

REVIEW

Open Access



# Review on radiomic analysis in $^{18}\text{F}$ -fluorodeoxyglucose positron emission tomography for prediction of melanoma outcomes

Karim Amrane<sup>1,2\*</sup> , Coline Le Meur<sup>3</sup>, Philippe Thuillier<sup>4,5</sup>, Christian Berthou<sup>2,6</sup>, Arnaud Uguen<sup>2,7</sup>, Désirée Deandreis<sup>8</sup>, David Bourhis<sup>5,9</sup>, Vincent Bourbonne<sup>3,10</sup> and Ronan Abgral<sup>5,9</sup>

## Abstract

Over the past decade, several strategies have revolutionized the clinical management of patients with cutaneous melanoma (CM), including immunotherapy and targeted tyrosine kinase inhibitor (TKI)-based therapies. Indeed, immune checkpoint inhibitors (ICIs), alone or in combination, represent the standard of care for patients with advanced disease without an actionable mutation. Notably BRAF combined with MEK inhibitors represent the therapeutic standard for disease disclosing BRAF mutation. At the same time, FDG PET/CT has become part of the routine staging and evaluation of patients with cutaneous melanoma. There is growing interest in using FDG PET/CT measurements to predict response to ICI therapy and/or target therapy. While semiquantitative values such as standardized uptake value (SUV) are limited for predicting outcome, new measures including tumor metabolic volume, total lesion glycolysis and radiomics seem promising as potential imaging biomarkers for nuclear medicine. The aim of this review, prepared by an interdisciplinary group of experts, is to take stock of the current literature on radiomics approaches that could improve outcomes in CM.

**Keywords** Melanoma, Radiomic, FDG-PET, Immune checkpoint inhibition, Immunotherapy, BRAF, NRAS

\*Correspondence:

Karim Amrane  
kamrane@ch-morlaix.fr

<sup>1</sup>Department of Oncology, Regional Hospital of Morlaix, Morlaix 29600, France

<sup>2</sup>Lymphocytes B et Autoimmunité, Inserm, UMR1227, Univ Brest, Inserm, LabEx IGO, Brest, France

<sup>3</sup>Department of Radiotherapy, University Hospital of Brest, Brest, France

<sup>4</sup>Department of Endocrinology, University Hospital of Brest, Brest, France

<sup>5</sup>UMR Inserm 1304 GETBO, University of Western Brittany, Brest IFR 148, France

<sup>6</sup>Department of Hematology, University Hospital of Brest, Brest, France

<sup>7</sup>Department of Pathology, University Hospital of Brest, Brest, France

<sup>8</sup>Department of Nuclear Medicine, Gustave Roussy Institute, University of Paris Saclay, Paris, France

<sup>9</sup>Department of Nuclear Medicine, University Hospital of Brest, Brest, France

<sup>10</sup>Inserm, UMR1101, LaTIM, University of Western Brittany, Brest, France

## Introduction

Cutaneous melanoma (CM) results from genetic mutations in melanocytes and accounts for approximately 1% of all malignant skin tumors.

It mainly affects the Caucasian population of both sexes [1], over 75 years of age [2] and is the 15th most common cancer worldwide. Its incidence has increased more rapidly than other cancers in recent decades [3] and varies considerably between countries due to differences in UV exposure and skin phototype [4]. The 3 countries with the highest incidence rate of CM are in Europe (Denmark, Sweden, and the Netherlands). In case of distant metastases, the prognosis is very poor; with a 5-year



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

survival rate of 99.4% (95% CI: 98.9–100.0) for stage I versus 23.0% (95% CI: 10.3–51.4) for stage IV [3] according to the AJCC melanoma staging system [5].

Several common prognostic factors have been highlighted in the literature, including age, gender, nutritional status, performance status (PS), tumor size, Breslow thickness, mitotic rate and ulceration, lymph node involvement, distant metastasis, anatomical site (axial localization), surgery of the primary lesion, time-to-progression after first-line therapy, and biomarkers (BRAF, NRAS) status [4, 6].

Immunotherapy, such as immune checkpoint inhibitors (ICIs) that block CTLA-4 (e.g., ipilimumab) [7], PD-1 (e.g., nivolumab, pembrolizumab) [8–10] and LAG3 (e.g., relatimab) [11], have demonstrated objective tumor regression in patients with advanced melanoma, using the activation of the immune system to generate an anti-tumor response.

The 5-year survival rate for CM with ipilimumab is estimated to be between 12.3% and 16.5%, with a prolonged response time [12], while nivolumab has a 5-year survival rate of 39% with a response rate of 42% [13]. On the other hand, the combination of anti-PD1 with anti-CTLA4 (ipilimumab) or anti-LAG3 (relatimab) has shown better results than monotherapy. Indeed, the nivolumab-ipilimumab combination showed a 5-year survival rate of 52%, and a response rate of 58%; while data on the nivolumab-relatimab combination are still immature. However, the 12-month progression-free survival (PFS) rate is 47.7% [14, 15]. Tyrosine kinase inhibitors (TKIs), which target kinases in the mitogen-activated protein kinase (MAPK) pathway, are approved for metastatic BRAF-mutant CM, and prolong patient survival. Response to these therapies is limited by drug resistance and is less durable than with ICIs. Early clinical trials have shown that BRAF inhibitors (BRAFi) (vemurafenib, dabrafenib and encorafenib) improve the median PFS compared with chemotherapy in patients with BRAF-activating somatic mutations [16, 17]. However, the efficacy of BRAFi as monotherapy is limited due to the development of drug resistance, resulting in little difference in long-term survival rates compared with chemotherapy [18, 19]. Targeted therapies combining BRAFi and MEKi (dabrafenib plus trametinib, vemurafenib plus cobimetinib, encorafenib plus binimetinib), which act downstream of BRAF in the MAPK pathway, have been tested to enhance inhibition of the MAPK pathway and delay acquired resistance to BRAFi. Combination therapy improved median PFS compared with BRAFi monotherapy [20, 21].

Although most patients respond immediately to BRAFi-MEKi combination therapy (objective response rate (ORR) of around 70%), long-term PFS after treatment affects only a subset of patients (5-year PFS rates range from 14–22.9% [22–24]). Whereas ICI combination

therapy has a slightly lower ORR (around 60%) but shows a more durable response in BRAF-mutant CM (36% 5-year PFS rate of 36% and 6.5-year PFS rate of 34% with nivolumab plus ipilimumab [15, 25]). Finally, as suggested by phase II studies, combinations with ICI have been evaluated to enhance MAPK inhibition and delay acquired resistance to BRAFi MEKi [26]. However, only one trial was able to compare the ICI-BRAFi-MEKi combination against BRAFi-MEKi, but did not show superiority in patients with advanced BRAFV600-mutated melanoma [27].

Despite the paradigm shift brought about by ICIs (prolonged survival and good tolerance [28–31]), 40–65% of metastatic melanomas do not respond to mono- or combo-ICIs and more than 43% of patients develop secondary resistance after an initial response after 3 years of treatment [28].

Pre-therapeutic  $^{18}\text{F}$ -fluorodeoxyglucose positron-emission tomography (FDG-PET/CT) is currently indicated for the staging of locally advanced CM [32]. In addition, several studies have shown that tumor uptake, assessed by different standard quantitative static (SUV=standardized uptake value), and volumetric (MTV=metabolic tumor volume, TLG=total lesion glycolysis) parameters, is a prognostic factor for survival in various solid cancers, including CM [33–35]. However, no clear cut-off value for the selection of patients with poor prognosis has been validated in practice.

More recently, radiomic approach has been developed to characterize tumor heterogeneity by extracting textural features (TFs) from different imaging modalities, including PET. CM also exhibits high biological heterogeneity with hypoxic areas, necrotic regions, zones of high cellular proliferation and intra-tumoral angiogenic heterogeneity [36, 37]. SUV, MTV and TLG can by definition be considered as first-order radiomic features, but do not provide information concerning on the spatial distribution of voxel values. Therefore, second and higher order parameters allow the modelling of this spatial relationship with numerous TFs [38]. This image-based approach may lead to a more personalized therapy in this particular poor prognosis cancer entity, where intensification of treatment could lead to improved clinical outcomes. Given the narrow therapeutic window of such approaches, efficient prognostic and predictive tools are needed. Texture analysis of the tumor on pre-therapeutic FDG-PET/CT may also be useful, although there are currently many pitfalls in radiomics in terms of reproducibility and calibration of PET machines. Nevertheless, the development of artificial intelligence and deep learning brings a new insights and could improve the robustness of this technique [39].

Given the significant potential of radiomics, we first present an overview of the basics of radiomics, then the

potential contribution of radiomics to prognosis, the predictive approach to mutation status and, finally, we discuss the methodological aspect and its limitations before turning to strategies for harmonizing practices.

### Basics of radiomics

Radiomics is a process based on the extraction of large sets of quantitative features (e.g., volume, shape, intensity, texture, etc.) from medical images, using a high-throughput computer system [40]. Radiomics aims to extract the full power of images in order to improve clinical decision support by integrating image-derived information into the clinic.

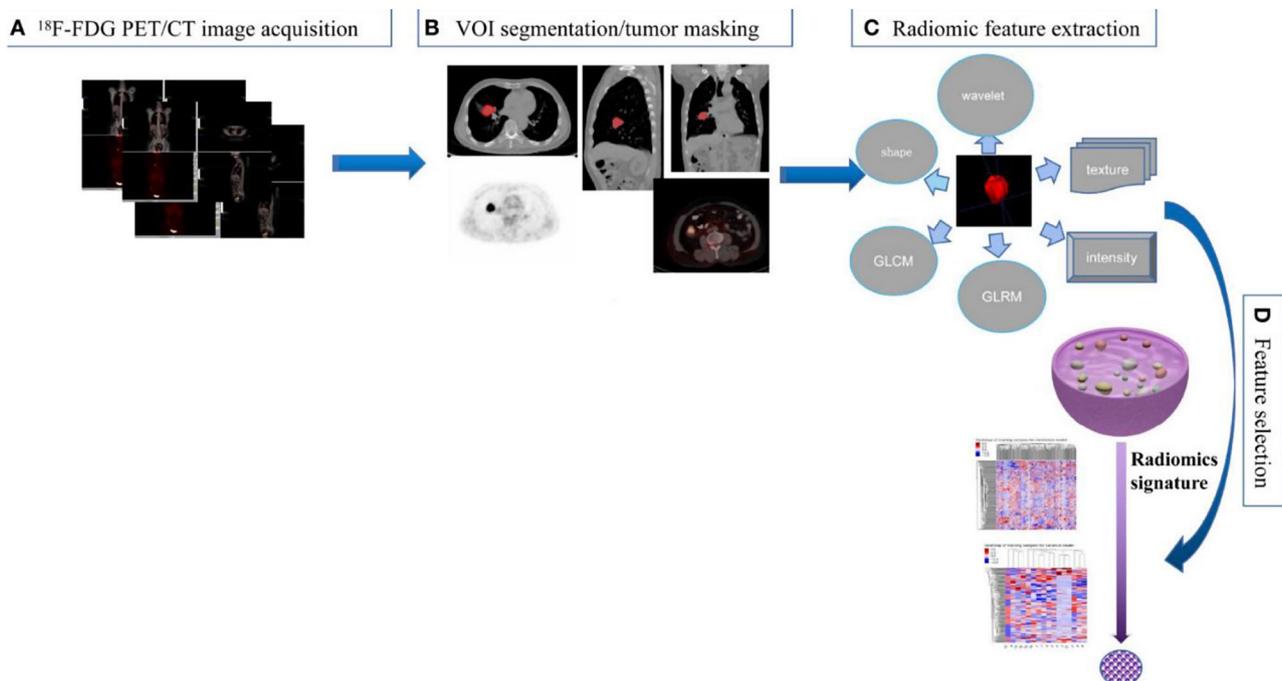
The radiomics workflow consists of several steps: (a) image acquisition, (b) VOI segmentation, (c) radiomic feature extraction and (d) feature selection (Fig. 1). [41]

Several types of data can be extracted from PET imaging. Qualitative semantic features are commonly used in the clinical lexicon to describe lesions [42]. They have different levels of complexity, expressing firstly the properties of the lesion shape and voxel intensity histogram, and then those of the spatial arrangement of intensity values at the voxel level (texture). They can be extracted directly from the images or after applying various filters or transformations (e.g., the wavelet transform) (Fig. 2– Table 1) [38].

