International Journal of Radiation Oncology biology • physics

www.redjournal.org

Critical Review

Radiomics in Nuclear Medicine Applied to Radiation Therapy: Methods, Pitfalls, and Challenges

Sylvain Reuzé, MSc,[*,†,‡] Antoine Schernberg, MD, MSc,[§] Fanny Orlhac, PhD,[‖] Roger Sun, MD, MSc,[*,†,§] Cyrus Chargari, MD, PhD,[*,†,§,¶,] Laurent Dercle, MD, MSc,[††,‡‡] Eric Deutsch, MD, PhD,[*,†,§] Irène Buvat, PhD,[‖] and Charlotte Robert, PhD[*,†,‡]**

Radiothérapie Moléculaire, Univ. Paris-Sud, Institut Gustave Roussy, Inserm, Université Paris-Saclay, 94800, Villejuif, France; [†]Université Paris Sud, Université Paris-Saclay, Le Kremlin-Bicêtre, France; [‡] Department of Radiotherapy — Medical Physics, Gustave Roussy, Université Paris-Saclay, Villejuif, France; [§]Department of Radiotherapy, Gustave Roussy, Université Paris-Saclay, Villejuif, France; [‖]IMIV, CEA, Inserm, CNRS, Université Paris-Sud, Université Paris-Saclay, Orsay, France; [¶]French Military Health Services Academy, Ecole du Val-de-Grâce, Paris, France; **Institut de Recherche Biomédicale des Armées, Bretigny-sur-Orge, France; [††]Department of Nuclear Medicine and Endocrine Oncology, Gustave Roussy, Université Paris-Saclay, Villejuif, France; and [‡‡]Immunologie des tumeurs et Immunothérapie, Univ. Paris-Sud, Institut Gustave Roussy, Inserm, Université Paris-Saclay, 94805, Villejuif cedex, France

Received Jan 5, 2018, and in revised form Apr 27, 2018. Accepted for publication May 2, 2018.

Radiomics is a recent area of research in precision medicine and is based on the extraction of a large variety of features from medical images. In the field of radiation oncology, comprehensive image analysis is crucial to personalization of treatments. A better characterization of local heterogeneity and the shape of the tumor, depicting individual cancer aggressiveness, could guide dose planning and suggest volumes in which a higher dose is needed for better tumor control. In addition, noninvasive imaging features that could predict treatment outcome from baseline scans could help the radiation oncologist to determine the best treatment strategies and to stratify patients as at low risk or high risk of recurrence. Nuclear medicine molecular imaging reflects information regarding biological processes in the tumor thanks to a wide range of radiotracers. Many studies involving ^{18}F-fluorodeoxyglucose positron emission tomography suggest an added value of radiomics compared with the use of conventional PET metrics such as standardized uptake value for both tumor diagnosis and prediction of recurrence or treatment outcome. However, these promising results should not hide technical difficulties that still currently prevent the approach from being widely studied or clinically used. These difficulties mostly pertain to the variability of the imaging features as a function of the acquisition device and protocol, the robustness of the models with respect to that variability, and the interpretation of the radiomic models. Addressing the impact of the variability in acquisition and reconstruction protocols is needed, as is harmonizing the radiomic feature calculation methods, to ensure the reproducibility of studies in a multicenter context and their implementation in a clinical workflow. In this review, we explain the potential impact of positron emission tomography radiomics for radiation therapy and underline the various aspects that need to be carefully addressed to make the most of this promising approach.

Reprint requests to: Charlotte Robert, PhD, Institut Gustave Roussy — Département de Radiothérapie, 114 rue Edouard Vaillant, 94800 Villejuif, France; E-mail: ch.robert@gustaveroussy.fr

Conflict of interest: none.

Supplementary material for this article can be found at www.redjournal.org.

