ORIGINAL ARTICLE

Assessing immune organs on ^{18}F-FDG PET/CT imaging for therapy monitoring of immune checkpoint inhibitors: inter-observer variability, prognostic value and evolution during the treatment course of melanoma patients

Kevin Prigent[1] · Charline Lasnon[2] · Emilien Ezine[3] · Mélanie Janson[4] · Nicolas Coudrais[1] · Elisa Joly[3] · Laure Césaire[3] · Andrea Stefan[3] · Michel Depontville[3] · Nicolas Aide[1,4]

Abstract

Background Immune checkpoint inhibitors (ICIs) have significantly improved survival in advanced melanoma. There is a need for robust biomarkers to identify patients who do not respond. We analysed 14 baseline ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET/CT) metrics and their evolution to assess their correlation with patient outcome, compared with 7 established biological markers and 7 clinical variables.

Methods We conducted a retrospective monocentric observational study of 29 patients with advanced melanoma who underwent baseline ^{18}F-FDG PET/CT, followed by an early monitoring PET/CT (iPET) scan after 1 month of treatment and follow-up studies at 3rd (M3PET) and 6th month (M6PET). ^{18}F-FDG uptake in immune organs (spleen, bone marrow, ileocecal valve) and derived spleen-to-liver (SLR) and bone-to-liver (BLR) ratios were reviewed by two PET readers for reproducibility analysis purposes including 14 PET variables. The most reproducible indexes were used for evaluation as predictors of overall survival (OS) in comparison with PET response using imPERCIST5, whole-body metabolic active tumour volume (WB-MATV) and biological parameters (lactate dehydrogenases (LDH), reactive protein c (CRP), white blood count (WBC), absolute lymphocyte count (ALC), neutrophil to lymphocyte ratio (NLR) and derived neutrophils to lymphocyte ratio).

Results Strong reproducibility's (intraclass coefficients of correlation (ICC) > 0.90) were observed for spleen anterior SUV_{peak}, spleen MV, spleen TLG, spleen length and BLR_{mean}. ICC for SLR_{mean} and ileocecal SUV_{mean} were 0.86 and 0.65, respectively. In the 1-year OS 1 group, SLR_{mean} tended to increase at each time point to reach a significant difference at M6-PET ($p = 0.019$). The same trends were observed with spleen SUV_{peak} anterior and spleen length. In the 1-year OS 0 group, a significative increase of spleen length was found at iPET, as compared with baseline PET ($p = 0.014$) and M3-PET ($p = 0.0239$). Univariable Kaplan-Meier survival analysis found that i%var spleen length, M3%var SLR_{mean}, baseline LDH, i%var NLR and response at M6PET were all predictors of 1-year OS.

Conclusions SLR_{mean} is recommended as a prognosticator in melanoma patients under immunotherapy: its increase greater than 25% at 3 months, compared with baseline, was associated with poor outcome after ICIs.

Keywords Immune checkpoint inhibitors · Melanoma · 18F-FDG · PET/CT · Spleen · Ileocecal valve · imPERCIST · irAE

Kevin Prigent, Charline Lasnon and Emilien Ezine contributed equally to this work.

This article is part of the Topical Collection on Oncology - General.

✉ Nicolas Aide
    aide-n@chu-caen.fr

1   Nuclear Medicine Department, University Hospital, Avenue Côte de Nacre, 14000 Caen, France

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

3   Dermatology Department, University Hospital, Caen, France

4   Normandy University, Caen, France

Introduction

Immunotherapy has recently emerged as an important advance in cancer treatment. To date, the main immunotherapeutic approach that has been translated into survival benefit and is currently used in practice is the immune checkpoint inhibitors (ICIs). Broadly, the two most effective classes of agent are directed, alone or in combination, towards cytotoxic T lymphocyte-associated protein 4 (CTLA-4) or the programmed cell death protein 1 (PD1) or the PD1/programmed cell death protein ligand 1 (PD1/PD-L1) axis, which are negative regulators of T cell immune function. The CTLA-4 inhibitor, ipilimumab, has been shown to improve survival rates in melanoma patients [1]. PD1/PD-L1 inhibitors (of which the first validated agents were pembrolizumab and nivolumab) have been shown to improve survival rates among patients with various tumour types [2, 3].

Response assessment with PET/CT during immunotherapy can be rather challenging, since ^{18}F-FDG avid inflammatory reactions can occur during the treatment and may be associated with pseudoprogression, leading to misinterpretation of ^{18}F-FDG PET images [4–6]. The PET community has been very active in the last decade, coming up with various response criteria including PERCIMT (PET Response Evaluation Criteria for Immunotherapy). PECRIT (PET/CT Criteria for Early Prediction of Response to Immune Checkpoint Inhibitor Therapy), iPERCIST or PERCIST (PET Response Criteria in Solid Tumours)-adapted criteria such as imPERCIST5 [7].

On top of these PET response criteria created or modified to meet the challenges raised by immunotherapy,, several groups have reported baseline prognostic factors of response such as the metabolic active tumour volume (MATV) [8] and signs of immune activation [9]. However, while inter-observer variability in MATV delineation is known [10] and percentage change in SUV used to discriminate between responders and non-responders when using PERCIST or PERCIST-adapted criteria are based on test-retest studies [11], little data is available on immune activation signs [12]. The first sign of immune activity to be checked is spleen enlargement and/or increased uptake leading to an inversion of the spleen to liver ratio (SLR) [13, 14]. Some groups also proposed other signs such as the bone marrow to liver ratio (BLR) [15] and uptake in the ileocecal valve [16].

The biological counterparts of PET signs of immune activation are absolute lymphocyte count (ALC) and neutrophils to lymphocytes ratio (NLR), the latter having been shown to be a prognosticator in melanoma patients [17]. Also correlated to a worst survival is serum lactate dehydrogenase (LDH), a biochemical marker of tumour load [18].

The main objective of the present study was to seek the most clinically useful signs of immune activation on ^{18}F-FDG PET images to predict the prognostic of melanoma patients' ongoing ICI therapy. Ideally, these signs should be highly reproducible

between observers, simple and fast to compute and should be independent factors of response, especially compared with widely available, cost-effective and established biological parameters.

Materials and methods

Study design

We conducted a retrospective monocentric observational study in melanoma patients aged over 18 and who were referred for a ^{18}F-FDG PET/CT for therapy monitoring of immunotherapy. After searching in our database, out of 63 patients screened, 36 patients met the criteria and were included. Inclusion criteria were: (i) immunotherapy treatment; (ii) availability of baseline ^{18}F-FDG PET/CT, followed by an early monitoring PET/CT scan (iPET) after 1 month of treatment and follow-up studies at 3rd (M3PET) and 6th month (M6PET). Seven patients were excluded based on the following criteria: (i) baseline whole-body metabolic active tumour volume (WB-MATV) or TLG (WB-TLG) at 0 and (ii) missing data. PET/CT acquisitions dated from 11/08/2017 to 04/12/2019. Institutional review board approval was not required in accordance with European regulation. French observational studies without any additional therapy or monitoring procedure do not need the approval of an ethics committee. Nonetheless, in accordance with the European General Data Protection Regulation, we sought approval to collect data for this work from the national committee for data privacy, with the registration no. 2081250 v 0.

^{18}F-FDG PET/CT protocol

Patient's preparation in the PET unit and PET acquisition and reconstructions were performed as per the European Association of Nuclear Medicine (EANM) guidelines for PET tumour imaging [19, 20], our PET unit being EANM research Ltd (EARL) accredited since 2015. ^{18}F-FDG was injected after glucose level had been checked to be < 200 mg/dL in patients who had been fasting for at least 4 h. Patients were tentatively scanned 60 min after tracer injection. Patients were scanned from the base of the skull to mid-thigh, except for patients with primary melanoma of the lower limb, for whom a whole-body acquisition was performed. Two different PET/CT scanners were used: a Vereos system (Philips Medical Systems. Cleveland OH) and a Biograph TrueV with extended field of view (Siemens Medical Solutions). On the Biograph system, the injected dose was 4.0 MBq/kg with time per bed position of 2 min and 40 s for body mass index (BMI) < 25 kg/m^2 and 3 min and 40 s and for patients with BMI $\geq$ 25 kg/m^2. To meet the EARL harmonizing standards, PET/CT raw data were reconstructed with a 3D-OSEM (ordered

subsets expectation-maximization) reconstruction algorithm with point spread function (PSF) modeling (3 iterations and 21 subsets) and a 6.3-mm post-reconstruction Gaussian filter, using a 128 × 128 matrix size. Concerning the Vereos system, the injected dose was 3.0 MBq/kg with a time per bed position of 2 min whatever the patient's weight. To fulfill the EARL accreditation, PET/CT raw data were reconstructed with a 3D-OSEM reconstruction algorithm with PSF modeling (2 iterations and 10 subsets) with a 5-mm post-reconstruction Gaussian filter, using 288 × 288 matrix size.

