European Journal of Nuclear Medicine and Molecular Imaging (2019) 46:421–428
https://doi.org/10.1007/s00259-018-4151-8

ORIGINAL ARTICLE

Why harmonization is needed when using FDG PET/CT as a prognosticator: demonstration with EARL-compliant SUV as an independent prognostic factor in lung cancer

Benjamin Houdu[1] · Charline Lasnon[2,3] · Idlir Licaj[4] · Guy Thomas[5] · Pascal Do[6] · Anne-Valerie Guizard[7] · Cédric Desmonts[1] · Nicolas Aide[1,3]

Received: 3 May 2018 / Accepted: 28 August 2018 / Published online: 14 September 2018

Abstract

Background To determine EARL-compliant prognostic SUV thresholds in a mature cohort of patients with locally advanced NSCLC, and to demonstrate how detrimental it is to use a threshold determined on an older-generation PET system with a newer PET/CT machine, and vice versa, or to use such a threshold with non-harmonized multicentre pooled data.

Materials and methods This was a single-centre retrospective study including 139 consecutive stage IIIA-IIIB patients. PET data were acquired as per the EANM guidelines and reconstructed with unfiltered point spread function (PSF) reconstruction. Subsequently, a 6.3 mm Gaussian filter was applied using the EQ.PET (Siemens Healthineers) methodology to meet the EANM/EARL harmonizing standards (PSF_{EARL}). A multicentre study including non-EARL-compliant systems was simulated by randomly creating four groups of patients whose images were reconstructed with unfiltered PSF and PSF with Gaussian post-filtering of 3, 5, and 10 mm. Identification of optimal SUV thresholds was based on a two-fold cross-validation process that partitioned the overall sample into learning and validation subsamples. Proportional Cox hazards models were used to estimate age-adjusted and multivariable-adjusted hazard ratios (HRs) and their 95% confidence intervals. Kaplan–Meier curves were compared using the log rank test.

Results Median follow-up was 28 months (1–104 months). For the whole population, the estimated overall survival rate at 36 months was 0.39 [0.31–0.47]. The optimal SUV_{max} cutoff value was 25.43 (95% CI: 23.41–26.31) and 8.47 (95% CI: 7.23–9.31) for the PSF and for the EARL-compliant dataset respectively. These SUV_{max} cutoff values were both significantly and independently associated with lung cancer mortality; HRs were 1.73 (1.05–2.84) and 1.92 (1.16–3.19) for the PSF and the EARL-compliant dataset respectively. When (i) applying the optimal PSF SUV_{max} cutoff on an EARL-compliant dataset and the optimal EARL SUV_{max} cutoff on a PSF dataset or (ii) applying the optimal EARL compliant SUV_{max} cutoff to a simulated multicentre dataset, the tumour SUV_{max} was no longer significantly associated with lung cancer mortality.

Conclusion The present study provides the PET community with an EARL-compliant SUV_{max} as an independent prognosticator for advanced NSCLC that should be confirmed in a larger cohort, ideally at other EARL accredited centres, and highlights the need to harmonize PET quantitative metrics when using them for risk stratification of patients.

Charline Lasnon and Idlir Licaj equally contributed to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00259-018-4151-8) contains supplementary material, which is available to authorized users.

✉ Nicolas Aide
aide-n@chu-Caen.fr

1 Nuclear Medicine Department, University Hospital, Caen, France

2 Nuclear Medicine Department, François Baclesse Cancer Centre, Caen, France

3 INSERM ANTICIPE, Normandie University, Caen, France

4 Clinical Research Department, François Baclesse Cancer Centre, Caen, France

5 Medical Informatics Department, François Baclesse Cancer Centre, Caen, France

6 Lung Cancer Unit, François Baclesse Cancer Centre, Caen, France

7 Regional Cancer Registry, François Baclesse Cancer Centre, Caen, France

Keywords Non-small-cell lung cancer · Survival · FDG · PET/CT · Prognosticator · Harmonization · EARL accreditation program

Introduction

Lung cancer is the leading cause of cancer-related deaths worldwide. The clinical management of lung cancer may combine chemotherapy, molecular-targeted therapy, radiotherapy and surgery used alone or in combination. The choice of the treatments and their sequence depends on accurate staging, particularly in terms of the nodal status in the thorax and the presence of distant metastases. In that setting, 18F-fluorodeoxyglucose positron emission tomography/computed tomography (FDG PET/CT) has become a standard tool for initial staging [1]. In addition to its role in staging, FDG PET/CT is increasingly used for radiotherapy planning [2], therapeutic monitoring [3] and as a prognosticator [4]. The last of these roles of FDG PET/CT relies on the use of a standardized uptake value (SUV) metric, such as SUV_{max}, for risk stratification of patients, i.e., to predict the overall survival (OS) in lung cancer patients.