First-order features, or histograms, describe the distribution of individual voxel values or gray-scale intensities

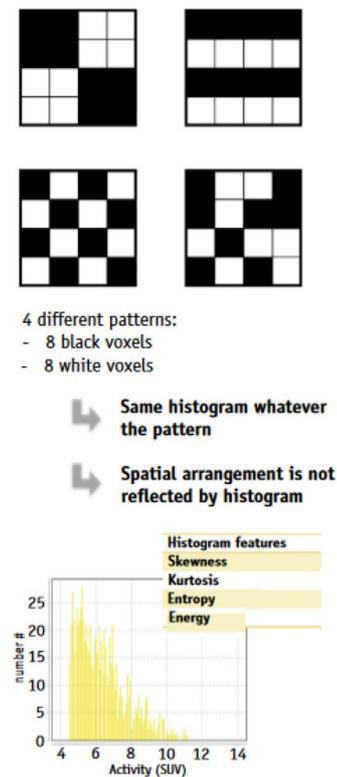
within a VOI, independent of spatial relationships. Histogram features are therefore not TFs, but describe the shape of the intensity histogram, such as kurtosis or flattening. Different spatial arrangements of gray levels, corresponding to the same histogram, can therefore give rise to identical index values. From the histogram, it is possible to calculate conventional parameters such as maximum (SUVmax), mean (SUVmean), peak (SUV peak defined as highest mean value in a spherical VOI of 1 cc), metabolic tumor volume (MTV), total lesion glycolysis (TLG=MTV x SUVmean), and skewness, kurtosis, energy and entropy.

Second- and third-order statistical features include so-called TFs [43, 44], which are obtained by calculating the statistical relationships between neighboring voxels [38]. They provide a measure of the spatial arrangement of voxel intensities, and thus of intra-lesional heterogeneity. These features can be derived from the Gray Level Co-occurrence Matrix (GLCM), which quantifies the occurrence of voxels of the same intensity at a given distance along a fixed direction, or the Gray Level Run Length Matrix (GLRLM), which quantifies successive voxels of the same intensity along fixed directions. The Gray Level Size-Zone Matrix (GLSZM) measures the number of neighboring voxels of the same intensity, for each intensity value. The Gray-Level Difference Matrix (NGDLM) measures the difference in intensity between neighboring voxels.

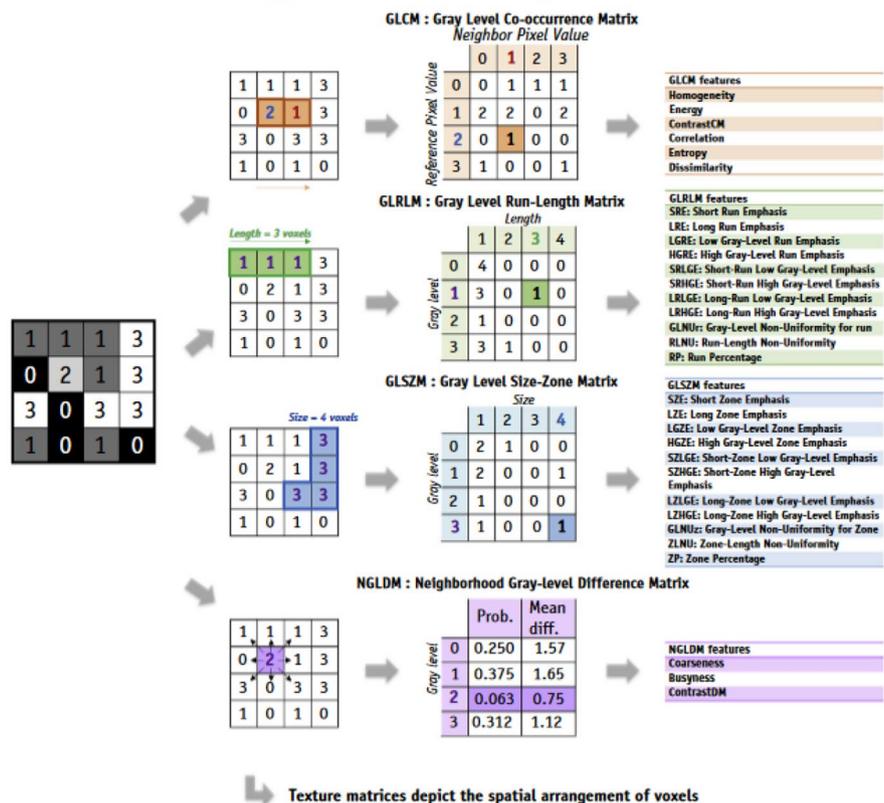


**Fig. 1** The radiomics workflow consists of several steps: (a) image acquisition, (b) VOI segmentation, (c) radiomic feature extraction and (d) feature selection [41]. Is taken from the article Radiomics in Nuclear Medicine Applied to Radiation Therapy: Methods, Pitfalls and Challenges. Sylvain Reuzé, Antoine Schernberg, Fanny Orhac, Roger Sun, Cyrus Chargari, Laurent Derclé, Eric Deutsch, Irène Buvat, Charlotte Robert [38]. License number authorizing re-use obtained from the publisher ELSEVIER / 5,660,690,860,547

**A Histogram or 1<sup>st</sup>-order features**



**B Textural or 2<sup>nd</sup>-order features**



**Fig. 2** Common texture matrices and derived texture indexes (Abbreviations: defined in the main text) [38]. Is from the article PET/CT Radiomic Features: A Potential Biomarker for EGFR Mutation Status and Survival Outcome Prediction in NSCLC Patients Treated With TKIs Published in *Front. Oncol.*, 21 June 2022 s.Thoracic Oncology Volume 12–2022 | <https://doi.org/10.3389/fonc.2022.894323>. Copyright © 2022 Yang, Xu, Li, Wang, Peng, Zhang, Wu, Chu, Wang, Meng and Zhang [41]

Higher-order statistical features are obtained using statistical methods after applying filters or mathematical transformations to the images, for example to identify repetitive or non-repetitive patterns, remove noise or highlight details. These include fractal analysis, Minkowski functions, wavelet transforms and Laplacian transforms of Gaussian-filtered images, which can extract areas of increasingly coarse texture patterns.

**Indications of FDG PET/CT**

FDG PET/CT plays a clear role in the diagnosis of advanced melanoma, and well-designed studies will determine whether it is appropriate to include FDG PET/CT in the management and follow-up of patients with metastatic CM. FDG PET is recommended for AJCC stage III (regional lymph nodes) and IV (systemic) melanoma for staging, restaging and follow-up treatment, with a significant implications on management [32]. For primary staging of stage IV melanoma, overall estimates of the sensitivity of FDG PET/CT over CT were 80% versus 51%, and for the specificity 87% versus 69%. However, although FDG PET/CT is an excellent tool, it

is not currently recommended for routine follow-up of asymptomatic patients with stage I to IIIA disease or for prognostic evaluation of CM [32]. FDG-PET can detect clinically relevant immune-related adverse events prior to clinical diagnosis, including endocrinopathies and enterocolitis [45].

Finally, it may help to understand the distinction between the categories of pseudoprogression, hyperprogression, dissociated/mixed response and sustained response in the interpretation of FDG-PET, several studies are underway [46].

**Methods**

**Search strategy**

Two authors (K.A. and C.L.M.) together performed a computerized bibliographic search of the Medline databases in September 2023. The following search terms (“Melanoma”) AND (“PET” OR “positron emission tomography”) AND (“radiomics” OR “radiomic” OR “texture” OR “textural”) AND (“Immune checkpoint-inhibitor” OR “Immunotherapy”) were used according to a free

**Table 1** Groups of features

Features class	Features
Conventional	SUVmean
	SUVmax
	SUVpeak
	Volume
	TLG
1 <sup>st</sup> order	Skewness
	Kurtosis
	Entropy
	Energy
GLCM	Homogeneity
	Energy
	Contrast
	Correlation
	Entropy
GLRLM	Dissimilarity
	SRE (short run emphasis)
	LRE (long run emphasis)
	LGRE (low gray-level run emphasis)
	HGRE (high gray-level run emphasis)
	SRLGE (short run low gray-level emphasis)
	SRHGE (short run high gray-level emphasis)
GLNUR (gray-level non-uniformity for run)	
NGLDM	RLNU (runlength non-uniformity)
	RP (run percentage)
	Coarseness
	Contrast
GLSZLM	Busyness
	SZE (short-zone emphasis)
	LZE (long-zone emphasis)
	LGZE (low grey-level zone emphasis)
	HGZE (high gray-level zone emphasis)
	SZLGE (short-zone low gray-level emphasis)
	SZHGE (short-zone high gray-level emphasis)
	LZLGE (large zone low gray-level emphasis)
	LZHGE (large zone high gray-level emphasis)
	GLNU (Gray level non-uniformity)
	ZLNU (zone length nonuniformity)
ZP (zone percentage)	

text protocol that applied only “Humans” and “English language” filters without time-period restriction.

### Criteria

Records were considered relevant to this review if they included melanoma FDG-PET/CT textural analysis and its prognosis impact, predictive mutational status approach with radiomic, methodological aspect of radiomic studies and practice harmonization. Case reports, editorials, letters, and meeting abstracts were excluded during the review process.

Studies were selected according to the following criteria: (i) only studies that included patients who had been treated for melanoma; (ii) only studies with patient numbers  $\geq 50$ ; (iii) only studies that included exclusive pre-therapeutic FDG-PET; (iiii) only studies that included textural features analysis.