FDG PET/CT analysis

PET readers were blinded to the results of any other imaging tests, clinical outcomes, or biological data. The analysis was carried out using LIFEx, an open source software [21]. All quantitative analyses were made using EARL-compliant PET images [20].

Fourteen PET variables involving the spleen, bone marrow and the ileocecal valve were reviewed by two PET readers as potential immune activation signs, at baseline, iPET and M3PET, for reproducibility analysis purposes. Other biomarkers (PET response and WB-MATV/TLG) were assessed by a single nuclear medicine physician. Absolute values were considered for the baseline PET and relative values expressed as percentage of variation (%var) for iPET and M3PET.

Evaluation of immune activation and immune-related side effects

First, spleen immune activation signs were analysed as follows:

- The maximum standardized uptake values (SUV_{max}), the SUV_{mean}, from a volume of interest (VOI) of 20 mm of diameter placed on the central part of the spleen.
- Three SUV_{peak} values were recorded in anterior, median and posterior part of the spleen.
- The spleen length was recorded as recommended as per the American Roentgen Ray Society on the CT [22] part of the PET/CT scan, using the greatest splenic measurement in the cranio-caudal axis from a coronal view.
- SLR was calculated from a central VOI of 20 mm of diameter on spleen and a 30 mm VOI on the right liver lobe using SUV_{max} and SUV_{mean} values, referred to as SLR_{max} and SLR_{mean}, respectively [15].
- The metabolic volume (spleen MV) and total lesion glycolysis (spleen TLG) from the whole spleen PET volume from interpolated regions of interest (ROI), using the 2D pencil tool around the spleen from LIFEx Software [21].

Secondly, other potential lymphocyte activation signs were recorded from the whole ileocecal valve using an adaptive VOI [16] and the bone marrow using SUV_{max} and SUV_{mean}. Osteomedullary activation was measured from a 20 mm VOI located on the right iliac bone. Then, BLR were calculated and referred to as BLR_{max} and BLR_{mean} still using a 30 mm VOI on right liver lobe.

Finally, all potential signs of immune-related adverse effects (irAEs) were recorded based on existing literature [13, 23–25]. Corticosteroid use for more than 2 months was also recorded.

Evaluation of MATV and TLG

The MATV protocol from LIFEx was used [26]. Each lesion was completely by-passed by an adapted VOI by hand; then, a threshold of 41% of SUV_{max} was applied as recommended by the EANM for solid tumours [19]. Thus, each MATV and TLG (TLG = SUV_{mean} × MATV) per lesion was summed to obtain WB-MATV and WB-TLG. Any sign of potential immune infiltrate organs was excluded. Figure 1 displays the methodology for MATV computation and immune organ evaluation.

Evaluation of PET response

Tumour response was assessed at 3 and 6 months. Recent immunotherapy-modified PERCIST using five lesions, imPERCIST5, was used [7].

Briefly, the recommendations of PERCIST5 were followed: up to five hottest target lesions were recorded to be summed (maximum of two per organ). The appearance of new lesions alone did not result in progressive metabolic disease (PMD); only an increase of the sum of SUL_{peak} by 30% was considered as PMD, with or without new lesions. Response to immunotherapy was classified as complete metabolic response (CMR), partial metabolic response (PMR), stable metabolic disease (SMD) or PMD. The responders concerned the patients classified as CMR or PMR and non-responders as SMD and PMD, as follows.

- CMR: complete resolution of ^{18}F-FDG uptake in the tumour volume, with tumour uptake lower than liver SUL and background blood pool, and disappearance of all lesions if multiple.
- PMR: at least 30% reduction in tumour uptake.
- SMD: less than 30% increase, or less than 30 decrease in tumour uptake
- PMD: greater than 30% increase in tumour uptake.

Biological biomarkers

The values of LDH (normal value 100–190 UI/L)), reactive protein C (CRP, normal value < 10 mg/L), white blood count
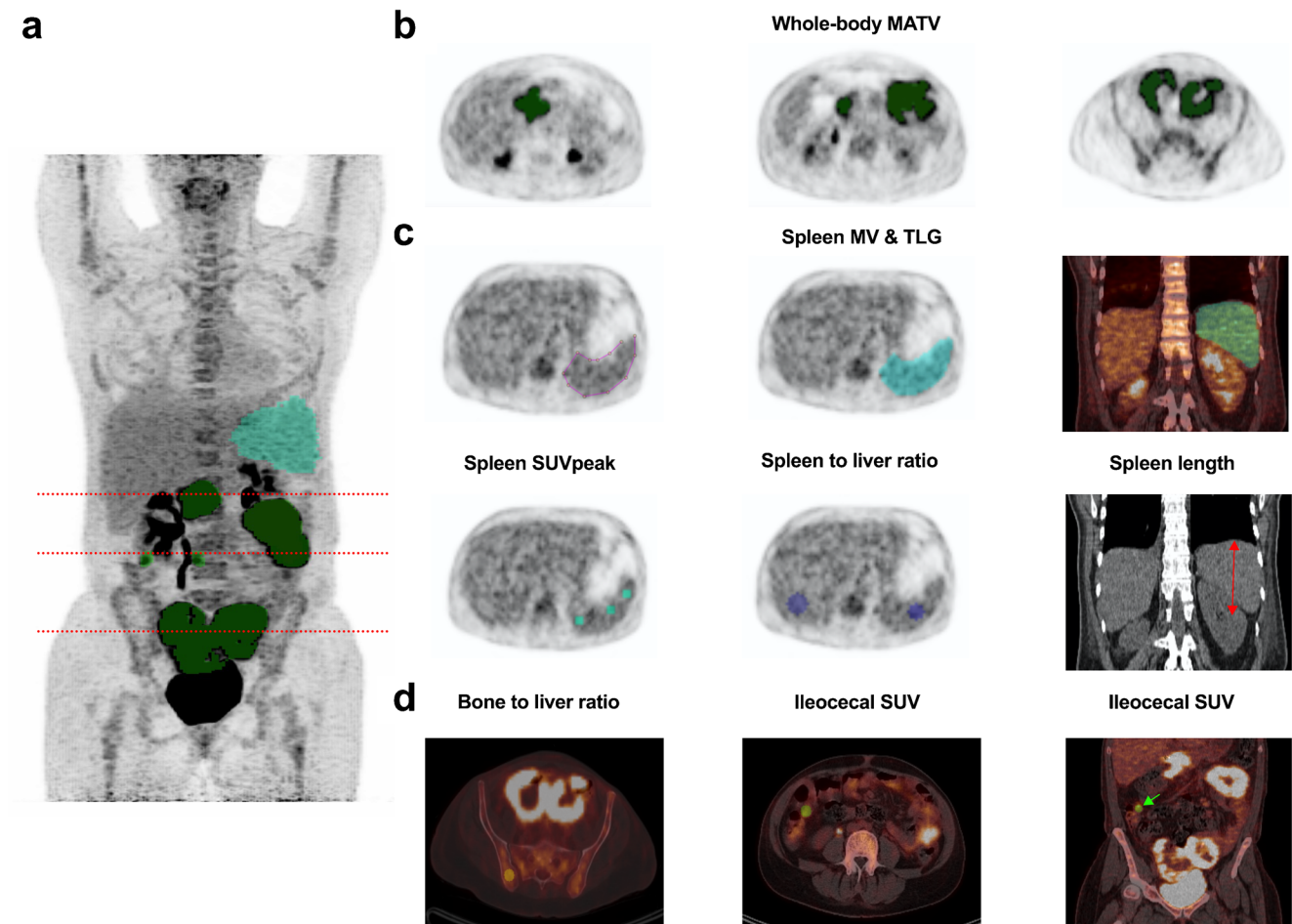
Fig. 1 Methodology for measurement of MATV and uptake in immune organs. ^{18}F-FDG-PET/CT monitoring in a 50-year-old woman treated for a metastatic melanoma under Nivolumab **a** MIP imaging and **b** PET transverse planes showing the spleen metabolic volume (blue) and the whole body MATV (green). **c** First line from the left to the right. The spleen MV and TLG recording process is shown in the PET transverse plane, beginning by drawing the spleen contours at different levels followed by interpolation to obtain the spleen MV. Second line from the left to the right showing the 3 three spleen SUV$_{peak}$ VOIs (blue) then the 2 VOIs (purple) used to compute the SLR$_{max}$ and SLR$_{mean}$. Finally, the CT in the coronal plane to record the spleen length (red arrow) **d**

(WBC, normal value 4000–10,000/mm^3), absolute neutrophils count (ANC, normal value 1500–7000/mm^3) and ALC (normal value 1500–4000/mm^3) were recorded. Each value comes from the biological assessment done systematically in approved laboratories 48 to 72 h before immunotherapy. Methods may differ depending on the laboratory.