In recent years, several studies have demonstrated the value of SUV as a prognostic factor in patients with non-small-cell lung cancer (NSCLC). However, the SUV thresholds for risk stratification of patients were determined before new algorithms, such as point spread function (PSF) modelling, were released. PSF, which is available from the major PET vendors, has been shown to improve the diagnostic performance of FDG PET/CT in various cancers. This algorithm also significantly increases the SUV metrics compared with standard algorithms such as ordered subset maximization and expectation (OSEM) reconstruction. As a consequence, optimal thresholds determined on older generation PET/CT systems cannot be applied to centres using newer systems equipped with PSF, and vice versa. Moreover, the heterogeneity in the quantitative capability that exists between PET systems equipped with conventional reconstruction algorithms and those with advanced algorithms is worsened by the fact that even centres using similar equipment tend to use different reconstruction parameters, which affects SUV metrics [5].

For those reasons, harmonization programmes, such as the EANM/EARL accreditation programme, have been developed to facilitate the use of quantitative PET parameters in multicentre trials or at sites equipped with several PET systems [5, 6]. Briefly, these harmonizing standards are based on the accurate calibration of PET systems, compliance to guidelines for PET imaging and the use of a reconstruction with a filter chosen so that quantitative metrics extracted from PET images are comparable between all accredited centres. Most of the EARL-accredited centres tend to use two reconstructions: one optimized for diagnostic purposes and one used for quantitative purposes, meeting harmonizing standards [5]. More recently, a software solution has been released that allows the interpretation of PET images optimized for diagnostics while providing harmonized SUVs without the need for a second reconstruction [7, 8]. In addition to being more convenient than the use of two reconstructions, this software retrospectively allows analyses of data acquired and stored before the era of harmonization programmes, providing the opportunity to produce mature series of EARL-compliant PET data, if other sources of variability in SUV measurements are controlled.

The aim of the present study was to provide the PET community with an EARL-compliant prognostic SUV threshold for locally advanced NSCLC. The second aim was to demonstrate how it is detrimental to use a threshold determined on an older-generation PET system with a newer PET/CT machine, and vice versa. Also investigated was the application of an EARL-compliant threshold to multicentre pooled data from various PET reconstruction algorithms and parameters, including systems not meeting the EARL-harmonizing standards.

Materials and methods

Patients' selection

A crosswise analysis between the database from our picture archiving and communications system and that of the Regional Cancer Registry was performed to identify all biopsy-proven non-small-cell lung cancer patients who were referred to our regional comprehensive cancer centre between June 2008 and June 2014 and presented at the thoracic multidisciplinary meeting. We selected patients with stage IIIA and IIIB NSCLC as determined by the multidisciplinary tumour board. Patients who did not have a baseline PET scan as part of their diagnostic workup, who were referred for recurrence of disease, and who had a known history of other malignancies were excluded (Fig. 1).

To have a minimum setback of 36 months for each patient, we decided to end the observational period on June 1st, 2017. French regulations do not mandate ethical committee approval for observational studies without any additional therapies or monitoring procedures. Nevertheless, we sought approval for the collection of data for our study at the National Commission on Informatics and Liberty (CNIL), with the registration n° 2,084,596 v 0.

Calibration and cross-calibration of the PET/CT system

The calibration of the PET system was performed daily with a ^{68}Ge cylinder with a known radioactive concentration. A cross-calibration procedure was performed every 3 months during the