Int J Radiation Oncol Biol Phys, Vol. ■, No. ■, pp. 1–26, 2018
0360-3016/$ - see front matter
https://doi.org/10.1016/j.ijrobp.2018.05.022

Introduction

Radiomics is a promising research area based on mapping multimodal imaging onto quantitative data and subsequent predictive model building to personalize decision making (1-3). Assessment of tumor heterogeneity and shape using medical imaging could enhance patient outcomes by using routinely available data and could be used as a substitute for repeated biopsies to monitor tumor evolution (4). In the field of radiation oncology, imaging features, now often called "radiomic features" (2, 5), computed before or during treatment could be used for patient stratification and dose planning guidance (6).

In recent years, nuclear medicine techniques such as positron emission tomography (PET) or single photon emission computed tomography (SPECT) have been used increasingly in oncology for diagnosis, detection of lymph nodes and distant metastases, characterization of tumor biology, and treatment response assessment (7). In addition to computed tomography (CT) and magnetic resonance imaging (MRI), the use of PET images in the radiation therapy workflow has been associated with more accurate tumor targeting, more reproducible segmentation, and significant adjustments in patient management and radiation therapy dose planning (8).

Several quantitative parameters can be extracted from PET images. Since the early 1990s, it has been well known that radiotracer uptake as quantified by standardized uptake value (SUV) is correlated with tumor aggressiveness in various cancer types (9-13). Volumetric features such as metabolic tumor volume (MTV) and total lesion glycolysis, which integrate both uptake and volumetric information, have also been related to patient outcomes (14-18). Although widely used, ^{18}F-fluorodeoxyglucose (FDG)-based PET SUV or volumetric parameters are insufficient to fully describe tumor characteristics related to tumor aggressiveness and poorer prognosis (eg, tumor heterogeneity, hypoxia, cellular proliferation, or necrosis) (19). Other tracers are required to target specific molecular features, and, for a given tracer, advanced image features should be calculated to extensively describe the tracer distribution within the tumor.

Three classes of radiomic features depicting the organization of voxel intensities are commonly used in PET. First-order features are derived directly from voxel values or from the histogram of intensities. They include several usual metrics such as the maximum or mean SUV value. Second-order features—"textural features," including Haralick's features (20)—are computed from gray-level matrices and reflect the spatial heterogeneity of voxel values within the region of interest (ROI). Higher-order features are extracted after image filtering. Shape features (eg, sphericity) are also of interest (21, 22).

The following goals can be targeted for the integration of PET radiomic features into radiation therapy clinical practice: (1) stratification of patients as a function of disease prognosis or outcome prediction, (2) improvement of tumor volume segmentation, (3) treatment planning optimization, and (4) response assessment (Fig. 1). In the past decade, despite tremendous efforts and great interest in using PET for radiation therapy, the major benefit has largely consisted of the optimization of tumor staging (23). The implementation of PET in adaptive radiation therapy has largely remained at the stage of ongoing clinical trials (24).

The aim of this article is to review the state of the art of implementation of PET radiomic studies in the context of personalized radiation therapy. To our knowledge, this is the first review focusing specifically on the use of PET imaging for the purpose of radiation therapy, for medical physicists and researchers in radiation therapy, including (1) an overview of published studies, (2) the state of the art of ongoing clinical trials involving PET radiomics for radiation therapy, (3) a didactic review of methodological pitfalls and challenges linked to radiomics analysis in PET imaging, and (4) an overview of statistical analysis and machine learning concepts. Physicians and physicists should indeed be aware of the large risks of biases generated by the lack of standardization in the acquisition process, reconstruction of images, postprocessing, or statistical learning. Whenever possible, a critical review of the proposed solutions is performed. Several recommendations are proposed at each step of the radiomic pipeline, and indications regarding future necessary studies are provided.

Applications of PET-based Radiomics in Radiation Oncology

This section briefly summarizes the pipeline for image feature calculation in PET (Fig. 2). The main applications of radiomics in the radiation oncology context are then described.