The NLR and the derived neutrophils to lymphocytes ratio (dNLR) were calculated as follows:

$$NLR = {^{ANC}}\!/\!_{ALC}$$

$$dNLR = {^{ANC}}\!/\!_{(WBC-ALC)}$$

Statistical analysis

Quantitative variables are presented as mean (SD).

To assess the inter-observer reproducibility and compare all PET quantitative measurements made by the two observers, intraclass correlation coefficients, Bland-Altman analyses as well as non-parametric Wilcoxon tests were used.

For the identification of PET and biological variables of interest to determine the 1-year OS status of patients, a univariable analysis by means of non-parametric Mann-Whitney test was used to compare quantitative variables of patients alive at 1 year (1 year OS 0 patients) and patients dead at 1 year (1 year OS 1 patients). For qualitative variable, Fischer exact tests were used. The landmarked specific overall survival (OS) was defined as the time from the first immunotherapy cycle to death from melanoma. The correlations between PET and biological variables of interest were studied using Spearman correlations. Afterward, ROC analyses were used to determine their optimal cut-off to discriminate between 1 year OS 0 and 1 year OS 1 patients based on the

optimisation of the Youden index and 1 year OS analyses were performed using univariable Kaplan-Meier survival analyses with log-rank test to compare survival curves.

The evolution biological and PET variables of interest during immunotherapy from baseline to M3PET were assessed using Friedman tests. A post hoc test was performed with the two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli test for multiple comparisons.

To confront the importance of biological PET variables in the prediction of the 1 year OS status, a multivariable random forest prediction model was used incorporating all PET and biological variables previously identified. The importance of each PET variable in the classification process was determined by the measurement of its mean decrease accuracy of class prediction. The method implemented classification and regression trees (CART, $n = 100$) and bootstrapping aggregating (bagging) method proposed by Breiman [27]. Moreover, a CHAIN algorithm classification tree was performed with a maximum tree depth set to 3 to avoid a too complex tree.

For all statistical tests, a two-tailed P value of less than 0.05 was considered statistically significant. Graphs and statistical analysis were performed on XLSTAT Software (XLSTAT 2017: Data Analysis and Solution for Microsoft Excel, Addinsoft, Paris, France (2017)).

Results

Population characteristics

Twenty-nine patients were included. Baseline patients' characteristics are displayed in Table 1. Fifteen patients (51.7%) had more than 2 metastatic sites. BRAF^{V600} mutation was found in 11 patients (38.0%), RAS mutation in 3 patients (10.3%) and NRAS mutation in 4 patients (13.8%). No KRAS mutation was observed. The median follow-up was 16.6 months and the 1 year OS was 65.5 % in the entire population. There were 10 recorded events in the entire population over the first year of follow-up. Only one patient was lost to follow-up at 6.4 months (Table 1).

Comparison and assessment of reproducibility of PET measured by the two observers

ICC values for all explored variables are displayed in Table 2. Strong reproducibilities (ICC > 0.90) were observed for spleen anterior SUV$_{peak}$, spleen MV, spleen TLG, spleen length and BLR$_{mean}$. Focusing on SLR, best performance was observed using mean values with an ICC equal to 0.86 as compared with 0.66 using maximum values. In the same trend, although ileocecal valve SUV ICC performances appeared lower, judged as moderate, better results were observed using SUV$_{mean}$ values. Bland-Altman analyses of these

seven PET variables, performed on all time points mixed together, are displayed on Fig. 2. Largest bias and 95%CI were observed for ileocecal valve SUV$_{mean}$ (7.45, 95%CI = 64.81–49.9) and BLR$_{mean}$ (2.59, 95%CI = 25.47–20.29). Moreover, Wilcoxon tests found that anterior spleen SUV$_{peak}$ values and ileocecal valve SUV$_{mean}$ values were significantly different between observers: Obs1 = 1.9 versus Obs2 = 1.89 (p = 0.02) and Obs1 = 1.53 versus Obs2 = 1.35 ($p < 0.001$), respectively. Descriptive analysis for others variables can be found in Table 3.

Overall, the more robust signs of immune activation were therefore SLR$_{mean}$, spleen MV and TLG, spleen length and BLR$_{mean}$.

Identification of PET variables of interest to predict the 1-year OS status

Evaluation of immune activation

Concerning quantitative PET variables, interim %var splenic length, M3%var SLR$_{mean}$ and M3%var ileocecal valve SUV$_{mean}$ were found to be significantly different between 1 year OS 0 and 1 year OS 1 patients. Interim %var splenic length was significantly lower in 1 year OS 1 patients [− 3.1 (14.3) versus 15.9 (31.4); p = 0.01] whereas M3%var SLR$_{mean}$ and M3%var ileocecal valve SUV$_{mean}$ were higher in 1 year OS 1 patients: − 19.1 (24.4) versus 3.4 (14.3); p = 0.03 and 2.38 (26.85) versus 106.91 (159.24); p = 0.03, respectively.

Evaluation of MATV and TLG

Characteristics of all quantitative PET variables as well as Mann-Whitney results between 1 year OS 0 and 1 year OS 1 patients can be found in Table 4. Of note, baseline WB-MATV almost reached significance with a p value equal to 0.06 with and values seemingly higher in 1 year OS 1 patients: 37.23 (55.08) versus 79.9 (78.5).

Correlation with biological biomarkers

The exploration of correlations between PET variables of interest and all biological variables (Fig. 3) found that SLR$_{mean}$ was positively correlated to M3%var LDH (ρ = 0.533, p = 0.005) and negatively correlated to M3%var ALC (ρ = − 0.415, p = 0.035). Of note, there were also significant positive correlations of M3%var SLR$_{mean}$ with M3%var WB-MATV (ρ = 0.382, p = 0.049) and with M3%var BLR (ρ = 0.610, p = 0.001).

Evaluation of PET response

Concerning qualitative variables, Fischer tests showed that 1 year OS 1 patients were mostly patients classified as non-responders on M6PET (p = 0.027) but no significant