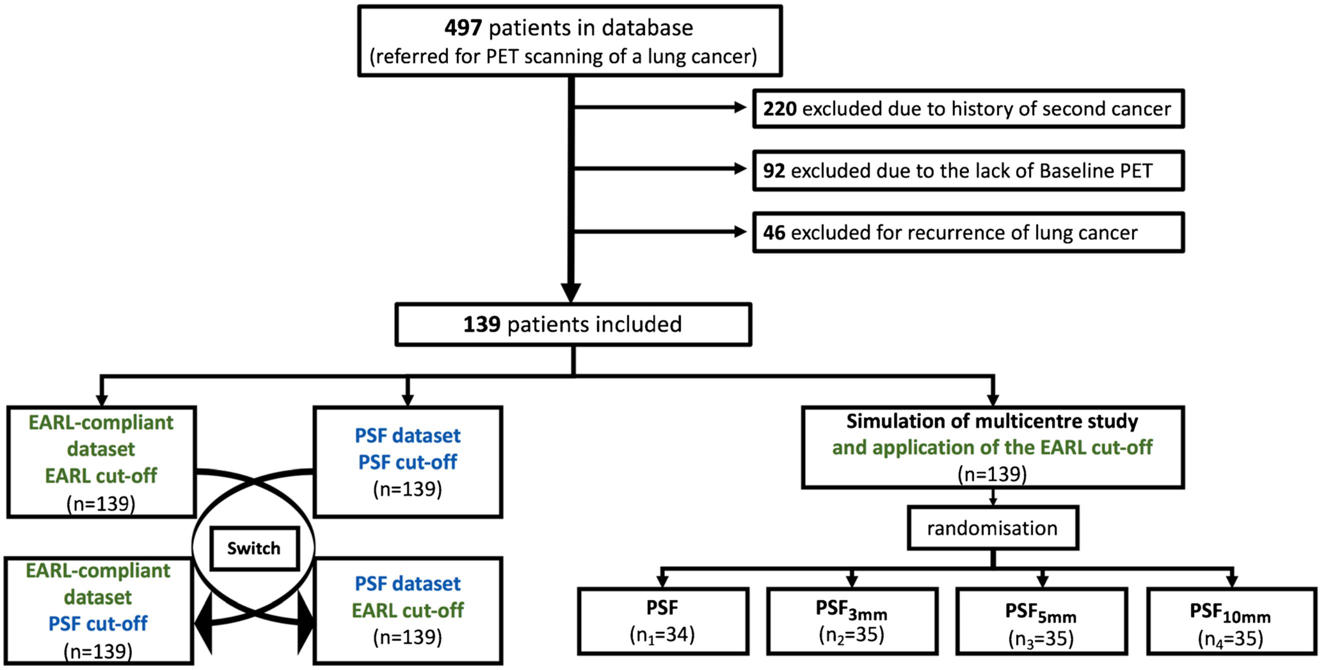

Fig. 1 Flow chart of patient selection and the conducted analyses

present study, as per the EANM guidelines [9]. The range of the cross-calibration factors was between 0.985 and 1.01.

FDG PET/CT acquisition

PET examinations were performed as per the EANM guidelines for PET tumour imaging. Patients were fasted for at least 6 h before intravenous injection of ^{18}F-FDG. All PET studies were tentatively performed after 60 ± 5 min rest in a warm room on a Biograph TrueV system (Siemens Healthineers) with a 6-slice spiral CT component.

First, a free-breathing CT acquisition was performed, using the following parameters: 60 mAs, 130 kVp, pitch 1, and 6 × 2 mm collimation. Then, a PET emission acquisition was performed with time per bed acquisitions of 160 s and 220 s for normal weight (BMI ≤ 25 kg/m^2) and overweight patients (BMI > 25 kg/m^2) respectively. Patients were scanned from the skull base to the mid-thighs.

Raw data were reconstructed with a point spread function (PSF) reconstruction algorithm (HD; TrueX, Siemens Healthineers), with three iterations and 21 subsets and no filtering. The matrix size was 168 × 168 voxels, resulting in voxels of 4.07 × 4.07 × 4.07 mm. Scatter and CT attenuation corrections were applied based on the CT scan.

PET analysis

PET analysis was performed on Syngo.via software equipped with the EQ.PET software (Siemens Healthineers). For each study, both PSF and PSF_{EARL} SUV_{max} values were extracted from the primary tumour, using a 41% isocontour method as per the EANM guidelines, excluding all nodal foci adjacent to the tumour.

$$SUV_{max} = \frac{maximum\ mesured\ activity\ \left(^{Bq}/_{cc}\right) \times body\ weight\ (g)}{injected\ dose\ (Bq)}$$

with the assumption that the density of the human body is 1 g/cc.

EARL quantitative values were recorded after the application of a Gaussian filter meeting the EARL requirements. The EQ.PET filter was calculated from NEMA NU2 phantom data from a PSF reconstruction, as described in detail elsewhere. Briefly, the recovery coefficients (RCs, defined as the ratio between the measured and true activity concentration for each sphere) of a National Electrical Manufacturers Association NU2 phantom scanned per the EANM guidelines were aligned to the EANM reference RCs by applying a Gaussian filter. As per the last recommendation from the EARL team for our centre, which has been accredited since 2015, the filter was 6.3 mm (Fig. 2).

Statistical analysis

Data are expressed as percentages (%) or means with standard deviations (± SD) or ranges when necessary. A Wilcoxon matched-pairs signed-rank test was used to compare the EARL-compliant and the PSF SUV_{max} values. Overall survival was defined as the time from the baseline PET scan to death of any cause or the end of the study, whichever came first.

We used a validated method to determine cutoffs described hereafter [10]. Potential cutoffs for each SUV measure were selected as described by Mazumdar et al. [11]. All possible values of the SUV measure were selected. Briefly, for each value

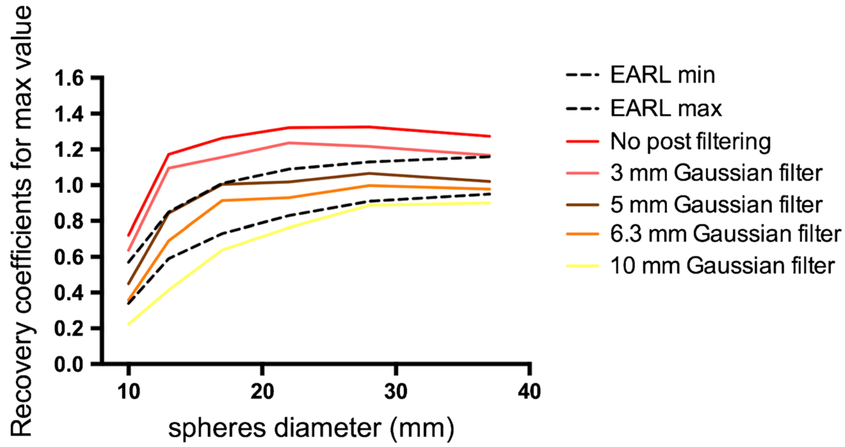

Fig. 2 Recovery coefficients based on maximum pixel values in a NEMA NU2 phantom scanned as per the EANM/EARM accreditation program. Unfiltered PSF reconstruction is used for diagnostic purposes, and PSF with a 6.3 Gaussian filter is used for EARL-compliant quantitation. Other kernels (3, 5, and 10 mm) have been used to simulate a multicentre study by randomly assigning these filters to four groups of patients

ø of the SUV scale the shortest time at which > 90% of the patients had died (T[ø]) was determined. Then, based of the ø / T[ø] curve, the relevant cutoffs were selected while avoiding very similar cutoffs ± 0.01. They corresponded to the constant proportion of the curve ø / T[ø].

