earlier identification of metastatic involvement is feasible and SUV measurement may provide such information in cases where it is important to assess whether a patient is responding to treatment^(7,8).

However, as far as we know, the normal SUV range for ¹⁸F-NaF PET/CT has been rarely analyzed in the scientific literature⁽¹³⁾ as well as the most appropriate cutoff values to distinguish normal from malignant bone uptake.

The present study was aimed at establishing the normal SUV range for ¹⁸F-NaF PET/CT and assessing the most appropriate cutoff values to classify ¹⁸F-NaF uptake as normal or malignant.

MATERIALS AND METHODS

Patient population

The present study was approved by the committee for Ethics in Research of School of Medicine of Universidade de São Paulo.

In this cross-sectional study, the images of the first 254 patients submitted to ¹⁸F-NaF PET/CT scans in the authors' department were analyzed. The following patients' characteristics were analyzed: weight, height, body mass index (BMI), age and sex.

Main tumor types were the following: 96, breast cancer; 28, prostate cancer; 16, lung cancer; 9, colorectal cancer; 8, melanoma; 8, liver cancer. Thirty-four patients presented with other tumors and in 55 patients the cancer type was not recorded.

PET/CT image acquisition

The patients were injected with 111 to 203 MBq (mean, 141 MBq) of ¹⁸F-NaF. Approximately 60 min after injection, all patients underwent whole-body (vertex to toes) three-

dimensional PET/CT. Images were acquired on a Discovery 690 GE with the time of flight technique (General Electric). The PET image reconstruction was performed using iterative technique with 24 subsets for all studies. Low-dose CT transmission scans were obtained (10 to 30 mAs) for attenuation correction. Other CT images parameters were the following: 120 kVp, 0.5-s rotation time; 1.375 pitch; and axial slice thickness of 3.75 mm. Emission PET images were obtained at 1 min per bed position (15 cm slice thickness with 3 cm of overlapping), with 13 to 15 bed positions per study.

Image analysis

Cuboid volumes of interest (VOIs) with edges around 2 to 3 cm were drawn on three bone regions, as follows: proximal right humerus diaphysis (HD), proximal right femoral diaphysis (FD) and first vertebral body (VB1) in the 254 patients, totalling 762 VOIs (Figure 1). The uptake in the VOIs was classified as normal or malignant on the basis of the radiopharmaceutical distribution pattern and on the CT images. Such a classification was established by consensus between two nuclear medicine physicians. A total of 675 volumes were classified as normal, and 52 were classified as malignant. Thirty-five VOIs classified as indeterminate or nonmalignant lesions were excluded from analysis.



Figure 1. Example of VOIs drawn on three bone regions: proximal right humerus diaphysis; proximal right femoral diaphysis; and first vertebral body.

Statistical analysis

For the whole study sample, the patients' characteristics (weight, height, BMI, age and sex) were described using mean and standard deviation.

The means and standard deviations of the maximal SUVs in each one of the three VOIs were calculated for the normal and malignant uptakes. ANOVA analysis was used to compare the maximal SUVs means among the three VOIs. The Tukey test was used to compare the maximal SUVs means between all pairs of regions.

The SUVs were also plotted on an ROC curve for each one of the three VOIs. The area under the ROC (AUC) as well as the most appropriate cutoff SUVs were calculated to classify the VOIs either as normal or malignant. The most appropriate cutoff values were established as the ones with higher result of the sum of sensitivity and specificity.

The statistical analyses were performed using Excel[®] worksheets and the SPSS 16.0[®].

RESULTS

The authors evaluated 254 patients, 66% of them, women. Mean and standard deviation of weight and height were, respectively, 69 ± 15 kg and 159 ± 8 cm. The group had a mean BMI of 27 ± 6 kg/m² and an average age of 60 ± 14 years.

Means and the standard deviations of the maximum SUVs for each one of the three VOIs for normal and malignant uptakes are shown on Table 1.

VOI	Classification	п	SUV (mean \pm SD)
VB1	Normal	207	14.4 ± 3.7
VB1	Malignant	27	39.3 ± 18.6
HD	Normal	236	3.8 ± 1.4
HD	Malignant	11	11.4 ± 4.9
FD	Normal	232	5.4 ± 2.0
FD	Malignant	14	14.4 ± 9.0

The ANOVA analysis comparing the means among the three VOIs in the regions classified as normal demonstrated a statistically significant difference (p < 0.01). The Tukey test demonstrated a statistically significant difference between all pairs of regions as compared with each other (p < 0.01).