### Review process

After duplicates were removed, two authors (K.A. and C.L.M.) completed an independent review of 66 abstracts to ultimately select 10 studies for separate full-text evaluation. Any discrepancies in study inclusion were resolved by consulting the supervisor of the review process (V.B. and R.A.).

### Prognostic approach

Radiomic application hold promise for treatment personalization. Several parameters can be extracted and combined, including standard descriptors (e.g., size, morphology, TNM stage (tumor, lymph nodes, metastasis)), qualitative, semi-quantitative physiological parameters (e.g., contrast enhancement, scattering characteristics, tracer uptake), and additional agnostic features that are otherwise “invisible” using bioinformatic approaches. Of these features, texture-based features have been the most widely studied, particularly in FDG-PET/CT [47] and several prognostic approaches have been investigated.

Ito et al. analyzed the prognostic value of pre-therapeutic metabolic FDG-PET/CT parameters in a retrospective study of 142 patients with only advanced CM who received second-line ipilimumab after chemotherapy. As criteria of tumor metabolic activity, the authors determined the SUV corrected for lean body mass (SUL) of all lesions and to assess tumor burden, they defined MTV as the volume bounded by a 42% SUVmax threshold of tumor lesions. Occasional manual adjustments were made when the defined volume exceeded the limits of the lesions seen on CT. Total body (TB) MTV and TLG, defined as the sum of MTV and TLG of each hypermetabolic metastatic lesion, were calculated for each patient. TB-MTV and TB-TLG were significantly associated with overall survival (OS) unlike SULmax ( $P=0.60$ ) and SULpeak ( $P=0.056$ ). The median OS for patients with

TB-MTV above the median was 10.84 months (95% CI, 5.88–15.81 months) as compared to 26.09 months (95% CI, 3.02–49.15 months) for patients with a TB-MTV below the median ( $P=0.002$ , HR=1.9). The corresponding 1-, 2- and 3-year survival rates were 47% vs. 72%, 28% vs. 51%, and 21% vs. 46% in these two subgroups, respectively. The prognostic value of TB-TLG and TB-MTV was equivalent. The median OS for patients with TB-TLG above the median was 10.84 months (95%CI: 5.50–16.19 months) as compared to 22.34 months (95%CI: 9.36–35.32 months) for patients with TB-TLG below the median. Differences in survival between patients with a sum of SULpeak above and below the median were not significant, but a trend was observed ( $P=0.056$ , HR=1.48). Image analysis was performed by a single experienced physician, certified in radiology and nuclear medicine, who had access to the patient's clinical data. Nevertheless, image interpretations were confirmed by another physician specialized in nuclear medicine, and doubtful results were resolved by consensus between the two investigators. Although this approach is not ideal, it does enable a certain quality of analysis to be maintained [48]. In this study, several FDG-PET/CT from the same manufacturer were used. A cross-calibration between the dose calibrator and the FDG-PET/CT was performed every month, in order to homogenize the results obtained, which is a guarantee of quality. This study had a selection bias because the population, although homogeneous and including brain metastasis were admitted, is not representative of the population eligible for immunotherapy, and the patients had previously received chemotherapy, which may bias the results under immunotherapy. Finally, there was no validation on an external cohort, which would provide greater robustness to the results [49].

In a retrospective single-center study, the analysis of volumetric parameters on initial PET prior to ICI induction was investigated in 56 patients with unresectable cutaneous ( $n=32$ ) or mucosal ( $n=24$ ) melanomas. Parameters extracted from tumor tissue were SUVmax, SUVmean, TB-MTV and from lymphoid tissue were bone marrow-to-liver ratio (BLR) and spleen-to-liver ratio (SLR). For the 2 matched melanoma subtypes (mucosal melanoma; cutaneous melanoma), (ORRs) were 33% and 42%, disease control rates (DCRs) were 56% and 69%, median follow-up were 25.0 and 28.9 months, median PFS were 4.7 and 10.7 months, and median OS were 23.9 and 28.3 months, respectively. For the CM subtype, increasing TB-MTV above 16.3 ml and BLR above 0.76 were independently associated with a shorter OS (both  $P<0.001$ ) and a shorter PFS ( $P=0.01$  and  $P=0.04$ , respectively) [50]. In this study, a single GE Discovery 690 FDG-PET/CT was used to ensure homogeneity of results and quality of analysis and findings. A pair of expert radiologists and nuclear medicine physicians analyzed the

FDG PET/CT scans and the contrast-enhanced CT scans and segmented the lesions. The readers were blinded and, unlike in the study conducted by Ito et al., had no knowledge of the patients' medical or pathological history, imaging characteristics or clinical outcome, [48]. This approach is of high methodological quality. However, it has a number of limitations: firstly, it is a heterogeneous population consisting of two melanoma subtypes with different prognosis and response to ICI [51, 52]. Secondly, the population is not homogeneous in terms of the treatments received, since in addition to the difference in the regimens administered in the first line, either anti-CTLA-4 or anti-PD1, there is also the fact of analyzing patients who received ICI in the second line. These elements help to explain the results, in addition to the small number of patients, in favor of a weak robustness, as the authors report. Finally, there was no validation in an external cohort.

In a meta-analysis of 24 studies analyzing the prognostic value of FDG-PET/CT in patients with melanoma under ICI, 13 studies used baseline scans. First-order radiomic variables, e.g. SUV, SUL, MTV, TLG, were the most commonly used. The pooled hazard ratios (HR) of MTV, SLR, SUV/SULmax, SUV/SULpeak, and TLG for OS were 1.777 (95% CI: 1.389–2.275,  $P<0.001$ ), 3.425 (95% CI: 1.707–6.869,  $P=0.001$ ), 0.941 (95% CI: 0.599–1.477,  $P=0.791$ ), 1.704 (95% CI: 1.253–2.316,  $P=0.016$ ), and 1.755 (95% CI: 1.315–2.342,  $P<0.001$ ), respectively. The conventional and modified response assessment criteria had a pooled sensitivity of 64% (95% CI: 46–79%) and 94% (95% CI: 81–99%) and a pooled specificity of 80% (95% CI: 59–93%) and 84% (95% CI: 64–95%), respectively [53].

Furthermore, in a retrospective series of 92 patients with advanced melanoma including 27% with BRAF mutation treated with anti-PD1 as first-line therapy, Nakamoto et al. also explored the association between FDG uptake by immune-mediating tissues, such as the bone marrow (BM) and spleen, and poor cancer outcomes. FDG/PET-CT scans were performed on different machines, which could lead to variability in SUV measurements of MTV. However, BM and spleen metabolism estimates were corrected for liver background as ratio (BLR and SLR, respectively), which harmonized the PET characteristics. MTV and TLG were calculated using an absolute threshold of SUV=2.5. Liver and spleen SUVmean were measured by drawing a spherical VOI in the center of a non-disease area of the right hepatic lobe (3 cm) and spleen (2 cm), respectively. For the BM, 1.5-cm spherical VOIs were placed in the center of a non-diseased L1–L4 vertebral bodies, and an average SUVmean was calculated for the lumbar vertebral bodies. After a median follow-up of 18.2 months, 57.6% (53/92) patients had disease progression, and 34.8% (32/90) of them died.

Median PFS and OS were 11.6 months (95% CI, 7.1–28.3 months) and more than 60 months, respectively. Both BLR and BRAF mutation were independent prognostic factors for PFS ( $P=0.017$  and  $0.018$ , respectively); and BLR, BRAF mutation, and lactate dehydrogenase (LDH) elevation were independent prognostic factors for OS ( $P=0.011$ ,  $0.0078$ , and  $0.013$ , respectively). The median PFS of the high BLR ( $>0.78$ ) group was 8.6 months (95% CI: 3.0–42.5 months), significantly shorter than that of the low BLR group at 28.3 months (95% CI: 7.7–54.9 months,  $P=0.027$ ). Similarly, the median OS of the high BLR group was 28.0 months (95% CI: 17.2–28.7 months), significantly shorter than that of the low BLR group  $>60$  months ( $P=0.019$ ) [54]. In this study, several FDG-PET/CT scanners from the same manufacturer were used as in the ITO study [48]. The use of different PET/CT scanners may have led to variability in SUV MTV measurements. However, the estimation of BM metabolism was assessed by normalizing the values to the liver background, thus harmonizing the PET characteristics and generalizing the model studied. Nevertheless, the authors did not specify the analysis modalities, which may have an impact on the quality of the study, unlike the previously cited studies [48, 50]. On the other hand, this was a generally homogeneous population, as brain metastases and BRAF-mutated patients were included, and all patients received first-line anti-PD1 therapy, although almost 12% received an anti-PD1-based combination. The retrospective nature of the study also induces selection bias. also note the absence of external validation, as in previous reported studies [48, 50].

In another retrospective study, two experienced nuclear medicine specialists performed a visual and semi-quantitative assessment of all lesions in 51 patients with CM who underwent FDG-PET/CT. Data were collected on the number of lesions in total and according to 5 body regions (head and neck, chest, abdomen and pelvis, skeleton and extremities, with bone lesions always counted in the skeleton). Usual metabolic parameters were collected (SULmax, SULpeak, SULmean, TB-MTV, TB-TLG) and tumor-to-background ratios (TBRs) calculated using mean activity in the liver (TLR), blood-pool (TBR), and muscle (TMR) at baseline, 3 months and 6 months after treatment initiation. The OS rates at 3 and 5 years were 49% and 43.1%, respectively. From baseline FDG-PET/CT, only the SULmax and SULpeak, as well as most TBRs, were predictive of 3- and 5-year OS rates. From 3 to 6 months FDG-PET/CT, TB-MTV, TB-TLG and most TBRs were predictive of survival. Changes in MTV, TLG, and most TBRs (delta values) between baseline and the 3- and 6-month follow-up scans had prognostic value. In the multivariate analysis, of all the parameters analyzed, the most predictive of OS were extracted from the 3-month follow-up scan. TBR was

the best reported predictive factor (cut-off value of 2.15, sensitivity of 88.5%, and specificity of 68.0% for 3-year survival) [55]. Two different FDG-PET/CT systems were used in this study. All studies were reanalyzed using Philips IntelliSpace Portal software, version 10.1. The authors also point out that their center is accredited by the European Association of Nuclear Medicine Forschungs GmbH (EARL) and that all acquisitions were performed according to EARL criteria (in order to maintain compatibility between data obtained from the two scanners), a quality guarantee not found in the previous reports cited. The authors point out that, despite the use of EARL standards to obtain the most accurate quantification possible, certain biases remain unavoidable. The reproducibility of TLG and MTV measurements had not been systematically studied in all patients. To eliminate the subjectivity associated with selecting a specific number of lesions to obtain reproducible quantitative results, the authors measured the sum of PET volumetric metabolic parameters for all lesions, specified as wMTV and wTLG. As in the study by Ito and Seban [50, 56], all images were analyzed by two experienced nuclear medicine specialists, who made a qualitative (visual) and semi-quantitative assessment of all lesions. The population is not sufficiently described to assess its homogeneity, in fact, there is no precision regarding mutation status or the presence of brain metastases. Nevertheless, all patients had received ICI as first-line treatment, but with different regimens. The majority of patients (62.7%) received anti-CTLA4, followed by anti-PD1 (31.4%) and a combination of anti-PD1 and anti-CTLA4 (5.9%). This treatment heterogeneity between anti-PD1 and CTLA4 may have influenced the results reported by the authors. In addition to the retrospective nature and relatively small sample size, which are sources of selection bias. It is important to note that there was no assessment of response rate and a lack of external validation.