The study sample was separated into training and validation datasets following Faraggi and Simon's method [12]. This two-fold cross-validation method consists of dividing the overall sample into learning and validation subsamples. An optimal cutoff is determined for each group using the minimum *p* value approach (the value associated with the maximal log-rank statistic or, equivalently, the minimum *p* value), and each patient was classified according to the cutoff of the subsample to which he/she did not belong. The final cutoff was the value (among all possible cutoffs) that minimized the *p* value in the overall sample, using a stratified log-rank test with subsample as the stratum. We evaluated the stability of the cutoffs using 500 bootstrap replications. Then the log–hazard ratio and its 95% CI were computed using the method described by Höllander et al. [13]. The final PSF_{EARL} and PSF SUV_{max} cutoff values were then used in Cox proportional hazard regression models. Cox proportional hazards models with age as the underlying time scale were used to estimate age-adjusted and multivariable-adjusted hazard ratios (HRs) and 95% confidence intervals (CIs) for the associations between lung cancer mortality and different measures of SUV metrics. Covariates that changed the HR estimate by > 5% were included: age (continuous), gender, and histological subtype. We found that the criteria for the proportional hazards assumption were met using Schoenfeld residuals. Survival curves were constructed using the Kaplan–Meier method.

We additionally conducted two sensitivity analyses: (i) we switched the PSF_{EARL} and PSF cutoff values and estimated their association with LC mortality using similar multivariable Cox regressions, and (ii) we created a virtual random multicentre study in which reconstruction was performed using different methods. To this end, we randomly generated four groups of patients ($n_1 = 34$, $n_2 = 35$, $n_3 = 35$ and $n_4 = 35$) whose images were reconstructed with a PSF reconstruction algorithm with no filtering and a Gaussian post-filtering of 3 mm, 5 mm, and 10 mm respectively (Fig. 2). We also applied the previous optimal EARL and PSF cutoff values to this dataset.

We performed the analyses using Stata 15 (Stata Corp, College Station, TX, USA) and considered two-sided *p* values < 0.05 as statistically significant.

Results

Population characteristics

We identified a database of 497 patients among which 358 were excluded because of the presence of a second cancer, the absence of a baseline PET examination, or because they were presented due to recurrence (Fig. 1). The characteristics of the 139 included patients are summarized in Table 1. With a median follow-up of 28 months (1–104 months), 110 patients (79.1%) died. In the whole population, the estimated OS rate at 36 months was 0.39 (0.31–0.47).

The technical characteristics of the PET examinations were as follows: the injected dose was 4.06 ± 0.48 MBq/kg; the blood

Table 1 Patients characteristics and corresponding univariable proportional HRs and 95% CIs for their associations with lung cancer mortality

Characteristics	Total population ($n = 139$)	Univariable HR (95% CI)
Age (years), *median (range)*	63.0 (57.5–69.0)	1.00 (0.98–1.02)
Gender, *n (%)*		
Females	35 (25.2)	1.00
Males	104 (74.8)	1.32 (0.84–2.06)
Histology, *n (%)*		
Adenocarcinoma	65 (46.8)	1.00
Squamous cell carcinoma	59 (42.4)	1.21 (0.82–1.79)
Others	15 (10.8)	0.80 (0.41–1.59)

glucose level at the time of injection was 5.67 ± 1.38 mmol/l; and the delay between injection and acquisition was 64.3 ± 6.9 min.

Cutoffs definition

Results of optimal cut points are summarized in Supplemental Table 1.

For the PSF, the most frequently selected cutoff across 500 bootstrap replications (436 of 500) was 25.43 (95% CI: 23.41–26.31). The learning and validation subsamples found identical cutoffs 348 times and 341 times out of 500 bootstrap replications for PSF and EARL respectively. We repeated the process for the EARL-compliant PET and found the corresponding numbers (408 of 500) with the most frequently selected cutoffs 8.47 (7.23–9.31).

Quantitative analysis of SUV_{max} with EARL-compliant cutoff values

The mean [± SD, (range)] tumour SUV_{max} values were 18.07 [±8.04, (3.50–45.31)] and 13.54 [±6.55, (2.63–32.55)] for the PSF and EARL-compliant PET datasets, respectively ($p < 0.0001$). In the univariable proportional hazards (Cox) regression, using the optimal PSF SUV_{max} cutoff value on the PSF dataset and the optimal EARL SUV_{max} cutoff on the EARL-compliant dataset, these SUV_{max} cutoff values were both significantly associated with an increase in the hazard of death (Table 2). Kaplan–Meier survival analyses showed that patients with a tumour $SUV_{max} > 25.43$ on PSF images or > 8.47 on EARL-compliant images had significantly poorer OS ($p = 0.03$ and $p < 0.01$, respectively, Fig. 3a). The mean survival times of patients with an $SUV_{max} > 25.43$ on PSF images and > 8.47 on EARL-compliant images were 23.4 months and 15.9 months respectively, compared with the mean survival time of 32.4 months for the overall population. In the univariable analysis, histology, age and gender were not found to be associated with lung cancer mortality (Table 1); however, a multivariable analysis was also conducted and is displayed in Supplemental Table 2.