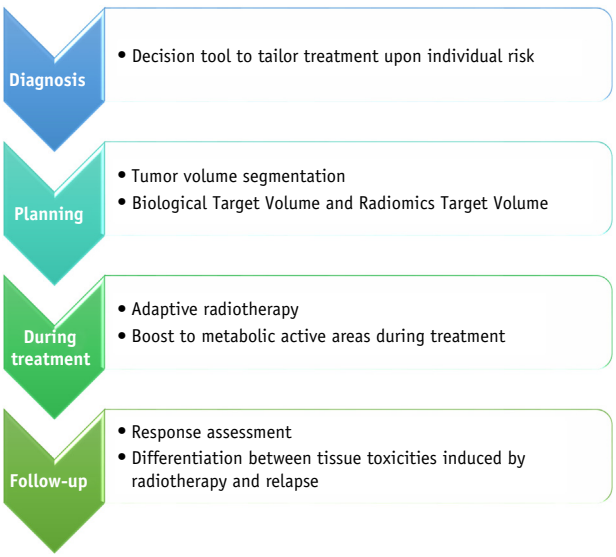

Fig. 1. Main applications of radiomics in the context of radiation oncology.

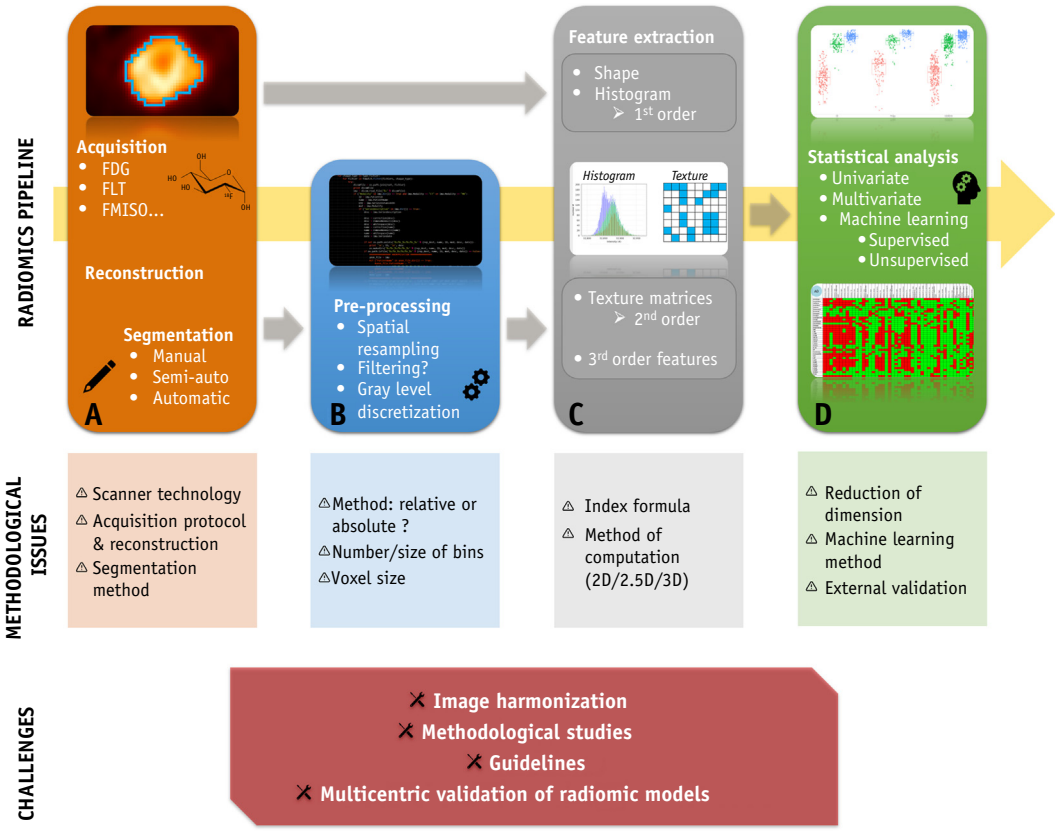

Fig. 2. Radiomics pipeline, methodological issues and challenges.