Annovazzi et al. reported on a retrospective study of 57 patients with advanced BRAF-mutated melanoma treated with BRAF±MEK inhibitors as first-line therapy. The authors analyzed the prognostic value of pre-therapeutic metabolic parameters and investigated the association between FDG uptake by immune-mediating tissues, such as BM and spleen, and poor cancer outcome. Thirty-four patients were classified as responders (metabolic partial/complete response [PR/CR]) and 23 as non-responders (metabolic progressive/stable disease [PD/SD]). Baseline MTV  $>56$  ml was independently associated with a shorter PFS. Patients who achieved a CR were associated with longer PFS compared to those who achieved a PR (median of 42.9 vs. 8.8 months respectively,  $P=0.009$ ). Disease progression occurred in new disease sites in 87.5% of CR, 7.1% of PR and 34.8% of PD/SD ( $P<0.001$ ). High baseline TB-MTV and lack of treatment response

were independent prognostic factors for OS, stratifying patients into 3 different prognostic classes (median OS 6.7, 18.3 and 102.2 months, respectively) [57]. Only one FDG-PET/CT scanner was used in this study which guarantees the quality for the analysis of the results. The images were analyzed by two nuclear medicine physicians, who were familiar with the clinical data but had no knowledge of the clinical outcome of the patients. The evaluation criteria were the same as for the Ito et al. study [48]. Although the population was relatively small, the study focused exclusively on BRAF-mutated patients treated with the BRAFi+/- MEKi combination, which is unprecedented in the literature. Although the population was drawn from a retrospective collection from a single center, it remains globally homogeneous, with cutaneous melanoma including patients with brain metastases. The patients included had received different treatment regimens including BRAFi as monotherapy (vemurafenib or dabrafenib) or in combination with MEKi (cobimetinib or trametinib, respectively). The retrospective design of the study may have led to selection bias. Another possible limitation of the study is that 17.5% of patients had received systemic therapy prior to BRAFi/MEKi targeted therapy.

In a recent single-center retrospective Australian study of 106 patients with advanced melanoma, who received the combination of ipilimumab and nivolumab, the authors also assessed the prognostic value of first-order parameters from pre-therapy FDG-PET/CT. For each patient, TB-MTV was automatically measured on FDG-PET/CT in the entire field of view (vertex to mid-thigh) using the PERCIST liver SUV threshold ( $1.5 \times$  mean liver SUV in a 3 cm spherical VOI on the right lobe+2 SD) [58]. The median TB-MTV and TB-TLG were estimated to be 43 mL and 359 g, respectively. The 12-month survival rate was significantly lower in patients with TB-MTV and TB-TLG higher than the median values (43% [95% CI: 32, 58] vs. 66% [95% CI: 55, 79],  $P < 0.001$  and 43% [95% CI: 32, 57] vs. 67% (95% CI: 56, 80),  $P < 0.001$ , respectively). Similarly, the subgroups of patients with high TB-MTV or TB-TLG had a higher risk of death, with HRs of 1.9 (95% CI: 1.2, 3.0;  $P = 0.007$ ) and 2.0 (95% CI: 1.2, 3.1;  $P = 0.004$ ), respectively [59]. Iravani et al. had used several FDG PET/CT scanners which were cross-calibrated every 3 months, in order to homogenize the results obtained, which is a guarantee of quality. The analysis procedure is qualitative, even though it relied on a single experienced nuclear medicine physician who was blinded to clinical information and patient outcomes. Although the population was large, it was not homogeneous. The authors did not specify the exclusive cutaneous nature of the melanoma, and even if all patients had received a dose of ipilimumab and nivolumab, only 46.2% had received it as first-line treatment, while the other

patients had received it as subsequent-line treatment. The lack of external validation reduces the robustness of the results obtained.

Finally, in a retrospective two-center study of 56 patients with metastatic BRAF wild-type melanoma who underwent initial FDG-PET/CT prior to ICI, the primary objective was to predict prognosis based on radiomic parameters. A semi-automated lesion segmentation method was used, based on a relative SUVmax threshold of 40%. Texture analysis was performed in the VOI of the lesion with the highest FDG uptake. Before calculating the TFs, the voxel intensities of these VOIs were resampled using 64 discrete values ranging from 0 to 32 SUV units. Only one TF, Long Zone Emphasis (LZE), was reported to be significantly correlated with OS (LZE cut-off value = -437) in their series. And both total MTV and LZE were independent prognostic factors for OS. Therefore, the population was stratified into 3 risk categories: low risk if TB-MTV  $\leq 5.6$  ml and LZE  $\leq -437$  ( $n = 23$ ; 41%), intermediate risk if TB-MTV  $> 5.6$  ml or LZE  $> -437$  ( $n = 19$ , 34%), and high risk if TB-MTV  $> 5.6$  ml and LZE  $> -437$  ( $n = 14$ ; 25%). Survival rates were respectively 91.1% (95% CI: 80–100) for the low-risk group, 56.1% (95% CI: 37.1–85) for the intermediate-risk group, and 19% (95% CI: 0.06–60.2) for the high-risk group. The HRs for OS were 0.11 for the low-risk group (95% CI: 0.025–0.46,  $P = 0.0028$ ), 1.2 for the intermediate-risk group (95% CI: 0.48–2.8,  $P = 0.74$ ) and 5.9 for the high-risk group (95% CI: 2.5–14,  $P < 0.0001$ ) [60] (Table 2). The PET/CT scans were acquired with different systems. As this was a retrospective study, no harmonization was performed prior to data collection. However, the ComBat post-reconstruction harmonization method was used to combine the conventional and textural characteristics of the four different PET/CT scanners, 22,23. After ComBat harmonization, a selection of features was used to shrink the model to reduce overfitting and co-variable correlation. The bidirectional correlation of features was assessed using Spearman's rank correlation coefficient, and those with correlation coefficients greater than 0.8 were removed. The use of the ComBat procedure adds considerable robustness to the study methodology [61].

An experienced nuclear medicine physician analyzed and segmented the FDG-PET/CT scans without knowledge of the patients' clinical findings, as in the study by Iravani et al. [59]. This is a bicentric study, which is not the case with the previously cited studies, giving it a certain strength, but it is a retrospective study on a relatively small number of patients, which introduces a selection bias. Nevertheless, the population remains homogeneous, as it is a wild-type BRAF population that has only received anti-PD1 as first-line treatment exclusively. Furthermore, this is the only report to have evaluated TFs

**Table 2** Summary of studies evaluating the prognostic value of pre-therapy FDG-PET/CT in metastatic melanoma

Author Number of patients	PET Time point	Therapy	Features assessed	Outcomes	Findings	PET/CT Scanners	External validation	Harmonization
Ito et al [48] <i>n</i> = 142	Pre-therapeutic	Ipilimumab (second line)	MTV, TLG, $SUL_{max}$ , $SUL_{peak}$	Survival prognosis	TB-MTV and TB-TLG predicted survival	GE Discovery Series (VCT, ST, STE, 600 and 690),	No	A cross-calibration between the dose calibrator and the FDG-PET/CT was carried out every month
Seban et al [50] <i>n</i> = 56	Pre-therapeutic	Ipilimumab Pembrolizumab	$SUV_{max}$ , $SUV_{mean}$ , TB-MTV, BLR, SLR	Survival prognosis	TB-MTV > 16.3 ml and BLR > 0.76 was independently associated with a shorter OS (both $P < 0.001$ ) and a shorter PFS ( $P = 0.01$ and $P = 0.04$ , respectively)	GE Discovery 690 FDG-PET/CT	No	NA
Nakamoto et al [54] <i>n</i> = 92	Pre-therapeutic	Pembrolizumab Nivolumab	$SUV_{max}$ , MTV, BLR	Survival prognosis	BLR > 0.78 was independently associated with a shorter PFS and OS ( $P = 0.017$ and $P = 0.011$ , respectively)	GE Discovery 600, 690, 710 or MI (GE Healthcare),	No	Normalization to hepatic background
Schweighofer-Zwink et al [55] <i>n</i> = 51	Pre-therapeutic	Ipilimumab Pembrolizumab Nivolumab	$SUL_{max}$ , $SUL_{peak}$ , $SUL_{mean}$ , TB-MTV, TB-TLG, TBR, TMR	Survival prognosis	TBR was independently associated with survival (cutoff value of 2.15, sensitivity of 88.5%, and specificity of 68.0% for 3-year survival)	Phillips TF Ingenuity and Siemens Biograph mCT scanners	No	Acquisitions were carried out on the basis of EARL criteria (in order to maintain compatibility between data obtained from the two scanners), Use of Philips IntelliSpace Portal software, version 10.1.
Annovazzi et al [57] <i>n</i> = 57	Pre-therapeutic	BRAF <sup>i</sup> MEK <sup>i</sup>	TB-MTV	Survival prognosis	TB-MTV > 56 ml was independently associated with a shorter PFS and OS	Siemens Biograph16 (Siemens Healthineers),	No	NA

**Table 2** (continued)

Author Num- ber of patients	PET Time point	Therapy	Features assessed	Outcomes	Findings	PET/CT Scanners	Externe validation	Harmonization
Iravani et al [59] n = 106	Pre-therapeutic	Ipilimumab- Nivolumab combination	TB-MTV TB-TLG	Survival prognosis	TB- MTV > 42 ml was inde- pendently associated with a shorter PFS and OS	FDG PET/CT scan- ners, including Biograph 16(Sie- mens Health- ineers), GE 690 or GE 710 Discovery (GE Healthcare) scanners	No	Cross-calibra- tion every 3 months
Flaus et al. [60] n = 56	Pre-therapeutic	Pembrolizumab Nivolumab	41 features including: First order statistics, GLCM, GLRLM, GLSZM, NGLDM	Survival prognosis Thera- peutic response	Both TB- MTV > 5,6 ml and LZE >-437were indep- endently associ- ated with a shorter OS ( $P < 0.0001$ )	Biograph mCT Flow 20 (Siemens Healthcare) and Biograph 6 HI-REZ (Siemens Healthcare) Biograph Horizon 16 (Siemens Healthcare) and Discovery 690 (General Electrics Healthcare)	No	ComBat procedure

**Abbreviations** GLCM=gray level co-occurrence matrices; GLRLM=gray-level run length matrix; GLSZM=gray level size zone matrix; NGLDM=neighborhood gray-level different matrix; LZE=long zone emphasis; SUV=standardized uptake value; SUL=SUV corrected by lean body mass; TB=total body; MTV=metabolic tumor volume; TLG=total lesion glycolysis; ACC=absolute correlation coefficient; BLR=bone marrow-to-liver ratio; SLR=spleen-to-liver ratio; TBR=tumor-to-background ratio; TMR=tumor-to-muscle ratio; OS=overall survival; PFS=progression free survival

Finally, as with all the studies cited in this chapter, there was no validation on an external cohort.