The AUCs were 0.933, 0.889 and 0.975 for UD, FD and VB1, respectively (Figure 2). The most appropriate SUV cutoffs were 9.0 (sensitivity: 73%; specificity: 99%), 8.4 (sensitivity: 79%; specificity: 94%), and 21 (sensitivity: 93%; specificity: 95%) for UD, FD and VB1, respectively.

DISCUSSION

 $^{18}\mbox{F-NaF}$ was introduced by Blau et al. $^{(14)}$ in 1962 as an imaging agent for bone lesions. Data from multiple small studies have shown that $^{18}\mbox{F-NaF}$ PET produces bone scans



Figure 2. The ROC curves for the three bone regions: proximal right humerus diaphysis (HD), proximal right femoral diaphysis (FD) and first vertebral body (VB1).

with higher sensitivity and specificity than 99m Tc-based bone scans^(15–19), and has also demonstrated to be superior to 18 F-FDG PET/CT and MRI⁽⁹⁾.

The SUV is commonly used as a relative measure of FDG uptake⁽²⁰⁾. The use of SUV is currently a routine in the oncological clinical practice of ¹⁸F-FDG PET/CT imaging. Ideally, the use of SUV removes the variability caused by differences in patients' size and the amount of injected activity⁽²¹⁾ and facilitates comparisons between patients.

Although the practice of using SUV thresholds for diagnosis is not widely accepted⁽²²⁾ and despite the fact that it has been repeatedly demonstrated that the use of SUV thresholds characterizes an uptake region as benign or malignant is often invalid, there are some situations where the use of SUV may be useful⁽²¹⁾. The SUV could be useful in cases where PET quantitative imaging is necessary in clinical research, clinical trials, and researches for development of drugs since it allows the standardization of imaging analysis⁽²³⁾. The SUV could also be useful in therapy monitoring scans as independent measures of changes in metabolic activity can provide an alternative approach to assess response to therapy⁽²¹⁾. Even for diagnostic purposes, the use of SUV thresholds could be valuable in some situations. For example, in cases where the FDG uptake in a tissue is no greater than the uptake in adjacent reference tissue and where the pretest likelihood of malignancy is low, it is considered to be safe to adopt an expectant strategy⁽²⁴⁾.

Most of the studies approaching the use of SUV in nuclear medicine were developed for ¹⁸F-FDG PET/CT, even in the field of skeletal pathologies⁽²⁵⁾. However, recent reports describe the use of SUV to analyze bone pathologies at ¹⁸F-NaF PET/CT. A recent scientific article has demonstrated that ¹⁸F-fluoride PET scans using SUV measurements have the potential to be a diagnostic tool in otosclerosis⁽²⁶⁾. It was also demonstrated that SUV can be considered as being as effective and accurate as kinetic modeling in measuring the response of pagetic bones to bisphosphonates by means of ¹⁸F-fluoride PET⁽²⁷⁾. Despite these studies, there is still little experience with the use of reference values of SUV in bone metabolism⁽¹³⁾.

In the present study, the authors analyzed the normal range of SUVs in a group of patients as well as tried to define the most appropriate values to distinguish between normal and malignant uptakes. The authors have observed that the normal SUVs range varies among the analyzed bone regions and that it is lower in the humerus than in the femur, additionally, it is lower in these two bones as compared with the first lumbar vertebral body. The authors also observed that, although the AUC is high in the three regions, the most appropriate cutoff value to classify bone uptake as normal or malignant varies among these regions and it is not possible to establish one single value for the whole body.

In a previous study, Win et al. have analyzed the ¹⁸Ffluoride maximum SUVs in the skeleton⁽¹³⁾. Their findings were similar to the findings of the present study. According to their study, various skeletal sites have different normal SUVs and vertebral bodies tend to show the highest values. They deeply discussed the causes of these differences in the SUVs and related some possible explanations as a higher flow in the spine as compared with the proximal femur⁽²⁸⁾, a greater bone turnover at the spine than at other skeletal sites⁽²⁹⁾, predominance of cortical bone in the humerus that has a lower level of bone metabolism as compared with the lumbar spine which is rich in trabecular bone⁽³⁰⁾, and the mechanical stress that lumbar vertebra is subjected as it is a primary weight bearing bone⁽³¹⁾. However, their mean SUVs in the same bone regions analyzed in our study are lower: 7.16 in the VB1, 2.16 in the femur and 1,71 in the humerus. The reasons for such differences between the two studies are still to be known and should be an object of fur-