Radiomic feature calculation

In FDG-PET, the only quantitative feature that is routinely extracted from the images is the SUV, most often calculated from a single voxel in the tumor (SUV_{max}) or from a small ROI (1 mL) corresponding to the most metabolically active part of the tumor (SUV_{peak}) (25). It is obvious that such a measurement is far too crude to reflect possible metabolic heterogeneity within the tumor that could be taken into account to optimize treatment planning. In the context of radiomics, advanced imaging features are therefore investigated to better characterize metabolic heterogeneity. These features include histogram-derived indices (Fig. 3A) and textural features (Fig. 3B). Shape features are not included in this review.

In the context of image analysis, a histogram is a graphical representation of voxel intensities. The range of values is divided into a series of intervals (called "bins"). Bin height is representative of the number of voxels with an SUV value that belong to the corresponding interval (Fig. 4A). From the histogram, several features may be extracted, depicting skewness, kurtosis, or variance of the distribution. Spatial arrangement of voxels is not reflected by histogram-derived indices (Fig. 3) (26).

Textural features are derived from the original PET images (yielding second- or third-order features) or from transformed versions of these images (yielding higher order features), and unlike histogram-based features, they do reflect the spatial arrangement of voxel values. The image transformation process for high-order feature calculation may be achieved by Laws filters, such as Laplacian or Gaussian transformation, which highlight edge structures, or by using wavelet decomposition, which reveals sharp transitions in the intensity frequency spectrum (27, 28). They are therefore good candidates to describe tumor heterogeneity. Four textural matrices are usually computed to derive textural features (Figs. 2C and 3):

1. The gray-level co-occurrence matrix reflects the probability of observing a pair of values in voxels at a given distance in a given direction (20);
2. The gray-level run-length matrix (GLRLM) measures the number of consecutive voxels with the same value aligned in a given direction for any voxel value (29);
3. The gray-level size zone matrix (GLSZM) measures the number of neighboring voxels with the same value for any voxel value (30);
4. The neighborhood gray-level difference matrix measures the difference between neighboring voxels based on voxel values (31).