Table 2 Univariable proportional HRs and 95% CIs of the association between SUV_{max} cutoff values and lung cancer mortality

Univariable proportional hazards (Cox) regression analyses		
Variables	Hazard ratio	95% CI for HR
EARL dataset		
EARL cutoff value > 8.47	1.64	1.04–2.62
PSF cutoff value > 25.43	1.35	0.65–2.75
PSF dataset		
EARL cutoff value > 8.47	1.75	0.89–3.47
PSF cutoff value > 25.43	1.87	1.16–3.03
Simulated multicentre dataset		
EARL cutoff value > 8.47	1.32	0.82–2.10

Quantitative analysis of SUV_{max} when switching cutoff values

In the univariable proportional hazards (Cox) regression, when switching the optimal SUV_{max} cutoff values between the two sets of quantitative data, i.e., applying the optimal PSF SUV_{max} cutoff to the EARL-compliant dataset and applying the optimal EARL SUV_{max} cutoff to the PSF dataset, the SUV_{max} cutoff values were no longer significantly associated with an increase in the hazard of death (Table 2). As expected and shown in Fig. 3b, Kaplan–Meier survival analyses did not demonstrate a significant worse overall survival in patients with a tumour $SUV_{max} > 8.47$ on PSF images or > 25.43 on EARL-compliant images ($p = 0.10$ and 0.43, respectively).

Quantitative analysis of SUV_{max} in a simulated multicentre dataset

As shown in Fig. 2, the application of a Gaussian filtering with various kernels led to recovery coefficients curves within the EARM/EARL bandwidth in one case and outside the EARM/EARL bandwidth in three cases.

When applying the optimal EARL SUV_{max} cutoff to the simulated multicentre dataset, no association was found between a tumour $SUV_{max} > 8.47$ and lung cancer mortality in the univariable proportional hazards (Cox) regression. Moreover, there was no significant difference in the overall survival of patients with a tumour SUV_{max} value ≤ 8.47 or > 8.47 in the Kaplan–Meier analysis ($p = 0.25$, Fig. 3c).

Discussion

Harmonization programmes, such as the EANM/EARL programme and the Quantitative Imaging Biomarker Alliance (QIBA/UPICT) [14, 15], have been set up to allow the comparison of SUV metrics across different centres, especially in multicentre studies [5, 6, 15]. These programmes are also useful in the context of departments that are running several different PET systems. Our group and other contributors have demonstrated that harmonization is necessary for therapeutic monitoring [16, 17]. The present study focused on another use of PET/CT: using SUV metrics as prognosticators for tailored treatment in patients with a higher risk of relapse or death who might benefit from more active surveillance and second line treatments.

Thus far, several retrospective and prospective studies have reported very different cutoff values depending on the PET system used and the reconstruction settings considered. We found cutoff values of 8.47 for the EARL-compliant dataset and 25.43 for unfiltered PSF images (percentage difference =