ther analysis. But, one must note that SUVs also depend on the measurement instruments and reconstruction methods^(32,33); therefore, differences between the images acquisition and reconstruction protocols could partially explain such discrepancies. A more recent study developed by Sabbah et al.⁽³⁴⁾ also analyzed ¹⁸F-fluoride SUVs in the skeleton. They found mean values of 10.07, 2.12 and 3.01 for maximum SUVs in the lumbar spine, humerus and femur, respectively. These values are also compatible with the results of the present study as they demonstrated that the mean SUVs vary among different bone structures and that the SUV in the lumbar spine is higher than the ones in the femur and humerus. Once again, the absolute values obtained are different from our results but they are also different from the values obtained by Win et al.⁽¹³⁾. The reasons for such differences between the three studies should be the same discussed above: variability in the images acquisition and reconstruction protocols.

Finally, an important aspect to be discussed is the use of ROC curves and threshold values in the present study. ROC curves were used in two aspects, namely, to define the most appropriate threshold to differentiate malignant from normal uptake areas in the three bone regions and to analyze the overall ¹⁸F-fluoride PET/CT accuracy to differentiate malignant from normal uptake in those regions. The definition of the thresholds was important to demonstrate that those values vary among the analyzed regions and corroborated the authors' findings that the normal SUVs range should be established for each one of the bone regions. However, is important to observe that the threshold values established in the present study should not be used to classify lesions as malignant since we did not analyze the SUVs in a group of benign lesions. Moreover, in the authors' clinical experience, the ¹⁸F-fluoride uptake intensity in benign bone lesions could be as intense as the uptake in malignant lesions, and some malignant lesions may present a very low uptake. Therefore, for the time being, SUVs in ¹⁸F-fluoride PET/CT studies should be used at most for follow-up purposes and not to classify the lesions as malign or benign.

CONCLUSION

The SUVs normal range and the most appropriate cutoff value to differentiate normal from malignant bone uptake vary according to bone region of analysis and it is not possible to establish a single value for the whole skeleton.

REFERENCES

- Segall G, Delbeke D, Stabin MG, et al. SNM practice guideline for sodium 18F-fluoride PET/CT bone scans 1.0. J Nucl Med. 2010; 51:1813–20.
- Grant FD, Fahey FH, Packard AB, et al. Skeletal PET with 18Ffluoride: applying new technology to an old tracer. J Nucl Med. 2008;49:68–78.
- 3. Perkins A, Hilson A, Hall J. Global shortage of medical isotopes threatens nuclear medicine services. BMJ. 2008;337:a1577.
- 4. Even-Sapir E, Metser U, Flusser G, et al. Assessment of malignant skeletal disease: initial experience with 18F-fluoride PET/CT and

Valadares AA et al. / ROC curve for classification of ¹⁸F-NaF uptake on PET/CT

comparison between 18F-fluoride PET and 18F-fluoride PET/CT. J Nucl Med. 2004;45:272–8.

- Even-Sapir E, Metser U, Mishani E, et al. The detection of bone metastases in patients with high-risk prostate cancer: 99mTc-MDP planar bone scintigraphy, single- and multi-field-of-view SPECT, 18F-fluoride PET, and 18F-fluoride PET/CT. J Nucl Med. 2006; 47:287–97.
- Krüger S, Buck AK, Mottaghy FM, et al. Detection of bone metastases in patients with lung cancer: 99mTc-MDP planar bone scintigraphy, 18F-fluoride PET or 18F-FDG PET/CT. Eur J Nucl Med Mol Imaging. 2009;36:1807–12.
- 7. Cook G Jr, Parker C, Chua S, et al. 18F-fluoride PET: changes in uptake as a method to assess response in bone metastases from castrate-resistant prostate cancer patients treated with 223Ra-chloride (Alpharadin). EJNMMI Res. 2011;1:4.
- Doot RK, Muzi M, Peterson LM, et al. Kinetic analysis of 18Ffluoride PET images of breast cancer bone metastases. J Nucl Med. 2010;51:521–7.
- Iagaru A, Young P, Mittra E, et al. Pilot prospective evaluation of 99mTc-MDP scintigraphy, 18F NaF PET/CT, 18F FDG PET/ CT and whole-body MRI for detection of skeletal metastases. Clin Nucl Med. 2013;38:e290–6.
- Brenner W, Vernon C, Muzi M, et al. Comparison of different quantitative approaches to 18F-fluoride PET scans. J Nucl Med. 2004; 45:1493–500.
- Puri T, Blake GM, Frost ML, et al. Comparison of six quantitative methods for the measurement of bone turnover at the hip and lumbar spine using 18F-fluoride PET-CT. Nucl Med Commun. 2012; 33:597–606.
- Even-Sapir E, Mishani E, Flusser G, et al. 18F-Fluoride positron emission tomography and positron emission tomography/computed tomography. Semin Nucl Med. 2007;37:462–9.
- Win AZ, Aparici CM. Normal SUV values measured from NaF18-PET/CT bone scan studies. PLoS One. 2014;9:e108429.
- Blau M, Nagler W, Bender MA. Fluorine-18: a new isotope for bone scanning. J Nucl Med. 1962;3:332–4.
- Hetzel M, Arslandemir C, König HH, et al. F-18 NaF PET for detection of bone metastases in lung cancer: accuracy, cost-effectiveness, and impact on patient management. J Bone Miner Res. 2003;18:2206–14.
- Hoh CK, Hawkins RA, Dahlbom M, et al. Whole body skeletal imaging with [18F]fluoride ion and PET. J Comput Assist Tomogr. 1993;17:34–41.
- Langsteger W, Heinisch M, Fogelman I. The role of fluorodeoxyglucose, 18F-dihydroxyphenylalanine, 18F-choline, and 18F-fluoride in bone imaging with emphasis on prostate and breast. Semin Nucl Med. 2006;36:73–92.
- 18. Schirrmeister H, Guhlmann A, Kotzerke J, et al. Early detection and accurate description of extent of metastatic bone disease in breast