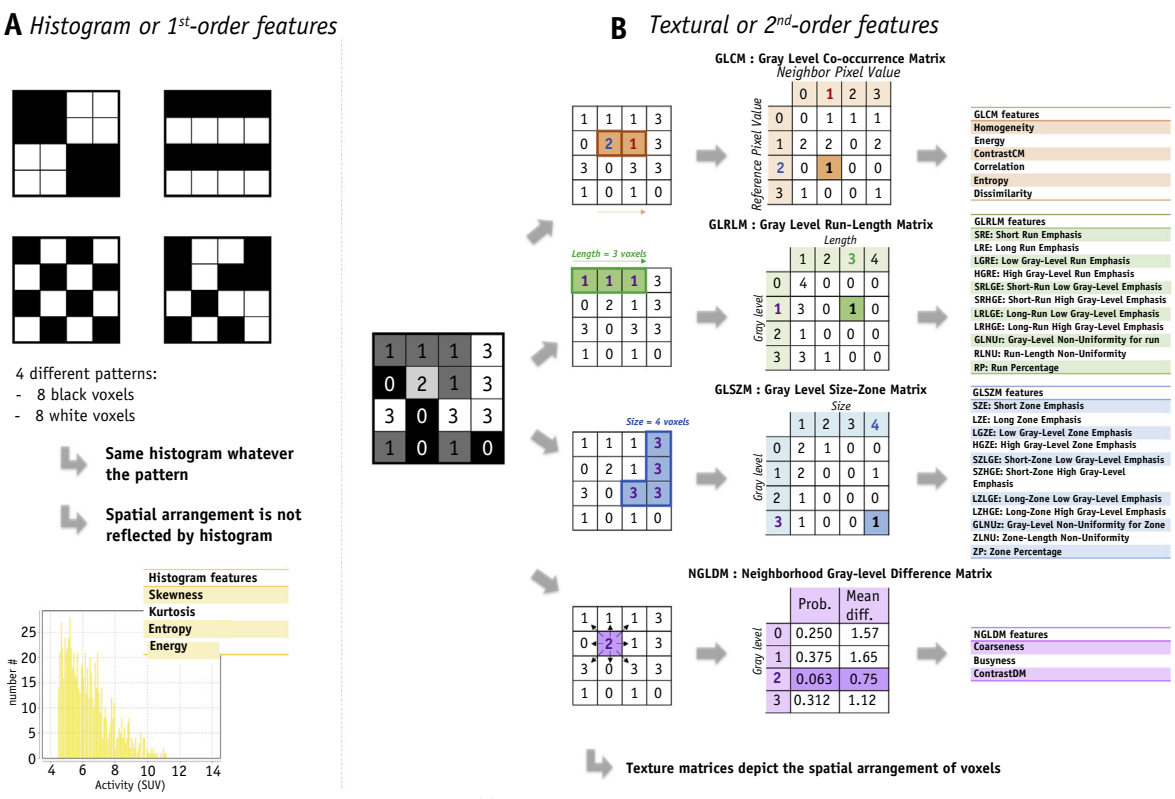

Fig. 3. Histogram or first-order features reflect voxel intensity distribution only (A). For different patterns containing the same number of black and white pixels, but not at the same location, the histogram is identical. Consequently, histogram-based feature values (eg, skewness, kurtosis, energy) are equal. Unlike histogram-based features, textural or second-order features (B) derived from texture matrices (eg, co-occurrence, run length, zone size matrices) reflect the complex and unique spatial arrangement of voxels.

Textural features can be computed in 2 dimensions (2D) or 3 dimensions (3D) using different methods (32), or in 2.5 dimensions, meaning that values are calculated in 2D for each slice and then averaged over all slices in the volume. Most PET radiomic studies are currently using 3D features to reflect the heterogeneity of the entire tumor (32-46).

With the advent of 3D imaging and knowing that tumors may include different heterogeneous uptake patterns, a 3D calculation in the whole volume seems more clinically relevant than 2D features computed in one unique slice of interest.

In addition, formulas can differ between studies, and features computed from different matrices can have the same name (47). For example, entropy can be computed either from a histogram of intensities or from the GLCM matrix accounting for spatial repartition of intensities, leading to different values (eg, 1.6 from the histogram and 2.8 from GLCM for a cervix tumor; Fig. 4B) (26).

To be widely applicable, a radiomic model should thus come with an unambiguous definition of the radiomic features involved in the model. To tackle that issue, a list of formulas for radiomic feature calculation is being proposed by the Image Biomarker Standardization Initiative (48).

Advanced PET quantitative features for outcome prediction in radiation therapy

Radiomics is mostly used as a tool for prediction of treatment outcome in terms of survival or pathologic tumor response. A list of clinical studies involving patients undergoing radiation therapy and radiomic analysis involving advanced features is presented in Table 1 (32-46, 49-62). All 28 studies were retrospective, and most (79%) included fewer than 100 patients. Only 3 studies included a validation of the radiomic model on data that differed from the data used to derive the model (42, 46, 59). The most frequent cancer types were lung (n = 8), esophageal (n = 6), head and neck (n = 5), and cervical (n = 3). Two studies evaluated radiation-induced damage in brain metastases and head and neck cancer (41, 49), whereas the main outcome for all other studies was overall survival or objective tumor response rate. Only 2 studies were conducted using tracers other than FDG: ^{18}F-fluoro-ethyl-tyrosine for brain metastasis (49) and ^{68}Ga-DOTATATE/DOTATOC (68Gallium-DOTA-Tyr(3)-Thr(8)-octreotate) for thyroid (62). A total of 25 studies investigated conventional (SUV and volumetric features), first-order and