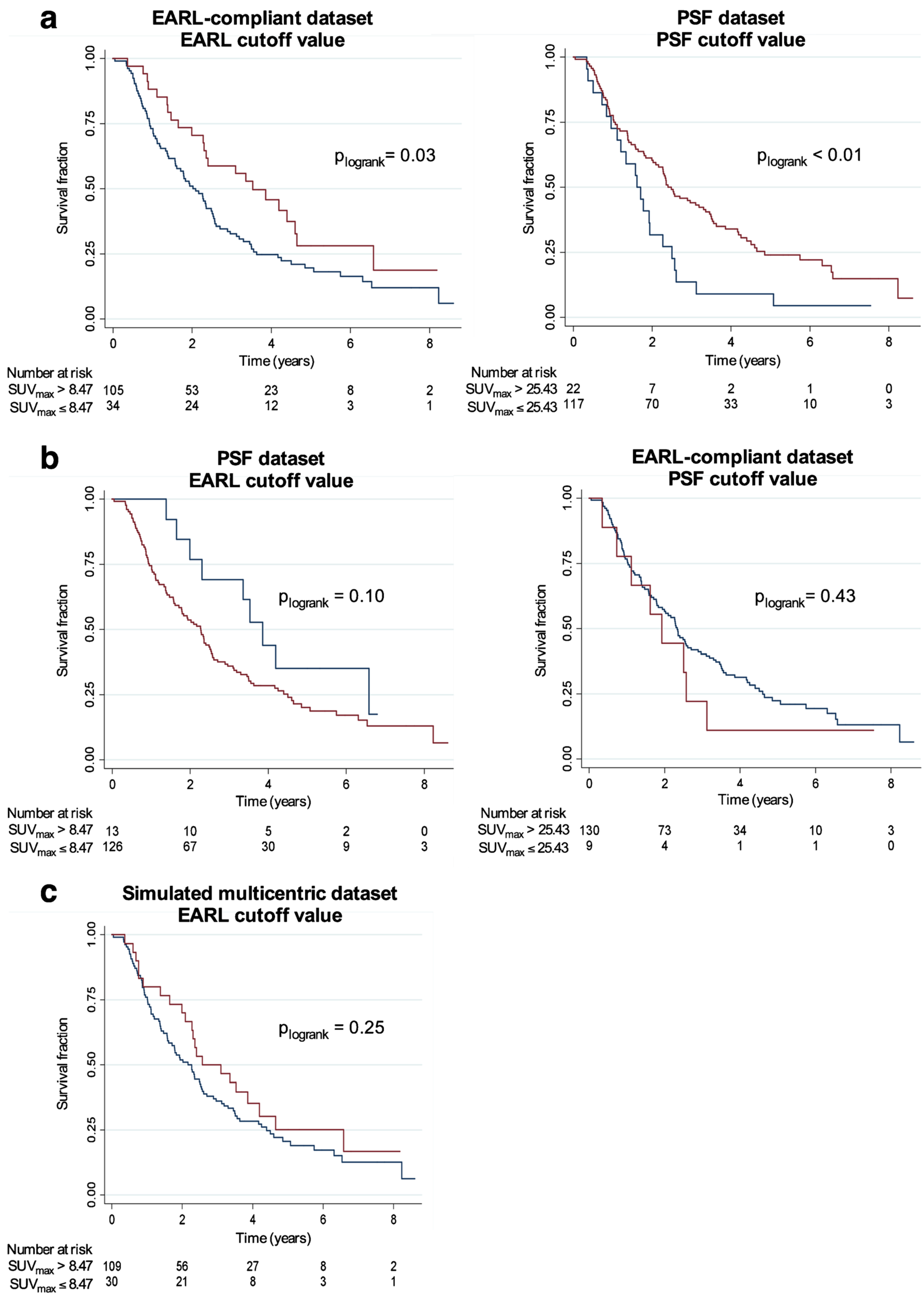

Fig. 3 Prognostic value of tumour SUV_{max}. Overall survival Kaplan–Meier curves according to different scenarios. **a** Use of the optimal PSF and EARL SUV_{max} cutoff values on their respective datasets. **b** Use of the PSF and EARL SUV_{max} cutoff values switched across datasets. **c** Use of the PSF and EARL SUV_{max} cutoff values on a simulated multicentre dataset

66.7). In NSCLC, all stages considered, many other SUV_{max} cutoff values have been described, ranging from 5 to 20 [18]. Our study found that these cutoff values that were determined for a given PET system using a specific reconstruction type could not be extrapolated to other PET centres, in both routine clinical and multicentre studies. It is noteworthy that the different SUV thresholds published for NSCLC patients may also be due to different repartitioning of histological subtypes (adenocarcinoma versus squamous cell carcinoma) and EGFR mutational status in American, European, and Asian cohorts of patients, as these parameters have been shown to influence SUV_{max} [19–21] and are known to depend on ethnicity.

It should be noted that the present study evaluated the worst-case scenario by comparing PSF reconstructions with no post-filtering to their EARL-compliant counterparts. Therefore, one could expect that the observed SUVmax variations would have been less important if post-filtering with a small kernel (2 or 3 mm) had been applied, which is what many centres tend to do [8].

Our study highlights the fact that that standardization of quantitative data is needed for multicentre studies mixing SUVs extracted from different PET systems when risk-stratifying patients. From a clinical perspective, this appears to be necessary for assessing pre-therapeutic tumour metabolic characteristics as a supplemental tool in the patient management process. The ongoing improvements in imaging standardization procedures such as the EANM/EARL programme [5, 6] have paved the way to standardizing SUV cutoff values. Determining SUV threshold values with harmonized quantitative data offers the opportunity to provide prognostic indices that can be extrapolated on a large scale across all accredited centres. One should also keep in mind that in addition to these reconstruction considerations, all the other physical and biological sources of SUV variability [22] must be adequately controlled. Thus, in the present study, cross-calibration between the PET system and the gamma counter used to prepare doses was performed regularly, and the injected activity per kg of body weight as well as the post-injection times were remarkably stable among patients and met the EANM guidelines for PET tumour imaging [9, 23].This could be considered as a strength but also as a weakness of our study, compared to a real multicentre study such as the recent one by Kaalep et al. [24] where cross calibration factors ranged from 0.88 to 1.12. A correction was proposed to account for system calibration bias in order to be able to compare the various reconstruction modes' impact on RCs and not to be affected by inter-scanner calibration errors. For this purpose, a correction factor, defined as the ratio between expected and measured activity concentration in the corresponding uniform background compartment, was applied to all RCs.