Table 1 Baseline patients' characteristics

All patients	All patients (n = 29)	1 year OS0 (n = 19)	1 year OS1 (n = 10)	P value
Gender, n (%)				
Male	15 (51.7)	9 (47.4)	6 (60)	P = 0.699
Female	14 (48.3)	10 (52.6)	4 (40)	
Age (years), mean [range]	67.4 [37–91]	70.4 [51–91]	61.70 [37–85]	P = 0.134
WHO performance, n (%)				
< 2	26 (89.6)	17 (89.4)	9 (90)	P > 0.99
≥ 2	2 (7)	1 (5.2)	1 (10)	
Missing	1 (3.4)	1 (5.2)	0	
Immunotherapy type, n (%)				
Nivolumab	19 (65.5)	13 (68.4)	6 (60)	P = 0.69
Pembrolizumab	9 (31)	6 (31.6)	3 (30)	
Ipilimumab and nivolumab	1 (3.5)	0	1 (10)	
Histology primary melanoma, n (%)				
Cutaneous	15 (51.7)	8 (42.1)	7 (70)	P = 0.49
Non cutaneous	5 (17.3)	4 (21)	1 (10)	
Primary melanoma unknown	9 (31)	7 (36.9)	2 (20)	
Organ involvement, n (%)				
(sub)cutaneous	11 (38)	10 (52.6)	1 (10)	P = 0.043
Lymph nodes	15 (51.7)	10 (52.6)	5 (50)	P > 0.99
Lungs	8 (27.6)	4 (21)	4 (40)	P = 0.39
Muscular	4 (13.8)	1 (5.2)	3 (30)	P = 0.1
Skeletal	2 (6.9)	2 (10.5)	0	P = 0.52
Liver	8 (27.6)	6 (31.6)	2 (20)	P = 0.67
Abdominal	9 (31)	5 (26.3)	4 (40)	P = 0.67
Adrenal	6 (20.7)	4 (21)	2 (20)	P > 0.99
Heart/pericardium	3 (10.3)	1 (5.2)	2 (20)	P = 0.26
Brain	8 (27.6)	4 (21)	4 (40)	P = 0.39
Cutaneous primary melanoma	All patients (n = 15)	1 year OS0 (n = 8)	1 year OS1 (n = 7)	P value
Histoprognostic factors, n (%)				
Breslow (mm), mean (SD)	4.65 (7.88)	6.27 (9.91)	2.06 (0.92)	P = 0.91
Clark				
<IV	3 (20)	2 (25)	1 (14.3)	P = 0.99
≥IV	8 (53.3)	4 (50)	4 (57.1)	
Missing	4 (26.7)	2 (25)	2 (28.6)	
Ulceration				
Yes	8 (53.3)	5 (62.5)	3 (43)	P = 0.8
No	4 (26.7)	2 (25)	2 (28.5)	
Missing	3 (20)	1 (12.5)	2 (28.5)	
Regression				
Yes	1 (6.7)	1 (12.5)	0	P = 0.99
No	11 (73.3)	6 (75)	5 (71.4)	
Missing	3 (20)	1 (12.5)	2 (28.6)	
Mitotic index				P = 0.99
High	3 (20)	2 (25)	1 (14.3)	
Low	9 (60)	5 (62.5)	4 (57.1)	
Missing	3 (20)	1 (12.5)	2 (28.6)	

relationship with imPERCIST M3PET response classification was found (p = 0.098).

Occurrence of irAEs

Overall, 33 patients displayed irAEs (44.8%) mostly on iPET (61.5%). Endocrine and digestive systems were the most affected. Overall, 5 colitis, 4 thyroiditis, 3 gastritis, 3 enthesopathies, 2 hypophysitis and 2 myositis were observed, all time point considered. Five patients had multiple irAEs (≥ 2). There was no treatment interruption and only four patients received immunosuppressants (30.8%) with long-term corticosteroid therapy (n = 3) or cyclosporine (n = 1). No relation between patient's outcome and i-irAEs, M3-irAEs or long-term use of corticosteroid was found with p value equal to 0.201, 0.398 and 0.128, respectively.
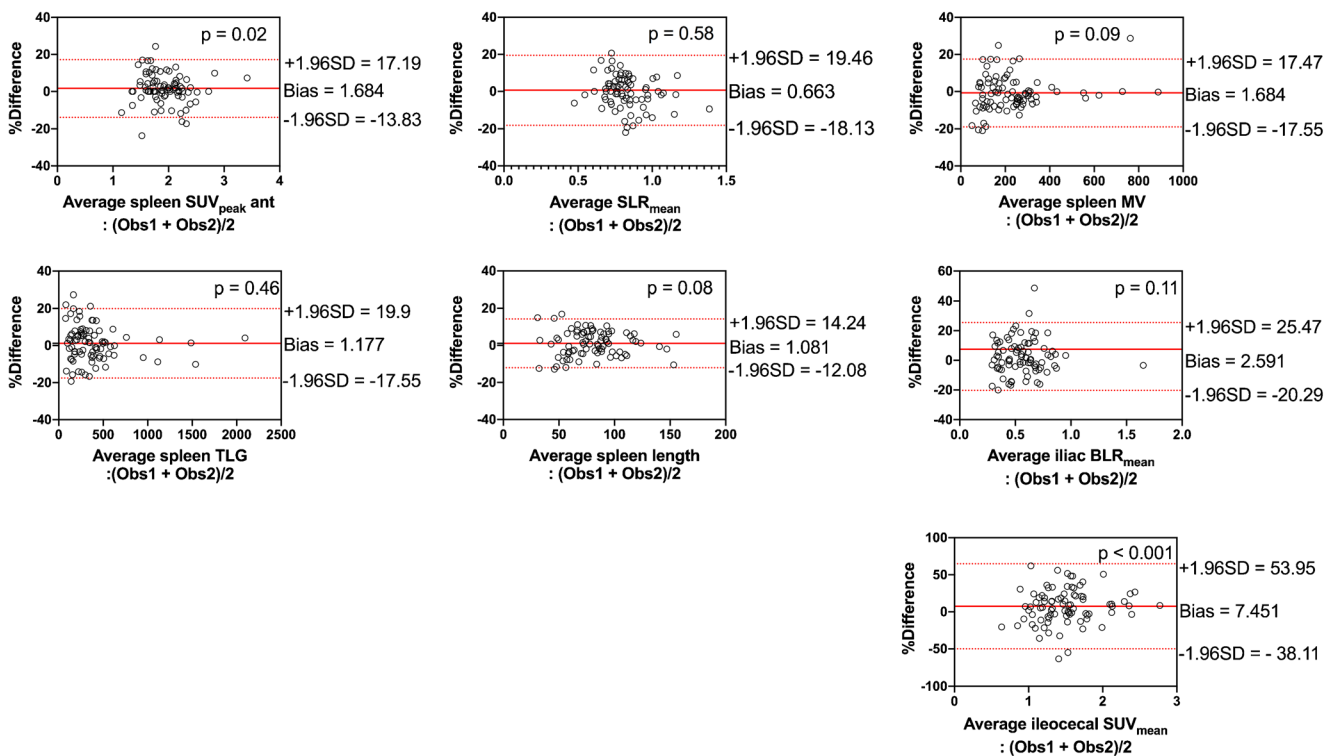
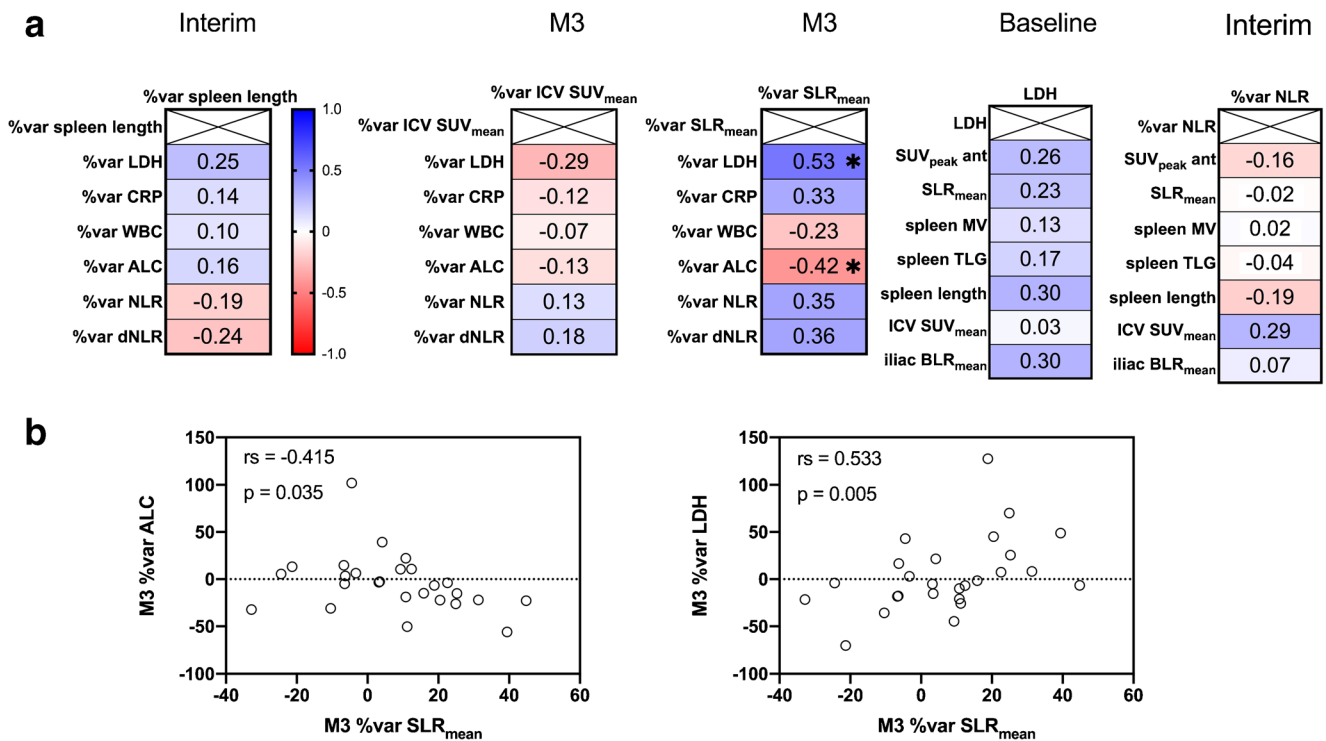**Fig. 2** Bland and Altman analysis between two observers for variables displaying the best ICC