### Predictive mutational status approach

Survival of patients with advanced CM is still limited, reaching barely 25–30% at 5 years [62]. However, this survival could be significantly improved by targeted therapies, which offer 3-year survival rates of 37%, depending mainly on genomic status [63], with BRAF- (50%), NRAS- (15–20%) and proto-oncogene (c-Kit) (2–3%) mutations being the most common [64]. However, these individualized therapies rely on genomic analysis of tumor cells, which requires prior biopsy and histopathological evaluation. Any imaging technique that provides information on mutational status could be of considerable benefit, avoiding time-consuming histopathological analysis and allowing for earlier initiation of targeted therapy, thereby optimizing outcomes. Few series have been reported on this subject in literature.

Saadani et al. [65] conducted the first study on a cohort of 70 patients with unresectable stage III-IV melanoma to predict BRAFV600 mutation status using selected radiomic TFs derived from FDG PET/CT with molecular biology as the gold standard. Patients were divided into 2 groups according to their mutation status: BRAFV600 group (35 patients, 100 lesions) and wild-type BRAF group (35 patients, 79 lesions). The size of the tumor lesion was measured in the axial plane. Measurable

disease was defined as lesions of at least 2 cm or (if the tumor was indistinguishable on CT) by an equivalent MTV of at least 4.2 ml. from each patient's measurable lesions, the 3 lesions with the highest SUVmax per organ were considered target lesions. Target lesions were delineated by a relative threshold of SUVmax > 50% without background correction. The specified organ regions were consistent with the melanoma metastasis model: lymph nodes, lung, liver, bone, subcutaneous, intramuscular, and others. For radiomic analysis, 485 features were extracted, including 5 conventional PET parameters (SUVmax, SUVmean, SUVpeak, MTV and TLG) and others related to morphology ( $n=22$ ), local intensity ( $n=2$ ), intensity-based statistics ( $n=18$ ), intensity-volume histogram ( $n=6$ ), intensity histogram ( $n=24$ ) and texture ( $n=408$ ). TFs were based on GLCM, GLRLM, GLSZM, gray level distance zone matrix (GLDZM), the neighborhood gray tone difference matrix (NGTDM) and NGLDM. For feature selection, the authors considered 6 different approaches, which they cross-validated 10 times. For each patient, 1 to 10 target lesions were analyzed. The BRAFV600 and wild-type BRAF groups were not statistically different with respect to SUV, MTV, TLG, longest diameter, or prior local intervention metrics. Stratification by organ region for SUV and TLG metrics yielded the same result. The best prediction model based on conventional PET features included all of them, i.e., SUVmean, SUVmax, SUVpeak, TLG, and MTV. Texture

analysis was performed on 176 lesions (3 lesions from scans with a different voxel matrix were excluded). AUCs predictive of BRAFV600 mutation ranged from 0.54 to 0.62. The authors conclude that BRAFV600 mutation status was neither associated with nor predicted by conventional PET features, whereas TFs had predictive value (AUC of 0.54 to 0.62). Patients were scanned on a Phillips Gemini TF 16 or Phillips Gemini TF big-bore FDG-PET/CT scanner, cross-calibrated (with calibration phantoms) for 1–3 min per bed position. The systems were from the same supplier and had the same type of image acquisition and reconstruction methods, and the same settings were used. This guarantees the quality of the image analysis. However, the methods used to analyze the FDG-PET/CT images were not specified, which is a methodological shortcoming. Regarding the population, it was derived from a monocentric retrospective collection, which is synonymous with selection bias. However, even if the number of patients is small, this population remains homogeneous. Finally, in the absence of an external validation cohort, the authors carried out 10 cross-validations, which could lead to highly biased results if the TFs are poorly applied [66].

Furthermore, Olthof et al. [67] evaluated the potential of CT texture analysis parameters and metabolic characteristics of melanoma metastases on FDG-PET/CT to predict relevant mutations of tumor cells prior to targeted therapy in correlation with histopathological specimens. Sixty-six patients undergoing contrast-enhanced FDG-PET/CT before planned metastasectomy without any prior systemic therapy were included in this single-centre retrospective analysis. Both quantitative analysis on FDG-PET (including MTV, SUVmax, SUVpeak and SUVmean) and CT analysis (including attenuation Standard Deviation (SD), Kurtosis, Skewness, Mean value of Positive Pixels (MPP), Uniformity of Distribution of Positive Pixels (UPP), Entropy, Uniformity) were calculated in the largest resected metastasis in a delineated area with a relative threshold of  $SUV_{max} < 40\%$ . Tissue samples were analysed for BRAF ( $n=21$ ), NRAS ( $n=11$ ), c-Kit ( $n=1$ ) and wild-type ( $n=23$ ). In 10 patients, including 3 patients with uveal melanoma, the mutation status was unknown. The attenuation SD within the target lesion showed a weak correlation with its SUVpeak ( $\rho = -0.292$ ,  $P=0.017$ ). However, no correlation between metabolic FDG-PET parameters and tumor cell mutation was found [68]. All patients were examined on the same whole-body FDG-PET/CT system (Biograph mCT, Siemens Healthineers, Knoxville TN, USA). This ensures the quality of image analysis. Patients were evaluated retrospectively by two readers blinded to any clinical data using dedicated software. With regard to the population, it was derived from a monocentric retrospective collection, which is synonymous with selection bias.

### Methodological aspect

The methodological aspect of radiomic studies, including texture analysis, remains essential to ensure reproducibility of results in clinical practice [38]. Furthermore, intratumoral heterogeneity in the case of multimetastatic disease is one of the problems of melanoma. Because of this characteristic, manual tumor delineation is an extremely time-consuming process, which can make it difficult to use clinically in routine practice. Typically, target lesions are quantified and used for follow-up. However, more recent results suggest the use of total MTV as a single parameter or in combination with the number of new lesions to improve therapy assessment and outcome prediction [69]. Many studies have suggested that TFs are affected by region size [70, 71], which is crucial for the classification of highly heterogeneous regions of varying size. The first threshold of 64-voxel was chosen after most research papers used this volume threshold for regions to be included, as TFs below this size provide less predictive information. Therefore, the methodological aspect in radiomic studies remains essential to ensure the reproducibility of results for clinical practice [38]. Recently, a standardization of radiomic procedures has been proposed to avoid the lack of reproducibility and to allow harmonization of practice, [72, 73]. In the future, radiomics will also undoubtedly include artificial intelligence to increase its robustness, [39].

The studies mentioned above in this review [60, 65, 67, 74] used different thresholding techniques for image segmentation. The most commonly used absolute threshold of  $SUV=2.5$  produces many false positive regions of interest, initially overestimating tumor regions and making manual correction very laborious and time consuming. Nevertheless, a relative threshold of  $40\%SUV_{max}$  has been reported as a more reproducible method of tumor delineation and the best approximation for estimating tumor volume [75]. Indeed, Guezennec et al. demonstrated high interobserver reproducibility with this method for calculating TFs in an analysis of 43 head and neck cancers, showing excellent agreement for 3 indices (GLCM Homogeneity, GLCM Correlation, and GLCM Entropy) with an intraclass correlation coefficient greater than 0.90 [76]. In practice, thresholding methods offer the following advantages: simple, fast calculation, good results on images with objects having uniform intensity values against a high-contrast background. They also work well on images with variable backgrounds and very small, sparse objects. However, these methods have a number of disadvantages, as they give poor results on images with low object/background contrast, generate more noise, a background intensity with significant variation across the image/high illumination gradient, and are quite computationally expensive [77]. Another commonly used segmentation technique is based on the

PERCIST approach for treatment response assessment in delineating the target tumor with a threshold of  $1.5 \times$  mean liver SUV in a 3-cm spherical VOI in the right lobe + 2 SD [58, 59]. This technique has the same limitations as absolute SUV-based thresholding, namely the difficulty in dealing with intratumoral heterogeneity. In addition, the gradient-based method allows assessment of tumor in easily excluding areas of necrosis, but is not widely available, although it is highly reproducible [76, 78].

Another pitfall is the size of the tumors. The disadvantage of calculating TFs for regions larger than 64 voxels [79] is not really relevant for a disease like advanced CM, which spreads throughout the body with heterogeneous regions of different sizes, even smaller than a few milliliters. Nevertheless, higher order features require more space to manifest their properties. Thus, some studies suggest that these second or higher order TFs are highly dependent on volume size [70, 71, 80], as gray-scale regions are less present in small volumes. However, it has previously been shown that quantification at 64 Gy levels provides the best compromise between sufficient sampling of voxel SUVs, preservation of original intensity data and potential additional information regarding MTV [80]. This suggests that for spherical lesions, a minimum volume of at least  $4 \text{ cm}^3$  using a voxel size of 4 mm ( $4 \times 4 \times 4 = 64$  voxels) is currently the retained size threshold [81, 82]. The effects of SUV discretization and spatial resampling were investigated by discretizing into a fixed number of bins (32, 64 or 128), over the full SUV range in each image, and either without spatial resampling ( $4.07 \times 4.07 \times 5.00$  mm), or by resampling into 3 mm isotropic voxels. Over 60% (60.3%) of the concordance correlation coefficient (CCC) values for TFs varied by less than 5% between bin number groups. The repeatability of GLCM features was largely insensitive to changes in SUV discretization, with 80.2% of CCC values varying by less than 5% within SUV bin groups. GLZLM features were significantly affected by changes in SUV bin number, with 69.3% of CCC values varying by 5% or more (and up to 228.6%) within SUV bin groups [83]. To study the impact of harmonized image reconstructions on feature consistency, the authors performed discretization using a fixed bin number (FBN) of 64 and a fixed bin width (FBW) of 0.25. They found that the effect of image discretization resulted in better repeatability and less sensitivity to boundary differences for the FBW discretization [84]. In addition, Orhac et al. showed that discretization with FBW resulted in more meaningful features, i.e. features capable of distinguishing tumor types well [85].

To address these specific problems in metastatic melanoma, Vagenas et al. [86] proposed a whole-body segmentation model capable of efficiently identifying highly heterogeneous tumor lesions in metastatic melanoma

from whole-body 3D FDG-PET/CT images. The decision support system provides unsupervised segmentation of high FDG uptake regions based on fuzzy C-means and a custom region growth algorithm to overcome size constraints. Subsequently, a region classification model based on radiomic features and neural networks was used to classify these regions as tumors or not. The experimental results showed high performance in identifying metastatic melanoma lesions with a sensitivity of 83.7%, specificity of 91.8% and accuracy of 87.75% with an AUC of 94.2%.