A limitation of this study was that the metabolic tumour volumes (MATV) could not be evaluated in the same way that we studied the impact of reconstruction parameters on SUV because the EQ.PET technology did not provide EARL-compliant metabolic volumes. Baseline MATV are being increasingly used; they have been shown to be predictive of OS in NSCLC patients [25, 26], and are sensitive to reconstruction parameters [27]. Even if the determination of the optimal cutoffs involved advanced statistical tools, they need to be validated in other prospective and bigger samples. Nevertheless, as cutoffs for all PET datasets were determined using the same methodology, this does not impact the conclusions of a detrimental effect of using a threshold determined on an older-generation PET system with a newer PET/CT machine, and vice versa; thus, as a consequence, the need for harmonization.

As a perspective, it should be noted that though a large number of the EARL-accredited centres do not have PSF or PSF and time-of-flight capabilities, this situation is expected to change over the next several years. A recent study evaluating recent systems with these reconstruction capabilities or Bayesian penalized-likelihood reconstruction [24] reported a potential upgrade of the EARL specifications. As shown in Supplemental Fig. 1, our unfiltered PSF reconstruction would not fulfil the requirements of this upgrade, but PSF using a post-filtering Gaussian filter with a 3 mm kernel would. The threshold extracted from PSF-filtered data with a 3 mm Gaussian filter and its prognostic capability are shown in Supplemental Fig. 2; this threshold can be used by EARL-accredited centres using advanced reconstruction algorithms when this upgraded EARL specification is released.

Conclusion

The present study including a mature cohort of 139 consecutive patients who presented with stages IIIA and IIIB NSCLC provides the PET community with an EARL-compliant SUV_{max} as an independent prognosticator, and demonstrates the need to harmonize PET protocols with harmonizing standards, such as the EANM/EARL programme, whenever SUV metrics are used as prognosticators.

Compliance with ethical standards

Ethical approval and consent to participate All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Competing interests None to declare.

Informed consent In accordance with European regulations, French observational studies without any additional therapy or monitoring procedure do not need the approval of an ethical committee. Nonetheless, we sought approval to collect data for our study from the national committee for data privacy, the National Commission on Informatics and Liberty (CNIL), with the registration n°2,080,317 v 0.