Fig. 3 Correlation matrix using the non-parametric spearman method. **a** Among the best ICC, the 5 variables displayed a significative difference after Mann and Whitney tests according to the 1-year OS status.

Spearman rho values are displayed in each cell. Those displaying a significative *p* value are marked as *. **b** The two significative correlations are shown

Exploration of established biological variables of interest

Mean (SD) time (days) between biological samples and PET baseline, iPET, M3-PET and M6-PET was 18 (18), 5 (4), 4 (3) and 5 (5), respectively.

Baseline LDH values were higher in 1 year OS 1 patients [272.8 (84.3) versus 226.1 (85.1); $p = 0.04$] as well as i%var NLR [98.7 (100.9) versus 29.5 (56.5); $p = 0.049$]. Other biological variables were not significantly different between 1 year OS 0 and 1 year OS 1 patients (Table 3).

The exploration of correlations between biological variables of interest and all PET variables found a single significant positive correlation between baseline LDH and baseline WB MATV ($\rho = 0.445$, $p = 0.01$) (Fig. 3).

Figure 4 b shows the evolution of the main biology data. In 1 year OS 1 patients, a significative increase of NLR was found at M6-PET vs baseline PET ($p = 0.03$).

Change in imaging biomarkers across the treatment sequence

The evolution of PET variables according to 1 year OS is displayed in Fig. 4a. In the 1 year OS 1 group, from baseline, SLR_{mean} tended to increase at each time point to achieve a significant difference at M6-PET ($p = 0.019$). The same trends were observed with spleen SUV_{peak} anterior and spleen length. In the 1 year-OS 0 group, a significative increase of spleen length was found at iPET, as compared with baseline PET ($p = 0.014$) and M3-PET ($p = 0.0239$).

Survival analysis

Prediction of survival status at 12 months

Interim %var spleen length, M3%var SLR_{mean}, M3%var SUV_{mean} ileocecal valve, baseline LDH and i%var NLR ROC analyses were all statistically significant and showed that optimal cut-off values to predict an event were < 5%, > 25%, > 47%, > 196 UI/L and > 35%, respectively. The variable displaying the best performance was i%var spleen length: area under the curve = 0.800, sensitivity = 80.0%, specificity = 84.2%, accuracy = 82.8% ($p = 0.001$). All ROC analyses results can be found in Table 5. Of note, an extra-ROC analysis was conducted on baseline WB-MATV and was found to be significant (AUC = 0.718, $p = 0.026$). With an optimal cut-off value > 29 cc, sensitivity was equal to 70.0%, specificity to 73.7% and accuracy to 72.4%.

Kaplan-Meier survival analysis

Univariable Kaplan-Meier survival analyses were run using these previously described cut-off values and found that i%var

spleen length, M3%var SLR_{mean}, baseline LDH and i%var NLR were all predictors of 1 year OS. Kaplan-Meier curves for all variables are shown on Fig. 5. The parameter that seemed the most discriminant between 1 year OS 0 and 1 year OS 1 patients was M3%var SLR_{mean}. Patients with high M3%var SLR_{mean} (> 25%) had a 1 year OS of 0% whereas patients with low M3%var SLR_{mean} (≤ 25%) had a 1 year OS of 85.9% ($p < 0.0001$). Moreover, M6-PET imPERCIST response was also predictive of 1 year OS with poorer prognostic in patients classified as non-responders on M6-PET. Of note, the same analysis conducted on baseline MATV was also significant with poorer 1 year OS observed in patients displaying the highest values (> 29 cc): 41.15% versus 80% ($p = 0.018$).

Random forest prediction model

The multivariable random forest prediction model including i%var spleen length, M3%var SLR_{mean}, M3%var SUV_{mean} ileocecal valve, baseline LDH, i%var NLR and M6-PET response showed that the variable displaying the highest importance expressed as mean decrease accuracy was M3%var SLR_{mean} followed by M3%var SUV_{mean} ileocecal valve. Respective mean decrease accuracies were 5.34, 8.54, 6.71, 4.65, − 1.92 and 2.51. Moreover, a CHAID classification tree including all previously mentioned PET and biological variables of interest found that 96.3% of patients were well-classified when using first M3%var SLR_{mean} > 25% and then ileocecal valve M3%var SUV_{mean} > 47% as criteria of classification. Only one 1 year OS 1 patient was misclassified. The representative case of a well-classified patient can be seen in supplemental Figure 1.

Discussion

ICIs have transformed the management and the prognostic of patients with metastatic melanoma. However, some of them developed resistance to ICIs [2] and the majority is still refractory [28]. Therefore, achieving a gain in survival while also accounting for health economics implications [29] requires biomarkers to predict the outcome of ICI treatment. Amongst these potential biomarkers are PET metrics, which can be split into three categories: PET response criteria adapted to ICIs in order to account for pseudo-progression patterns, MATV and signs of immune activation.