Another problem with radiomics is the difference in significant TFs from study to study, probably due to the “rebound beta”, which would result from a high correlation between the plethora of TFs often analyzed. To minimize this problem, Deleu et al. [87] used the principal component analysis (PCA) technique on 123 melanoma lesions from 26 patients. This technique, which reduces the dimensionality of large datasets containing highly correlated variables, such as texture feature datasets derived from FDG-PET images, improves the interpretability of the data while minimizing the loss of information by creating new uncorrelated variables that successively maximize variance. Thirty-eight TFs were extracted from different VOIs, including GLCM, NGLDM, GLRLM and GLSZM. Using Principal Component Analysis, their dataset of 38 generated TFs could be compressed to a dataset of 5 new uncorrelated variables or PCA that explained approximately 82% of the total variance. Although this study was conducted on a small population and retrospectively, it allowed the identification of 5 new uncorrelated variables that provide summarized information on the locoregional FDG distribution with an emphasis on both high and low FDG uptake regions, contrast in FDG uptake values (steepness), tumor volume, locoregional tracer distribution and on the speed of change of SUV intensity between different regions.

The analysis of radiomic features has shown particular promise in cancer research. However, traditional radiomic feature analysis has had limited utility for patients with metastatic or multifocal disease because there are few established methods for aggregating radiomic features from multiple tumors to produce patient-level outcome estimates. Chang et al. [88] compared different radiomic feature aggregation methods and tested their performance in different survival models in a population of patients with secondary brain metastases from various solid tumors (non-small cell lung cancer (NSCLC), small cell lung cancer (SCLC), breast, melanoma, renal, gastrointestinal). The following aggregation methods were compared to estimate patient-level risk for patients with multiple metastases: (1) Unweighted average: radiomic size and shape characteristics were

summed, and all other characteristics were averaged for each patient. (2) Weighted average: radiomic size and shape characteristics were summed, while all other characteristics were averaged for each patient based on a weighted proportion of the total volume of all metastases per patient. (3) Weighted average of the three largest metastases: the radiomic characteristics of size and shape of the three largest metastases in each patient were summed, while all other characteristics of the three largest metastases were averaged based on a weighted proportion of the total volume of the three largest metastases. The three largest metastases were chosen because Graded Prognostic Assessment, a widely accepted prognostic model, uses a threshold of three metastases for prognostic risk stratification. (4) Largest metastasis+number of metastases: the characteristics of each patient's largest metastasis were selected, and the total number of metastases for each patient was included as an additional variable. (5) Largest metastasis alone. (6) Smallest metastasis alone as a control with the hypothesis that the smallest tumor would have a reduced prognostic value compared to larger tumors. The results show that a volume-weighted average of the radiomic characteristics of the three largest brain metastases is the most effective technique for modelling survival across different survival analysis methods. It also suggests that in patients with multifocal disease, the largest tumors may influence prognosis. The application of this aggregation method was used in this study, which evaluated whether the combination of radiomic and clinical parameters was superior to clinical parameters alone in predicting therapeutic response at 3 months and OS at 6 and 12 months, in 261 patients with stage IV malignant melanoma undergoing immunotherapy with PD-1 and CTLA-4 checkpoint inhibitors. Radiomic assessment was performed using CT. For each segmented lesion, 14 radiomic shape features, 18 first-order statistical features and 75 TFs were extracted. A total of 1316 features were extracted per lesion. The features were not harmonized to account for the different scanner types, each radiomic feature at the lesion level was aggregated per patient in four different ways: (1) feature value of the most important lesion; (2) volume-weighted average of the feature values of the three most important lesions; (3) feature value of the most predictive lesion, and (4) volume-weighted average of the feature values of the three most predictive lesions. Although the study was not statistically significant ( $P=0.08$ ) and the resulting AUCs were low ( $AUC<0.7$ ), it has the merit of following a whole-body segmentation approach of all visible metastases in order to obtain as much information as possible [89].

Finally, radiomics suffers from the problem of reproducibility. Indeed, model stability and reproducibility clearly need to be assessed before applying a predictive

model in a clinical context. Indeed, it is well known that model fit is optimal in the training set used to build the model, whereas validation in an external cohort provides more reliable fit estimates. The first step in model validation is internal cross-validation. However, the best way to assess the potential clinical value of a model is to validate it with independent, prospectively collected cohorts, ideally in clinical trials. This raises the problem of data sharing between different institutions, requiring to use of shared databases as validation sets [42, 49]. One approach to solving this problem is the development of large, publicly accessible databases, such as The Cancer Imaging Archive (TCIA) [90], a public resource that hosts patient imaging data from The Cancer Genome Atlas (TCGA) database, which contains comprehensive multidimensional genomic data and clinical annotations for over 30 cancer types [91]. These images can be used as a valuable resource to generate and validate hypotheses.

### Practice harmonization strategies

The field of radiomics is expanding and offers real hope for prediction in the field of cancer. However, studies are mainly carried out in monocentric settings, with the problem of reproducibility. Indeed, the variability of radiomic steps can affect the absolute values and statistical distributions of radiomic features, and thus the interchangeability of radiomic features [36, 92]. Radiomics is a rapidly expanding field of research [93], but to date the vast majority of studies have been conducted out in a single institution, creating a "batch effect", which is characterized by the disparity between the biological [40]. This "batch effect" is conceptually similar to variations induced in radiomic features induced by the medical imaging model, acquisition protocol and/or reconstruction settings, sometimes referred to as the "center effect".

Various techniques for eliminating "batch effects" have been widely used when combining data from different experiments consisting of two or more data sets. The "batch effect" removal methods can be classified into three groups: matrix factorization (MF) methods, discretization methods and location-scale (LS) methods [94].

Matrix factorization methods (MF) assume that the variation in the data corresponding to batch effects is independent of the variation corresponding to the biological variable of interest and can be captured in a small set of factors that can be estimated by matrix factorization methods.

Discretization methods transform expression data into persistently defined classes based on their expression levels. This merging method can be done trivially by concatenating discretized matrices. discretized matrices. Some loss of information during discretization is unavoidable. However, it has been shown that these strategies can

occasionally lead to comparable or improved accuracy, depending on the type of downstream analysis.

Location-Scale (LS) methods transform the data of each batch in order to obtain an equal mean and/or variance. Among these LS methods, there is the ComBat method, which is widely used in pathological imaging [95, 96]. ComBat identifies a lot-specific transformation to express all data in a common space without center effects. The ComBat algorithm follows a three-step procedure: (i) Data standardization, (ii) Empirical Bayes estimation of prior distribution hyperparameters from the standardized data and subsequent estimation of batch effect parameters, which are used in (iii) Batch correction [97]. It eliminates batch effects mainly on the basis of an empirical Bayes framework. It has been shown to be robust with smaller sample sizes and remains a widely used approach. When compared with five other popular “batch effect” removal methods, ComBat was found to be “most capable of reducing and eliminating batch effects while increasing precision and accuracy”. In the context of harmonizing radiomics characteristics, ComBat works by normalizing them. It has several interesting properties for radiomics harmonization as it does not require feature definitions to be modified, so it can be used with any algorithm. It’s fast and easy to use. It makes use of all available information, as none of the features are eliminated before analysis. It is machine-specific, based solely on patient data acquired in the different centers, and does not require phantom experience and it can be used for prospective or retrospective data, provided that the different centers have the same cohorts of patients with the same disease who have received the corresponding treatment. One of the limitations is that it centers the data on the overall mean of all samples, resulting in an adjusted data matrix that is shifted to an arbitrary location that no longer coincides with the location of any of the original centers. This can cause harmonized features to lose their original physical meaning (and impossible values, such as negative volumes or SUVs) [61].

A modified version of ComBat, called M-ComBat [98], centers the data on the location and scale of a predetermined “reference” batch, providing the ability to harmonize the feature set with a chosen reference. Da-Ano proposed a parametric bootstrap addition for the parameters in the ComBat and M-ComBat models, respectively [61]. The proposed use of bootstrapping for initial estimates reduced the variances within each center and facilitated bias reduction in the estimation of center effect parameters by ComBat and M-ComBat, respectively, resulting in improved center effect elimination and predictive performance. Although relatively small, the improvement from the use of Da-Ano M-ComBat was consistent.

In total, these studies were all retrospective and essentially monocentric, with the exception of the Flaus study, which involved two centers. This makes them less robust.

The overall mean number of patients in the studies reviewed was 77, with 4 studies having numbers of around 50 patients, which is relatively small for a radiomics study design. The studies [50, 54–56, 59, 60] studies used different machines with harmonization procedures to ensure the quality of analysis. Flaus et al [60], on the other hand, used the ComBat technique, which serves as a reference technique to overcome the biases introduced by different machines.

Only two studies considered TFs [60, 65].

It should also be noted that no external validation was performed. However, a cross-validation was carried out by Saadani’s team [65], but this was less robust than an external validation.

## Conclusion

Radiomics holds promise for assessing melanoma and predicting response to TKIs and ICI immunotherapy, and also has the potential to predict mutational status. but certain limitations need to be overcome. Firstly, study design must be based on a methodology with specific criteria, which would harmonize practices and promote reproducibility. Secondly, it would be essential to design prospective studies with a large number of patients, in order to overcome the heterogeneity generated by retrospective studies. This would make it possible to include in clinical practice a simple tool capable of correctly analyzing images and aiding therapeutic decision-making.

## Acknowledgements

Not applicable.

## Author contributions

Conceptualization, K.A, D.D, V.B and R.A.; methodology, K.A, C.L.M, V.B and R.A.; software, K.A; validation, D.D and R.A. formal analysis, P.T and D.B; investigation, C.L.M. and D.B.; resources, R.A. and C.L.M.; data curation, K.A; writing—original draft preparation, K.A; writing—review and editing, R.A, D.D and V.B; visualization, A.U and C.B; supervision, R.A; project administration, D.D and R.A. All authors have read and agreed to the published version of the manuscript.

## Funding

This research received no external funding.

## Data availability

Not applicable.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

All authors declare that they have no conflict of interest in relation to the results of this review.