References

1. Madsen PH, Holdgaard PC, Christensen JB, Hoilund-Carlsen PF. Clinical utility of F-18 FDG PET-CT in the initial evaluation of lung cancer. Eur J Nucl Med Mol Imaging. 2016;43:2084–97. https://doi.org/10.1007/s00259-016-3407-4.
2. Mac Manus MP, Everitt S, Bayne M, Ball D, Plumridge N, Binns D, et al. The use of fused PET/CT images for patient selection and radical radiotherapy target volume definition in patients with non-small cell lung cancer: results of a prospective study with mature survival data. Radiother Oncol. 2013;106:292–8. https://doi.org/10.1016/j.radonc.2012.12.018.
3. Hicks RJ. Role of 18F-FDG PET in assessment of response in non-small cell lung cancer. J Nucl Med. 2009;50(Suppl 1):31S–42S. https://doi.org/10.2967/jnumed.108.057216.
4. Eschmann SM, Friedel G, Paulsen F, Reimold M, Hehr T, Budach W, et al. Is standardised (18)F-FDG uptake value an outcome predictor in patients with stage III non-small cell lung cancer? Eur J Nucl Med Mol Imaging. 2006;33:263–9. https://doi.org/10.1007/s00259-005-1953-2.
5. Aide N, Lasnon C, Veit-Haibach P, Sera T, Sattler B, Boellaard R. EANM/EARL harmonization strategies in PET quantification: from daily practice to multicentre oncological studies. Eur J Nucl Med Mol Imaging. 2017;44:17–31. https://doi.org/10.1007/s00259-017-3740-2.
6. Kaalep A, Sera T, Oyen W, Krause BJ, Chiti A, Liu Y, et al. EANM/EARL FDG-PET/CT accreditation — summary results from the first 200 accredited imaging systems. Eur J Nucl Med Mol Imaging. 2018;45:412–22. https://doi.org/10.1007/s00259-017-3853-7.
7. Lasnon C, Salomon T, Desmonts C, Do P, Oulkhouir Y, Madelaine J, et al. Generating harmonized SUV within the EANM EARL accreditation program: software approach versus EARL-compliant reconstruction. Ann Nucl Med. 2017;31:125–34. https://doi.org/10.1007/s12149-016-1135-2.
8. Quak E, Le Roux PY, Hofman MS, Robin P, Bourhis D, Callahan J, et al. Harmonizing FDG PET quantification while maintaining optimal lesion detection: prospective multicentre validation in 517 oncology patients. Eur J Nucl Med Mol Imaging. 2015;42:2072–82. https://doi.org/10.1007/s00259-015-3128-0.
9. Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging. 2015;42:328–54. https://doi.org/10.1007/s00259-014-2961-x.
10. Diouf M, Bonnetain F, Barbare JC, Bouche O, Dahan L, Paoletti X, et al. Optimal cut points for Quality of Life Questionnaire–core 30 (QLQ–C30) scales: utility for clinical trials and updates of prognostic systems in advanced hepatocellular carcinoma. Oncologist. 2015;20:62–71. https://doi.org/10.1634/theoncologist.2014-0175.
11. Mazumdar M, Glassman JR. Categorizing a prognostic variable: review of methods, code for easy implementation and applications to decision-making about cancer treatments. Stat Med. 2000;19:113–32.
12. Faraggi D, Simon R. A simulation study of cross-validation for selecting an optimal cutpoint in univariate survival analysis. Stat Med. 1996;15:2203–13. https://doi.org/10.1002/(sici)1097-0258(19961030)15:20<2203::aid-sim357>3.0.co;2-g.
13. Hollander N, Sauerbrei W, Schumacher M. Confidence intervals for the effect of a prognostic factor after selection of an 'optimal' cutpoint. Stat Med. 2004;23:1701–13. https://doi.org/10.1002/sim.1611.
14. Graham MM, Wahl RL, Hoffman JM, Yap JT, Sunderland JJ, Boellaard R, et al. Summary of the UPICT Protocol for 18F-FDG PET/CT Imaging in Oncology Clinical Trials. J Nucl Med. 2015;56:955–61. https://doi.org/10.2967/jnumed.115.158402.
15. Makris NE, Huisman MC, Kinahan PE, Lammertsma AA, Boellaard R. Evaluation of strategies towards harmonization of FDG PET/CT studies in multicentre trials: comparison of scanner validation phantoms and data analysis procedures. Eur J Nucl Med Mol Imaging. 2013;40:1507–15. https://doi.org/10.1007/s00259-013-2465-0.
16. Lasnon C, Quak E, Le Roux PY, Robin P, Hofman MS, Bourhis D, et al. EORTC PET response criteria are more influenced by reconstruction inconsistencies than PERCIST but both benefit from the EARL harmonization program. EJNMMI Phys. 2017;4:17. https://doi.org/10.1186/s40658-017-0185-4.
17. Quak E, Le Roux PY, Lasnon C, Robin P, Hofman MS, Bourhis D, et al. Does PET SUV harmonization affect PERCIST response classification? J Nucl Med. 2016;57:1699–706. https://doi.org/10.2967/jnumed.115.171983.
18. Berghmans T, Dusart M, Paesmans M, Hossein-Foucher C, Buvat I, Castaigne C, et al. Primary tumor standardized uptake value (SUVmax) measured on fluorodeoxyglucose positron emission tomography (FDG-PET) is of prognostic value for survival in non-small cell lung cancer (NSCLC): a systematic review and meta-analysis (MA) by the European Lung Cancer Working Party for the IASLC Lung Cancer Staging Project. J Thorac Oncol. 2008;3:6–12. https://doi.org/10.1097/JTO.0b013e31815e6d6b.
19. Lv Z, Fan J, Xu J, Wu F, Huang Q, Guo M, et al. Value of (18)F-FDG PET/CT for predicting EGFR mutations and positive ALK expression in patients with non-small cell lung cancer: a retrospective analysis of 849 Chinese patients. Eur J Nucl Med Mol Imaging. 2017;45(5):735–750 https://doi.org/10.1007/s00259-017-3885-z.
20. Wong CY, Nunez R, Bohdiewicz P, Welsh RJ, Chmielewski GW, Ravikrishnan KP, et al. Patterns of abnormal FDG uptake by various histological types of non-small cell lung cancer at initial staging by PET. Eur J Nucl Med. 2001;28:1702–5. https://doi.org/10.1007/s002590100638.
21. Yoshida T, Tanaka H, Kuroda H, Shimizu J, Horio Y, Sakao Y, et al. Standardized uptake value on (18)F-FDG-PET/CT is a predictor of EGFR T790M mutation status in patients with acquired resistance to EGFR-TKIs. Lung Cancer. 2016;100:14–19. https://doi.org/10.1016/j.lungcan.2016.07.022.
22. Boellaard R. Standards for PET image acquisition and quantitative data analysis. J Nucl Med. 2009;50(Suppl 1):11S–20S. https://doi.org/10.2967/jnumed.108.057182.
23. 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. https://doi.org/10.1007/s00259-009-1297-4.
24. Kaalep A, Sera T, Rijnsdorp S, Yaqub M, Talsma A, Lodge MA, et al. Feasibility of state of the art PET/CT systems performance harmonisation. Eur J Nucl Med Mol Imaging. 2018;45(8):1344–1361 https://doi.org/10.1007/s00259-018-3977-4.
25. Salavati A, Duan F, Snyder BS, Wei B, Houshmand S, Khiewvan B, et al. Optimal FDG PET/CT volumetric parameters for risk stratification in patients with locally advanced non-small cell lung cancer: results from the ACRIN 6668/RTOG 0235 trial. Eur J Nucl Med Mol Imaging. 2017;44:1969–83. https://doi.org/10.1007/s00259-017-3753-x.
26. Soussan M, Chouahnia K, Maisonobe JA, Boubaya M, Eder V, Morere JF, et al. Prognostic implications of volume-based measurements on FDG PET/CT in stage III non-small-cell lung cancer after induction chemotherapy. Eur J Nucl Med Mol Imaging. 2013;40:668–76. https://doi.org/10.1007/s00259-012-2321-7.
27. Lasnon C, Enilorac B, Popotte H, Aide N. Impact of the EARL harmonization program on automatic delineation of metabolic active tumour volumes (MATVs). EJNMMI Res. 2017;7:30. https://doi.org/10.1186/s13550-017-0279-y.