To the best of our knowledge, this study is the first to report inter-observer variability in a wide range of signs of immune activation on ^{18}F-FDG PET. Strong reproducibility was observed for spleen and bone marrow measurements and showed that the SUV mean metric should be preferred. Ileocecal valve SUV performances appeared lower, with ICCs judged as moderate. In addition to harbouring a low
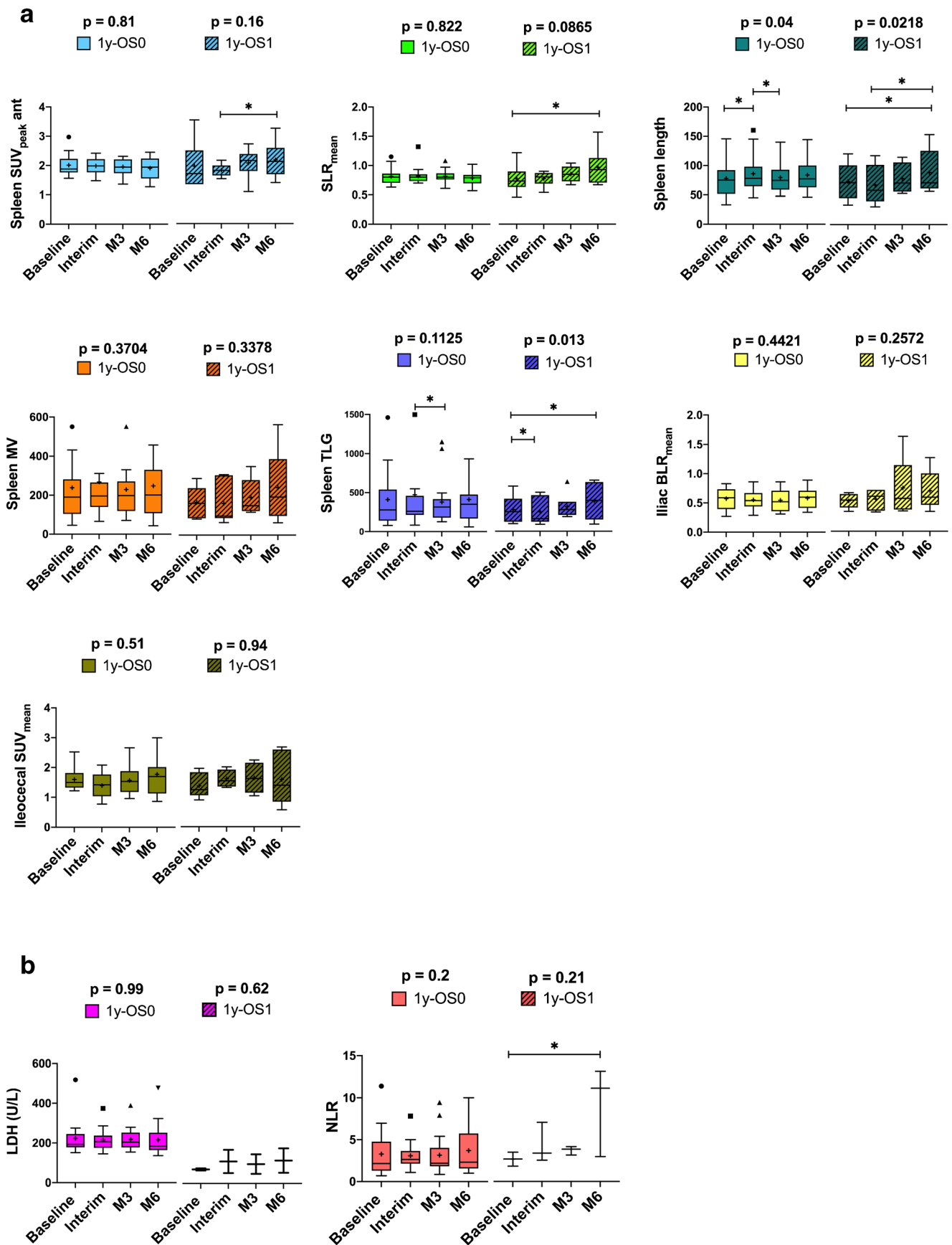
Fig. 4 Evolution of PET (**a**) and biological (**b**) variables during treatment with immune checkpoint inhibitors. Results are split according to the survival status at 1 year: 1 year OS as OS0 (survivors) on the left side and OS1 (deceased) on the right side of each graph. Friedman tests were used to compare each variable at different time points
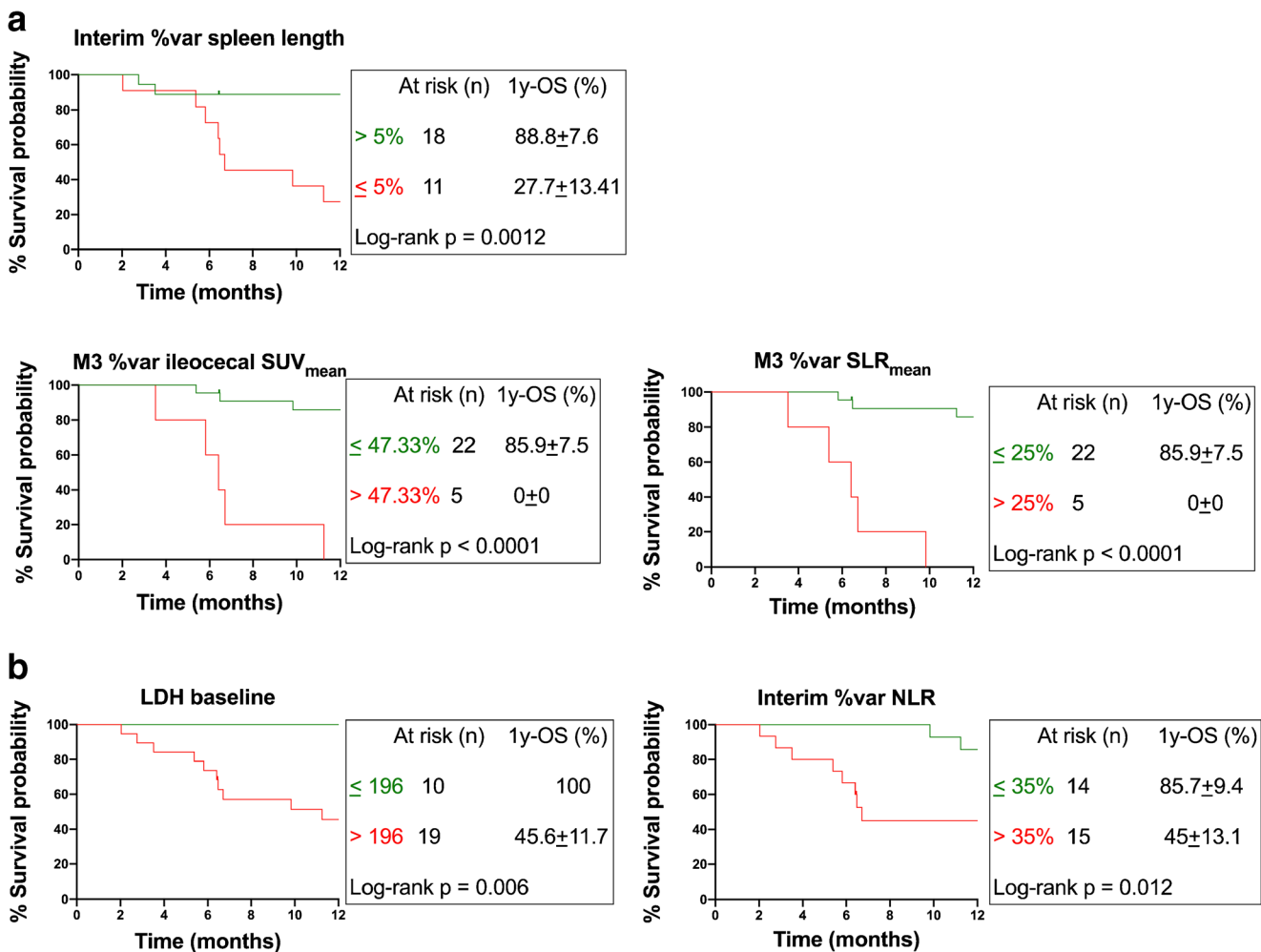
Fig. 5 Landmarked Kaplan-Meier estimates of 1-year overall survival (1 year OS) according to **a** 3 PET variables reported as the %var and **b** 2 biology variables

inter-observer variability, computation of signs of immune activation has to be user-friendly if one wants to use them for daily routine in busy PET centres. Therefore, despite an excellent ICC, spleen MV and TLG are unlikely to be adopted. Overall, SLR, BLR and spleen length appear the most suitable.

While understanding the baseline host anti-tumour immune response on baseline scan is important, its evolution under treatment has not yet been studied. Our study focused on the longitudinal evolution of several potential biomarkers extracted from ^{18}F-FDG PET/CT imaging, at baseline, after the 2 first injections of ICIs, and after 3 and 6 months of treatment.

The main finding of our study is that M3%var SLR$_{mean}$ was higher in patients who died during the first year of treatment and was the variable displaying the highest importance in a multivariable random forest prediction model including PET and biological variables significantly different in survivors and non-survivors on univariable analysis (i%var spleen length, M3%var SLR$_{mean}$, M3%var SUV$_{mean}$ ileocecal valve, baseline LDH, i%var NLR and M6-PET response). While the

attention of the PET community has been mainly focused on the capability of SLR to predict immune activation [13, 30] (an increased spleen uptake being considered to reflect "unleashed" T lymphocytes with an expected better outcome), several findings of our study show that the detrimental increase in SLR on M3PET may be related to inflammation. The spleen, being the largest lymphoid organ in the human body, is the site of immune cell activation and maturation [31]. The mechanism for high splenic avidity is not well understood but high splenic uptake or high SLR has been observed in patients with infections, granulomatous diseases as well as patients with cancers associated with an inflammatory state such as Hodgkin's lymphoma [16, 32–37]. Increased splenic avidity was also reported in GCSF injections and in the post-chemotherapy period [38]. In our study, SLR correlated well with the evolution of tumour burden WB-MATV at M3 and displayed a negative correlation with ALC at the same time point, known to be correlated with inflammation (Fig. 4).