Received: 2 November 2023 / Accepted: 24 June 2024

Published online: 05 July 2024

## References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136:E359–386.
2. Matthews NH, Li W-Q, Qureshi AA, Weinstock MA, Cho E. Epidemiology of Melanoma. In: Ward WH, Farma JM, editors. *Cutan Melanoma Etiol Ther* [Internet]. Brisbane (AU): Codon Publications; 2017 [cited 2022 Feb 6]. <http://www.ncbi.nlm.nih.gov/books/NBK481862/>.
3. Buja A, Bardin A, Damiani G, Zorzi M, De Toni C, Fusinato R, et al. Prognosis for cutaneous melanoma by Clinical and Pathological Profile: a Population-based study. *Front Oncol*. 2021;11:737399.
4. Hartman RI, Lin JY. Cutaneous Melanoma-A review in detection, staging, and management. *Hematol Oncol Clin North Am*. 2019;33:25–38.
5. Keung EZ, Gershenwald JE. The eighth edition American Joint Committee on Cancer (AJCC) melanoma staging system: implications for melanoma treatment and care. *Expert Rev Anticancer Ther*. 2018;18:775–84.
6. Cherobin ACFP, Wainstein AJA, Colosimo EA, Goulart EMA, Bittencourt FV. Prognostic factors for metastasis in cutaneous melanoma. *Bras Dermatol*. 2018;93:19–26.
7. 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.
8. Robert C, Ribas A, Schachter J, Arance A, Grob J-J, Mortier L, et al. Pembrolizumab versus ipilimumab in advanced melanoma (KEYNOTE-006): post-hoc 5-year results from an open-label, multicentre, randomised, controlled, phase 3 study. *Lancet Oncol*. 2019;20:1239–51.
9. Robert C, Long GV, Brady B, Dutriaux C, Maio M, Mortier L, et al. Nivolumab in previously untreated melanoma without BRAF mutation. *N Engl J Med*. 2015;372:320–30.
10. Larkin J, Chiarion-Sileni V, Gonzalez R, Grob JJ, Cowey CL, Lao CD, et al. Combined Nivolumab and Ipilimumab or Monotherapy in untreated melanoma. *N Engl J Med*. 2015;373:23–34.
11. Tawbi HA, Schadendorf D, Lipson EJ, Ascierto PA, Matamala L, Castillo Gutiérrez E, et al. Relatlimab and Nivolumab versus Nivolumab in Untreated Advanced Melanoma. *N Engl J Med*. 2022;386:24–34.
12. Lebbé C, Weber JS, Maio M, Neyns B, Harmankaya K, Hamid O, et al. Survival follow-up and ipilimumab retreatment of patients with advanced melanoma who received ipilimumab in prior phase II studies. *Ann Oncol off J Eur Soc Med Oncol*. 2014;25:2277–84.
13. Robert C, Long GV, Brady B, Dutriaux C, Di Giacomo AM, Mortier L, et al. Five-year outcomes with nivolumab in patients with Wild-Type BRAF Advanced Melanoma. *J Clin Oncol off J Am Soc Clin Oncol*. 2020;38:3937–46.
14. LAG3-PD-1 combo impresses in Melanoma. *Cancer Discov*. 2021;11:1605–6.
15. Larkin J, Chiarion-Sileni V, Gonzalez R, Grob J-J, Rutkowski P, Lao CD, et al. Five-year survival with combined Nivolumab and Ipilimumab in Advanced Melanoma. *N Engl J Med*. 2019;381:1535–46.
16. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, et al. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Med*. 2011;364:2507–16.
17. Hauschild A, Grob J-J, Demidov LV, Jouary T, Gutzmer R, Millward M, et al. Dabrafenib in BRAF-mutated metastatic melanoma: a multicentre, open-label, phase 3 randomised controlled trial. *Lancet Lond Engl*. 2012;380:358–65.
18. Chapman PB, Robert C, Larkin J, Haanen JB, Ribas A, Hogg D, et al. Vemurafenib in patients with BRAFV600 mutation-positive metastatic melanoma: final overall survival results of the randomized BRIM-3 study. *Ann Oncol*. 2017;28:2581–7.
19. Hauschild A, Ascierto PA, Schadendorf D, Grob JJ, Ribas A, Kiecker F, et al. Long-term outcomes in patients with BRAF V600-mutant metastatic melanoma receiving dabrafenib monotherapy: analysis from phase 2 and 3 clinical trials. *Eur J Cancer Oxf Engl*. 1990. 2020;125:114–20.
20. Larkin J, Ascierto PA, Dréno B, Atkinson V, Liszkay G, Maio M, et al. Combined vemurafenib and cobimetinib in BRAF-mutated melanoma. *N Engl J Med*. 2014;371:1867–76.
21. Long GV, Stroyakovskiy D, Gogas H, Levchenko E, de Braud F, Larkin J, et al. Combined BRAF and MEK inhibition versus BRAF inhibition alone in melanoma. *N Engl J Med*. 2014;371:1877–88.
22. Ascierto PA, Dréno B, Larkin J, Ribas A, Liszkay G, Maio M, et al. 5-Year outcomes with Cobimetinib plus Vemurafenib in BRAFV600 mutation-positive Advanced Melanoma: Extended follow-up of the coBRIM study. *Clin Cancer Res off J Am Assoc Cancer Res*. 2021;27:5225–35.
23. Dummer R, Flaherty KT, Robert C, Arance A, de Groot JWB, Garbe C, et al. COLUMBUS 5-Year update: a randomized, Open-Label, phase III trial of Encorafenib Plus Binimetinib Versus Vemurafenib or Encorafenib in patients with BRAF V600-Mutant melanoma. *J Clin Oncol*. 2022;40:4178–88.
24. Robert C, Grob JJ, Stroyakovskiy D, Karaszewska B, Hauschild A, Levchenko E, et al. Five-year outcomes with Dabrafenib plus Trametinib in Metastatic Melanoma. *N Engl J Med*. 2019;381:626–36.
25. Wolchok JD, Chiarion-Sileni V, Gonzalez R, Grob J-J, Rutkowski P, Lao CD, et al. Long-term outcomes with Nivolumab Plus Ipilimumab or Nivolumab alone Versus Ipilimumab in patients with Advanced Melanoma. *J Clin Oncol off J Am Soc Clin Oncol*. 2022;40:127–37.
26. Ferrucci PF, Di Giacomo AM, Del Vecchio M, Atkinson V, Schmidt H, Schachter J, et al. KEYNOTE-022 part 3: a randomized, double-blind, phase 2 study of pembrolizumab, dabrafenib, and trametinib in BRAF-mutant melanoma. *J Immunother Cancer*. 2020;8:e001806.
27. Ascierto PA, Stroyakovskiy D, Gogas H, Robert C, Lewis K, Protsenko S, et al. Overall survival with first-line atezolizumab in combination with vemurafenib and cobimetinib in BRAFV600 mutation-positive advanced melanoma (IMspire150): second interim analysis of a multicentre, randomised, phase 3 study. *Lancet Oncol*. 2023;24:33–44.
28. Michielin O, Atkins MB, Koon HB, Dummer R, Ascierto PA. Evolving impact of long-term survival results on metastatic melanoma treatment. *J Immunother Cancer*. 2020;8:e000948.
29. Huang M, Lou Y, Pellissier J, Burke T, Liu FX, Xu R, et al. Cost effectiveness of Pembrolizumab vs. Standard-of-care chemotherapy as first-line treatment for metastatic NSCLC that expresses high levels of PD-L1 in the United States. *Pharmacoeconomics*. 2017;35:831–44.
30. Miguel LS, Lopes FV, Pinheiro B, Wang J, Xu R, Pellissier J, et al. Cost effectiveness of Pembrolizumab for Advanced Melanoma Treatment in Portugal. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2017;20:1065–73.
31. Courtney PT, Yip AT, Cherry DR, Salans MA, Kumar A, Murphy JD. Cost-effectiveness of Nivolumab-Ipilimumab Combination Therapy for the treatment of Advanced Non-small Cell Lung Cancer. *JAMA Netw Open*. 2021;4:e218787.
32. Salaün P-Y, Abgral R, Malard O, Querellou-Lefranc S, Quere G, Wartski M, et al. Good clinical practice recommendations for the use of PET/CT in oncology. *Eur J Nucl Med Mol Imaging*. 2020;47:28–50.
33. Hicks RJ, Iravani A, Sandhu S. 18F-fluorodeoxyglucose Positron Emission Tomography/Computed tomography for assessing Tumor response to Immunotherapy in Solid tumors: Melanoma and Beyond. *PET Clin*. 2020;15:11–22.
34. Bisschop C, de Heer EC, Brouwers AH, Hospers G, a. P, Jalving M. Rational use of 18F-FDG PET/CT in patients with advanced cutaneous melanoma: a systematic review. *Crit Rev Oncol Hematol*. 2020;153:103044.
35. Van de Wiele C, Juanito G, Vander BK, Lawal I, Sathekge M, Maes A, et al. Practical considerations when interpreting FDG PET/CT Imaging for Staging and Treatment Response Assessment in Melanoma patients. *Semin Nucl Med*. 2021;51:544–53.
36. Hatt M, Tixier F, Pierce L, Kinahan PE, Le Rest CC, Visvikis D. Characterization of PET/CT images using texture analysis: the past, the present... any future? *Eur J Nucl Med Mol Imaging*. 2017;44:151–65.
37. Cook GJR, Goh V. A role for FDG PET Radiomics in Personalized Medicine? *Semin Nucl Med*. 2020;50:532–40.
38. Reuzé S, Schernberg A, Orhac F, Sun R, Chargari C, Derclé L, et al. Radiomics in Nuclear Medicine Applied to Radiation Therapy: methods, pitfalls, and challenges. *Int J Radiat Oncol Biol Phys*. 2018;102:1117–42.
39. Hatt M, Le Cheze C, Antonorsi N, Tixier F, Tankyevych O, Jaouen Y, et al. Radiomics in PET/CT: current status and future AI-Based evolutions. *Semin Nucl Med*. 2021;51:126–33.
40. Da-Ano R, Visvikis D, Hatt M. Harmonization strategies for multicenter radiomics investigations. *Phys Med Biol*. 2020.
41. Yang L, Xu P, Li M, Wang M, Peng M, Zhang Y et al. PET/CT Radiomic Features: A Potential Biomarker for EGFR Mutation Status and Survival Outcome Prediction in NSCLC Patients Treated With TKIs. *Front Oncol* [Internet]. 2022 [cited 2023 Jun 4];12. <https://www.frontiersin.org/articles/https://doi.org/10.3389/fonc.2022.894323>.
42. Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp*. 2018;2:36.