cancer with fluoride ion and positron emission tomography. J Clin Oncol. 1999;17:2381–9.

- Schirrmeister H, Glatting G, Hetzel J, et al. Prospective evaluation of the clinical value of planar bone scans, SPECT, and (18)F-labeled NaF PET in newly diagnosed lung cancer. J Nucl Med. 2001;42:1800–4.
- 20. Thie JA. Understanding the standardized uptake value, its methods, and implications for usage. J Nucl Med. 2004;45:1431–4.
- Kinahan PE, Fletcher JW. Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy. Semin Ultrasound CT MR. 2010;31: 496–505.
- 22. Keyes JW Jr. SUV: standard uptake or silly useless value? J Nucl Med. 1995;36:1836–9.
- Hallett WA, Maguire RP, McCarthy TJ, et al. Considerations for generic oncology FDG-PET/CT protocol preparation in drug development. Idrugs. 2007;10:791–6.
- 24. Gambhir SS. Molecular imaging of cancer with positron emission tomography. Nat Rev Cancer. 2002;2:683–93.
- Duarte PS, Zhuang H, Castellucci P, et al. The receiver operating characteristic curve for the standard uptake value in a group of patients with bone marrow metastasis. Mol Imaging Biol. 2002;4:157–60.
- Waterval JJ, Vallinga M, Brans B, et al. 18F-fluoride PET/CT scan for quantification of bone metabolism in the inner ear in patients with otosclerosis – a pilot study. Clin Nucl Med. 2013;38:677–85.
- Installé J, Nzeusseu A, Bol A, et al. (18)F-fluoride PET for monitoring therapeutic response in Paget's disease of bone. J Nucl Med. 2005;46:1650–8.
- Puri T, Frost ML, Curran KM, et al. Differences in regional bone metabolism at the spine and hip: a quantitative study using (18)Ffluoride positron emission tomography. Osteoporos Int. 2013;24: 633–9.
- 29. Cheng C, Heiss C, Dimitrakopoulou-Strauss A, et al. Evaluation of bone remodeling with (18)F-fluoride and correlation with the glucose metabolism measured by (18)F-FDG in lumbar spine with time in an experimental nude rat model with osteoporosis using dynamic PET-CT. Am J Nucl Med Mol Imaging. 2013;3:118–28.
- Frost ML, Siddique M, Blake GM, et al. Differential effects of teriparatide on regional bone formation using (18)F-fluoride positron emission tomography. J Bone Miner Res. 2011;26:1002– 11.
- Histed SN, Lindenberg ML, Mena E, et al. Review of functional/ anatomical imaging in oncology. Nucl Med Commun. 2012;33: 349–61.
- 32. Boellaard R. Standards for PET image acquisition and quantitative data analysis. J Nucl Med. 2009;50 Suppl 1:11S–20S.
- Fahey FH, Kinahan PE, Doot RK, et al. Variability in PET quantitation within a multicenter consortium. Med Phys. 2010;37: 3660–6.