Immune inflammatory cells can be actively tumour promoting and be associated with tumour aggressiveness and

immune suppression, resulting in failure of T cells to protect from cancer [39]. Inflammation promotes tumour through many mechanisms [40]. Currently, it is suggested that myeloid-derived suppressor cells (MDSCs) and maybe other inflammatory processes, which induce expression on MDSCs, could lead to a resistance to ICIs. MDSCs are a heterogeneity population of cells [41] and could explain the increase of SLR. A major mechanism recently discovered is that inflammation induces the production of MDSCs which inhibit both adaptive and innate immunity to suppress T cells, leading to ICI failure [41]. MDSCs accumulated in blood, bone marrow lymph nodes, spleen and tumour sites. Preclinical studies have shown that MDSCs represent up to 20% of splenocytes in cancer disease vs 3% in healthy mice [42].

Also important as a prognosticator was M3%var ileocecal valve SUV_{mean}, with a CHAID classification tree including PET and biological variables of interest demonstrating that 96.3% of patients were well-classified when using first M3%var SLR_{mean} > 25% and then M3%var ileocecal valve SUV_{mean} > 47% as criteria of classification.

In addition to signs of immune activation and conventional or ICI-adapted PET response criteria, the third piece of information assessable on PET imaging is the occurrence of irAEs. While several studies have shown that patients experiencing irAEs may have better survival [43], studies on PET irAEs are scarce [44, 45]. Recently, Iravani et al. showed in melanoma patients receiving a combination of two ICIs that ^{18}F-FDG PET/CT often could detect relevant irAEs which may precede clinical diagnosis [46]. According to our study, most of irAEs appeared early in the course of treatment but were not found to predict 1 year OS.

This warrants validation of %varSLR as a prognosticator with prospective studies or larger series by pooling our EARL-compliant data with those of other EARL-accredited PET centres.

Our study has several limitations. First, its retrospective nature and its relatively small sample size, except for the reproducibility study where 14 variables involving immune organs were studied at 3 time points. Secondly, we did not use time-dependent ROC curves [47]. Thirdly, a random forest method was used on a small amount of data and including multiple variables. Given the risk of over fitting, these results should be considered as a proof of concept requiring prospective evaluation.

Conclusion

Based on its good reproducibility and practicality, SLR_{mean} is recommended as a prognosticator in melanoma patients under immunotherapy: its increase greater than 25% at 3 months, compared with baseline, was associated with poor outcome after ICIs. This study provides a proof of concept concerning this parameter and warrants prospective evaluation.

Compliance with ethical standards

Competing interest The authors declare that they have no competing interest.

Ethical approval 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.

Consent to participate and consent for publication Institutional review board approval was not required because in accordance with European regulation, French observational studies without any additional therapy or monitoring procedure do not need the approval of an ethics committee. Nonetheless, in accordance with the European General Data Protection Regulation, we sought approval to collect and publish data for this work from the national committee for data privacy, with the registration no. 2081250 v 0.

References

1. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med. 2010;363:711–23. https://doi.org/10.1056/NEJMoa1003466.

2. Robert C, Schachter J, Long GV, Arance A, Grob JJ, Mortier L, et al. Pembrolizumab versus ipilimumab in advanced melanoma. N Engl J Med. 2015;372:2521–32. https://doi.org/10.1056/NEJMoa1503093.

3. Herbst RS, Baas P, Kim DW, Felip E, Pérez-Gracia JL, Han JY, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. Lancet (London, England). 2016;387:1540–50. https://doi.org/10.1016/s0140-6736(15)01281-7.

4. Tomita M, Yasui H, Higashikawa K, Nakajima K, Takakura H, Shiga T, et al. Anti PD-1 treatment increases [(18)F]FDG uptake by cancer cells in a mouse B16F10 melanoma model. EJNMMI Res. 2018;8:82. https://doi.org/10.1186/s13550-018-0433-1.

5. Sachpekidis C, Larribere L, Pan L, Haberkorn U, Dimitrakopoulou-Strauss A, Hassel JC. Predictive value of early 18F-FDG PET/CT studies for treatment response evaluation to ipilimumab in metastatic melanoma: preliminary results of an ongoing study. Eur J Nucl Med Mol Imaging. 2015;42:386–96. https://doi.org/10.1007/s00259-014-2944-y.

6. Hodi FS, Hwu WJ, Kefford R, Weber JS, Daud A, Hamid O, et al. Evaluation of immune-related response criteria and RECIST v1.1 in patients with advanced melanoma treated with pembrolizumab. J Clin Oncol: Off J Am Soc Clin Oncol. 2016;34:1510–7. https://doi.org/10.1200/jco.2015.64.0391.

7. Ito K, Teng R, Schoder H, Humm JL, Ni A, Michaud L, et al. (18)F-FDG PET/CT for Monitoring of ipilimumab therapy in patients with metastatic melanoma. Journal of nuclear medicine : official publication. Soc Nucl Med. 2019;60:335–41. https://doi.org/10.2967/jnumed.118.213652.

8. Seban RD, Moya-Plana A, Antonios L, Yeh R, Marabelle A, Deutsch E, et al. Prognostic 18F-FDG PET biomarkers in metastatic mucosal and cutaneous melanoma treated with immune checkpoint inhibitors targeting PD-1 and CTLA-4. Eur J Nucl Med Mol Imaging. 2020. https://doi.org/10.1007/s00259-020-04757-3.

9. Ayati N, Sadeghi R, Kiamanesh Z, Lee ST, Zakavi SR, Scott AM. The value of (18)F-FDG PET/CT for predicting or monitoring immunotherapy response in patients with metastatic melanoma: a systematic review and meta-analysis. Eur J Nucl Med Mol Imaging. 2020. https://doi.org/10.1007/s00259-020-04967-9.

10. Kolinger GD, Vallez Garcia D, Kramer GM, Frings V, Smit EF, de Langen AJ, et al. Repeatability of [(18)F]FDG PET/CT total metabolic active tumour volume and total tumour burden in NSCLC patients. EJNMMI research. 2019;9:14. https://doi.org/10.1186/s13550-019-0481-1.

11. Lodge MA. Repeatability of SUV in Oncologic (18)F-FDG PET. J Nucl Med: Off Publ, Soc Nucl Med. 2017;58:523–32. https://doi.org/10.2967/jnumed.116.186353.

12. Yeh R, Trager MH, Rizk EM, Finkel GG, Barker LW, Carvajal RD, et al. FLT-PET at 6 weeks predicts response assessed by CT at 12 weeks in melanoma patients treated with pembrolizumab. Clin Nucl Med. 2020;45:267–75. https://doi.org/10.1097/rlu.0000000000002967.

13. Aide N, Hicks RJ, Le Tourneau C, Lheureux S, Fanti S, Lopci E. FDG PET/CT for assessing tumour response to immunotherapy: report on the EANM symposium on immune modulation and recent review of the literature. Eur J Nucl Med Mol Imaging. 2019;46:238–50. https://doi.org/10.1007/s00259-018-4171-4.

14. Wong A, Callahan J, Keyaerts M, Neyns B, Mangana J, Aberle S, et al. (18)F-FDG PET/CT based spleen to liver ratio associates with clinical outcome to ipilimumab in patients with metastatic melanoma. Cancer Imaging: Off Publ Int Cancer Imaging Soc. 2020;20:36. https://doi.org/10.1186/s40644-020-00313-2.

15. Seban RD, Nemer JS, Marabelle A, Yeh R, Deutsch E, Ammari S, et al. Prognostic and theranostic 18F-FDG PET biomarkers for anti-PD1 immunotherapy in metastatic melanoma: association with outcome and transcriptomics. Eur J Nucl Med Mol Imaging. 2019;46:2298–310. https://doi.org/10.1007/s00259-019-04411-7.

16. Dercle L, Seban RD, Lazarovici J, Schwartz LH, Houot R, Ammari S, et al. (18)F-FDG PET and CT scans detect new imaging patterns of response and progression in patients with Hodgkin lymphoma treated by anti-programmed death 1 immune checkpoint inhibitor. J Nucl Med: Off Publ Soc Nucl Med. 2018;59:15–24. https://doi.org/10.2967/jnumed.117.193011.

17. Capone M, Giannarelli D, Mallardo D, Madonna G, Festino L, Grimaldi AM, et al. Baseline neutrophil-to-lymphocyte ratio (NLR) and derived NLR could predict overall survival in patients with advanced melanoma treated with nivolumab. J Immunother Cancer. 2018;6:74. https://doi.org/10.1186/s40425-018-0383-1.