43. Bundschuh RA, Dinges J, Neumann L, Seyfried M, Zsótér N, Papp L, et al. Textural Parameters of Tumor Heterogeneity in <sup>18</sup>F-FDG PET/CT for Therapy Response Assessment and Prognosis in patients with locally advanced rectal Cancer. *J Nucl Med off Publ Soc Nucl Med*. 2014;55:891–7.
44. Lovinfosse P, Polus M, Van Daele D, Martinive P, Daenen F, Hatt M, et al. FDG PET/CT radiomics for predicting the outcome of locally advanced rectal cancer. *Eur J Nucl Med Mol Imaging*. 2018;45:365–75.
45. Lee S-T, Kovaleva N, Senko C, Kee D, Scott AM. Positron Emission Tomography/Computed Tomography Transformation of Oncology: Melanoma and skin malignancies. *PET Clin*. 2024;19:231–48.
46. Abgral R, Bourhis D, Salaun P-Y. Clinical perspectives for the use of total body PET/CT. *Eur J Nucl Med Mol Imaging*. 2021;48:1712–8.
47. Sah B-R, Owczarczyk K, Siddique M, Cook GJR, Goh V. Radiomics in esophageal and gastric cancer. *Abdom Radiol N Y*. 2019;44:2048–58.
48. Ito K, Schöder H, Teng R, Humm JL, Ni A, Wolchok JD, et al. Prognostic value of baseline metabolic tumor volume measured on 18F-fluorodeoxyglucose positron emission tomography/computed tomography in melanoma patients treated with ipilimumab therapy. *Eur J Nucl Med Mol Imaging*. 2019;46:930–9.
49. Hatt M, Le Rest CC, Tixier F, Badic B, Schick U, Visvikis D. Radiomics: Data are also images. *J Nucl Med off Publ Soc Nucl Med*. 2019;60:S38–44.
50. Seban R-D, 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;47:2301–12.
51. Ma Y, Xia R, Ma X, Judson-Torres RL, Zeng H. Mucosal melanoma: pathological evolution, pathway dependency and targeted therapy. *Front Oncol*. 2021;11:702287.
52. Long GV, Swetter SM, Menzies AM, Gershenwald JE, Scolyer RA. Cutaneous melanoma. *Lancet Lond Engl*. 2023;402:485–502.
53. Ayati N, Sadeghi R, Kiamanesh Z, Lee ST, Zakavi SR, Scott AM. The value of 18F-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*. 2021;48:428–48.
54. Nakamoto R, Zaba LC, Liang T, Reddy SA, Davidzon G, Aparici CM, et al. Prognostic value of bone marrow metabolism on pretreatment 18F-FDG PET/CT in patients with metastatic melanoma treated with Anti-PD-1 therapy. *J Nucl Med off Publ Soc Nucl Med*. 2021;62:1380–3.
55. Schweighofer-Zwink G, Manaf-Farid R, Kölblinger P, Hehenwarter L, Harsini S, Pirich C, et al. Prognostic value of 2-[18F]FDG PET-CT in metastatic melanoma patients receiving immunotherapy. *Eur J Radiol*. 2022;146:110107.
56. Ito K, Teng R, Schöder H, Humm JL, Ni A, Michaud L, et al. 18F-FDG PET/CT for Monitoring of Ipilimumab Therapy in patients with metastatic melanoma. *J Nucl Med off Publ Soc Nucl Med*. 2019;60:335–41.
57. Annovazzi A, Ferraresi V, Rea S, Russillo M, Renna D, Carpano S, et al. Prognostic value of total metabolic tumour volume and therapy-response assessment by [18F]FDG PET/CT in patients with metastatic melanoma treated with BRAF/MEK inhibitors. *Eur Radiol*. 2022;32:3398–407.
58. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med*. 2009;50:S122–50.
59. Iravani A, Wallace R, Lo SN, Galligan A, Weppler AM, Hicks RJ, et al. FDG PET/CT prognostic markers in patients with Advanced Melanoma treated with Ipilimumab and Nivolumab. *Radiology*. 2023;307:e221180.
60. Flaus A, Habouzit V, De Leiris N, Vuillez JP, Leccia MT, Perrot JL, et al. FDG PET biomarkers for prediction of survival in metastatic melanoma prior to anti-PD1 immunotherapy. *Sci Rep*. 2021;11:18795.
61. Da-Ano R, Masson I, Lucia F, Doré M, Robin P, Alfieri J, et al. Performance comparison of modified ComBat for harmonization of radiomic features for multicenter studies. *Sci Rep*. 2020;10:10248.
62. Svedman FC, Pillas D, Taylor A, Kaur M, Linder R, Hansson J. Stage-specific survival and recurrence in patients with cutaneous malignant melanoma in Europe - a systematic review of the literature. *Clin Epidemiol*. 2016;8:109–22.
63. Forschner A, Eichner F, Amaral T, Keim U, Garbe C, Eigentler TK. Improvement of overall survival in stage IV melanoma patients during 2011–2014: analysis of real-world data in 441 patients of the German Central Malignant Melanoma Registry (CMMR). *J Cancer Res Clin Oncol*. 2017;143:533–40.
64. Kunz M. The genetic basis of new treatment modalities in melanoma. *Curr Drug Targets*. 2015;16:233–48.
65. Saadani H, van der Hiel B, Aalbersberg EA, Zavrakidis I, Haanen JBAG, Hoeks-tra OS, et al. Metabolic biomarker-based BRAFV600 Mutation Association and Prediction in Melanoma. *J Nucl Med off Publ Soc Nucl Med*. 2019;60:1545–52.
66. Demircioğlu A. Measuring the bias of incorrect application of feature selection when using cross-validation in radiomics. *Insights Imaging*. 2021;12:172.
67. Olthof S-C, Krumm P, Weichold O, Eigentler T, Bösmüller H, la Fougère C, et al. CT texture analysis compared to Positron Emission Tomography (PET) and mutational status in resected melanoma metastases. *Eur J Radiol*. 2020;131:109242.
68. Aoude LG, Wong BZY, Bonazzi VF, Brosda S, Walters SB, Koufariotis LT, et al. Radiomics biomarkers correlate with CD8 expression and predict Immune signatures in Melanoma patients. *Mol Cancer Res MCR*. 2021;19:950–6.
69. Dimitrakopoulou-Strauss A. Monitoring of patients with metastatic melanoma treated with immune checkpoint inhibitors using PET-CT. *Cancer Immunol Immunother CII*. 2019;68:813–22.
70. Brooks FJ, Grigsby PW. The effect of small Tumor volumes upon intra-tumoral Tracer Uptake Heterogeneity studies. *J Nucl Med off Publ Soc Nucl Med*. 2014;55:37–42.
71. Brooks FJ, Grigsby PW. Low-order non-spatial effects dominate second-order spatial effects in the texture quantifier analysis of 18F-FDG-PET images. *PLoS ONE*. 2015;10:e0116574.
72. Zwanenburg A, Vallières M, Abdallah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image Biomarker Standardization Initiative: standardized quantitative Radiomics for High-Throughput Image-based phenotyping. *Radiology*. 2020;295:328–38.
73. Keller H, Shek T, Driscoll B, Xu Y, Nghiem B, Nehmeh S, et al. Noise-based image harmonization significantly increases repeatability and reproducibility of Radiomics Features in PET images: a Phantom Study. *Tomogr Ann Arbor Mich*. 2022;8:1113–28.
74. Dittrich D, Pyka T, Scheidhauer K, Lütje S, Essler M, Bundschuh RA. Textural features in FDG-PET/CT can predict outcome in melanoma patients to treatment with Vemurafenib and Ipilimumab. *Nucl Nucl Med*. 2020;59:228–34.
75. Van de Wiele C, Kruse V, Smeets P, Sathekge M, Maes A. Predictive and prognostic value of metabolic tumour volume and total lesion glycolysis in solid tumours. *Eur J Nucl Med Mol Imaging*. 2013;40:290–301.
76. Guezennec C, Bourhis D, Orlhac F, Robin P, Corre J-B, Delcroix O, et al. Inter-observer and segmentation method variability of textural analysis in pre-therapeutic FDG PET/CT in head and neck cancer. *PLoS ONE*. 2019;14:e0214299.
77. Jardim S, António J, Mora C. Image thresholding approaches for medical image segmentation - short literature review. *Procedia Comput Sci*. 2023;219:1485–92.
78. Lasnon C, Enilrac 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.
79. 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.
80. Hatt M, Majdoub M, Vallières M, Tixier F, Le Rest CC, Groheux D, et al. 18F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort. *J Nucl Med off Publ Soc Nucl Med*. 2015;56:38–44.
81. Orlhac F, Soussan M, Maisonobe J-A, Garcia CA, Vanderlinden B, Buvat I. Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. *J Nucl Med off Publ Soc Nucl Med*. 2014;55:414–22.
82. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges J-P, et al. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med off Publ Soc Nucl Med*. 2011;52:369–78.
83. Crandall JP, Fraum TJ, Lee M, Jiang L, Grigsby P, Wahl RL. Repeatability of 18F-FDG PET Radiomic features in Cervical Cancer. *J Nucl Med off Publ Soc Nucl Med*. 2021;62:707–15.
84. Pfähler E, van Sluis J, Merema BBJ, van Ooijen P, Berendsen RCM, van Velden FHP, et al. Experimental multicenter and multivendor evaluation of the performance of PET radiomic features using 3-Dimensionally printed Phantom inserts. *J Nucl Med*. 2020;61:469–76.
85. Orlhac F, Soussan M, Chouahnia K, Martinod E, Buvat I. 18F-FDG PET-Derived Textural Indices reflect tissue-specific Uptake Pattern in Non-small Cell Lung Cancer. *PLoS ONE*. 2015;10:e0145063.
86. Vagenas TP, Economopoulos TL, Sachpekidis C, Dimitrakopoulou-Strauss A, Pan L, Provata A et al. A decision support system for the identification of

- metastases of Metastatic Melanoma using whole-body FDG PET/CT images. *IEEE J Biomed Health Inf.* 2022;PP.
87. Anne-Leen D, Machaba S, Alex M, Bart DS, Laurence B, Mike S, et al. Principal component analysis of texture features derived from FDG PET images of melanoma lesions. *EJNMMI Phys.* 2022;9:64.
  88. Chang E, Joel MZ, Chang HY, Du J, Khanna O, Omuro A, et al. Comparison of radiomic feature aggregation methods for patients with multiple tumors. *Sci Rep.* 2021;11:9758.
  89. Peisen F, Hänsch A, Hering A, Brendlin AS, Afat S, Nikolaou K, et al. Combination of whole-body baseline CT Radiomics and Clinical parameters to predict response and survival in a Stage-IV Melanoma Cohort Undergoing Immunotherapy. *Cancers.* 2022;14:2992.
  90. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging.* 2013;26:1045–57.
  91. Lee H, Palm J, Grimes SM, Ji HP. The Cancer Genome Atlas Clinical Explorer: a web and mobile interface for identifying clinical-genomic driver associations. *Genome Med.* 2015;7:112.
  92. Yip SSF, Aerts HJWL. Applications and limitations of radiomics. *Phys Med Biol.* 2016;61:R150–166.
  93. Bogowicz M, Leijenaar RTH, Tanadini-Lang S, Riesterer O, Pruschy M, Studer G, et al. Post-radiochemotherapy PET radiomics in head and neck cancer - the influence of radiomics implementation on the reproducibility of local control tumor models. *Radiother Oncol J Eur Soc Ther Radiol Oncol.* 2017;125:385–91.
  94. Da-Ano R, Visvikis D, Hatt M. Harmonization strategies for multicenter radiomics investigations. *Phys Med Biol.* 2020;65:24TR02.
  95. Kothari S, Phan JH, Stokes TH, Osunkoya AO, Young AN, Wang MD. Removing batch effects from histopathological images for enhanced cancer diagnosis. *IEEE J Biomed Health Inf.* 2014;18:765–72.
  96. Fortin J-P, Cullen N, Sheline YI, Taylor WD, Aselcioglu I, Cook PA, et al. Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage.* 2018;167:104–20.
  97. Müller C, Schillert A, Røthemeier C, Trégouët D-A, Proust C, Binder H, et al. Removing batch effects from Longitudinal Gene expression - quantile normalization plus ComBat as Best Approach for microarray Transcriptome Data. *PLoS ONE.* 2016;11:e0156594.
  98. Stein CK, Qu P, Epstein J, Buros A, Rosenthal A, Crowley J, et al. Removing batch effects from purified plasma cell gene expression microarrays with modified ComBat. *BMC Bioinformatics.* 2015;16:63.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.