18. Wagner NB, Forschner A, Leiter U, Garbe C, Eigentler TK. S100B and LDH as early prognostic markers for response and overall survival in melanoma patients treated with anti-PD-1 or combined anti-PD-1 plus anti-CTLA-4 antibodies. Br J Cancer. 2018;119:339–46. https://doi.org/10.1038/s41416-018-0167-x.

19. 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.

20. 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.

21. Nioche C, Orlhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, et al. LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. Cancer Res. 2018;78:4786–9. https://doi.org/10.1158/0008-5472.Can-18-0125.

22. Bezerra AS, D'Ippolito G, Faintuch S, Szejnfeld J, Ahmed M. Determination of splenomegaly by CT: is there a place for a single measurement? AJR Am J Roentgenol. 2005;184:1510–3. https://doi.org/10.2214/ajr.184.5.01841510.

23. Mekki A, Dercle L, Lichtenstein P, Marabelle A, Michot JM, Lambotte O, et al. Detection of immune-related adverse events by medical imaging in patients treated with anti-programmed cell death 1. Eur J Cancer (Oxf, Engl: 1990). 2018;96:91–104. https://doi.org/10.1016/j.ejca.2018.03.006.

24. Mekki A, Dercle L, Lichtenstein P, Nasser G, Marabelle A, Champiat S, et al. Machine learning defined diagnostic criteria for differentiating pituitary metastasis from autoimmune hypophysitis in patients undergoing immune checkpoint blockade therapy. Eur J Cancer (Oxf, Engl: 1990). 2019;119:44–56. https://doi.org/10.1016/j.ejca.2019.06.020.

25. Prigent K, Aide N. (18)F-Fludeoxyglucose PET/computed tomography for assessing tumor response to immunotherapy and detecting immune-related side effects: a checklist for the PET reader. PET Clin. 2020;15:1–10. https://doi.org/10.1016/j.cpet.2019.08.006.

26. Barrington SF, Meignan M. Time to Prepare for Risk Adaptation in Lymphoma by Standardizing Measurement of Metabolic Tumor Burden. J Nucl Med: Off Publ Soc Nucl Med. 2019;60:1096–102. https://doi.org/10.2967/jnumed.119.227249.

27. Breiman L. Classification and regression trees. New York: Routledge; 1984.

28. Zhao X, Subramanian S. Intrinsic resistance of solid tumors to immune checkpoint blockade therapy. Cancer Res. 2017;77:817–22. https://doi.org/10.1158/0008-5472.Can-16-2379.

29. Tartari F, Santoni M, Burattini L, Mazzanti P, Onofri A, Berardi R. Economic sustainability of anti-PD-1 agents nivolumab and pembrolizumab in cancer patients: recent insights and future challenges. Cancer Treat Rev. 2016;48:20–4. https://doi.org/10.1016/j.ctrv.2016.06.002.

30. Wong ANM, McArthur GA, Hofman MS, Hicks RJ. The advantages and challenges of using FDG PET/CT for response assessment in melanoma in the era of targeted agents and immunotherapy. Eur J Nucl Med Mol Imaging. 2017;44:67–77. https://doi.org/10.1007/s00259-017-3691-7.

31. Mebius RE, Kraal G. Structure and function of the spleen. Nat Rev Immunol. 2005;5:606–16. https://doi.org/10.1038/nri1669.

32. Salaun PY, Gastinne T, Bodet-Milin C, Campion L, Cambefort P, Moreau A, et al. Analysis of 18F-FDG PET diffuse bone marrow uptake and splenic uptake in staging of Hodgkin's lymphoma: a reflection of disease infiltration or just inflammation? Eur J Nucl Med Mol Imaging. 2009;36:1813–21. https://doi.org/10.1007/s00259-009-1183-0.

33. Chen A, Mokrane FZ, Schwartz LH, Morschhauser F, Stamatoullas A, Schiano de Colella JM, et al. Early (18)F-FDG PET/CT response predicts survival in relapsed or refractory Hodgkin lymphoma treated with nivolumab. J Nucl Med. 2020;61:649–54. https://doi.org/10.2967/jnumed.119.232827.

34. Mokrane FZ, Chen A, Schwartz LH, Morschhauser F, Stamatoullas A, Schiano de Colella JM, et al. Performance of CT Compared with (18)F-FDG PET in predicting the efficacy of nivolumab in relapsed or refractory Hodgkin lymphoma. Radiology. 2020;295:651–61. https://doi.org/10.1148/radiol.2020192056.

35. Dercle L, Ammari S, Seban RD, Schwartz LH, Houot R, Labaied N, et al. Kinetics and nadir of responses to immune checkpoint blockade by anti-PD1 in patients with classical Hodgkin lymphoma. Eur J Cancer (Oxf, Engl: 1990). 2018;91:136–44. https://doi.org/10.1016/j.ejca.2017.12.015.

36. Dercle L, Mokrane FZ, Schiano de Colella JM, Stamatoullas A, Morschhauser F, Brice P, et al. Unconventional immune-related phenomena observed using 18F-FDG PET/CT in Hodgkin lymphoma treated with anti PD-1 monoclonal antibodies. Eur J Nucl Med Mol Imaging. 2019;46:1391–2. https://doi.org/10.1007/s00259-019-04310-x.

37. Kim SY, Moon CM, Yoon HJ, Kim BS, Lim JY, Kim TO, et al. Diffuse splenic FDG uptake is predictive of clinical outcomes in patients with rectal cancer. Sci Rep. 2019;9:1313. https://doi.org/10.1038/s41598-018-35912-4.

38. Oliveira M, Lasnon C, Nganoa C, Gac AC, Damaj G, Aide N. Comprehensive analysis of the influence of G-CSF on the biodistribution of (18)F-FDG in lymphoma patients: insights for PET/CT scheduling. EJNMMI Res. 2019;9:79. https://doi.org/10.1186/s13550-019-0546-1.

39. Chang CH, Qiu J, O'Sullivan D, Buck MD, Noguchi T, Curtis JD, et al. Metabolic competition in the tumor microenvironment is a driver of cancer progression. Cell. 2015;162:1229–41. https://doi.org/10.1016/j.cell.2015.08.016.

40. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144:646–74. https://doi.org/10.1016/j.cell.2011.02.013.

41. Ostrand-Rosenberg S, Sinha P. Myeloid-derived suppressor cells: linking inflammation and cancer. J Immunol (Baltimore, Md : 1950). 2009;182:4499–506. https://doi.org/10.4049/jimmunol.0802740.

42. Jordan KR, Kapoor P, Spongberg E, Tobin RP, Gao D, Borges VF, et al. Immunosuppressive myeloid-derived suppressor cells are increased in splenocytes from cancer patients. Cancer Immunol Immunother. 2017;66:503–13. https://doi.org/10.1007/s00262-016-1953-z.

43. Kotwal A, Kottschade L, Ryder M. PD-L1 inhibitor-induced thyroiditis is associated with better overall survival in cancer patients. Thyroid: Off J Am Thyroid Assoc. 2020;30:177–84. https://doi.org/10.1089/thy.2019.0250.

44. Nobashi T, Baratto L, Reddy SA, Srinivas S, Toriihara A, Hatami N, et al. Predicting response to immunotherapy by evaluating tumors, lymphoid cell-rich organs, and immune-related adverse events using FDG-PET/CT. Clin Nucl Med. 2019;44:e272–e9. https://doi.org/10.1097/rlu.0000000000002453.

45. Wachsmann JW, Ganti R, Peng F. Immune-mediated disease in ipilimumab immunotherapy of melanoma with FDG PET-CT. Acad Radiol. 2017;24:111–5. https://doi.org/10.1016/j.acra.2016.08.005.

46. Iravani A, Osman MM, Weppler AM, Wallace R, Galligan A, Lasocki A, et al. FDG PET/CT for tumoral and systemic immune response monitoring of advanced melanoma during first-line combination ipilimumab and nivolumab treatment. Eur J Nucl Med Mol Imaging. 2020. https://doi.org/10.1007/s00259-020-04815-w.

47. Kamarudin AN, Cox T, Kolamunnage-Dona R. Time-dependent ROC curve analysis in medical research: current methods and applications. BMC Med Res Methodol. 2017;17:53. https://doi.org/10.1186/s12874-017-0332-6.