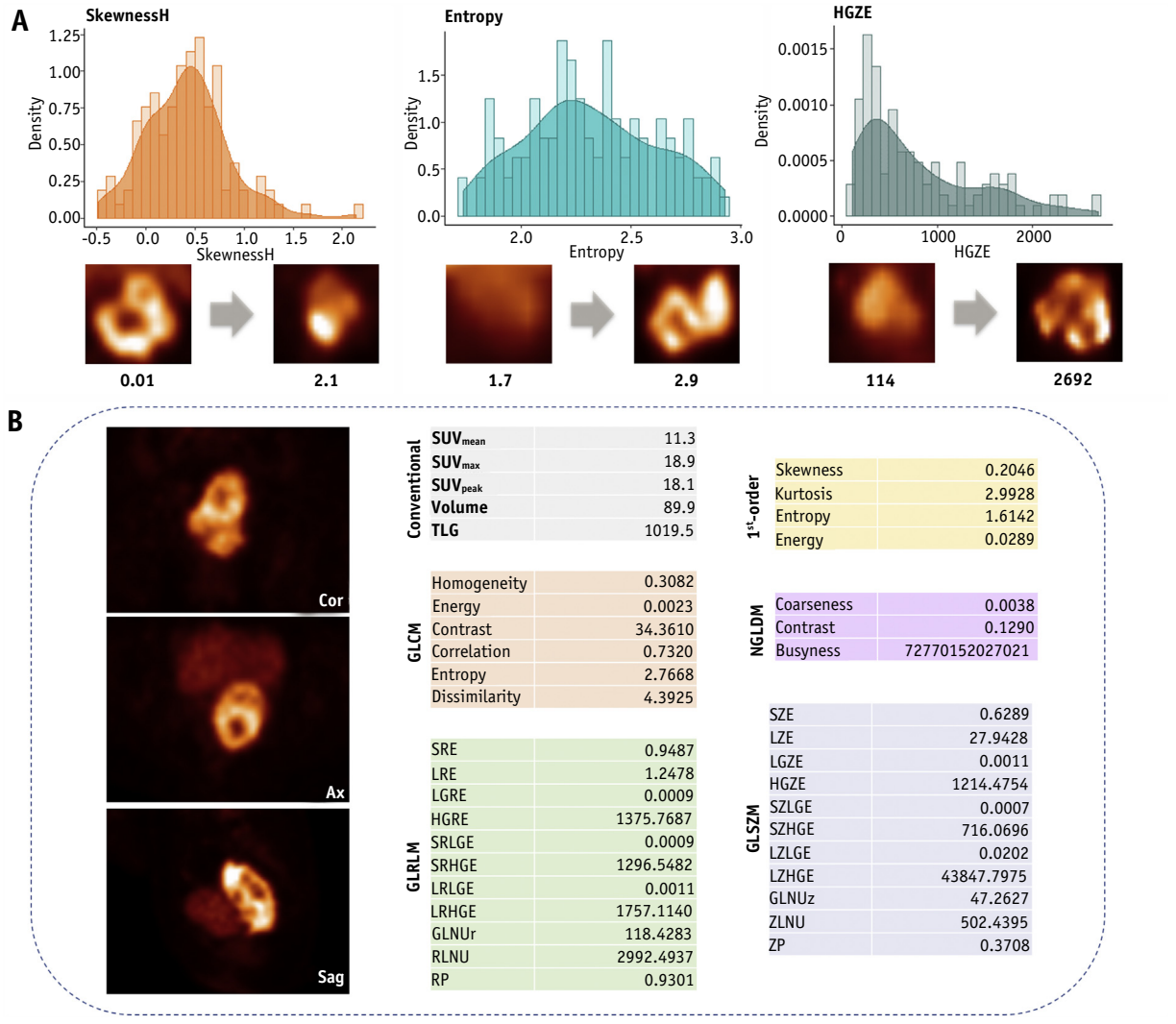

Fig. 4. (A) Comparison of distributions for 3 features—SkewnessH from the histogram of intensities, entropy from GLCM matrix, and HGZE from GLSZM matrix—with examples from patients with locally advanced cervical cancer.(B) Example of radiomic feature values: squamous cell cervical carcinoma, FIGO stage IIb, 78 years old; General Electric Discovery-690 PET-CT, voxel size: 2.7 mm × 2.7 mm × 3.4 mm; absolute discretization in 128 gray levels between 0 and 40 SUV. Features are classified by type (conventional, other first-order features, and features from textural matrices). *Abbreviations:* FIGO = fédération internationale de gynécologie et d'obstétrique (international federation of gynecology and obstetrics); GLCM = gray-level co-occurrence matrix; GLSZM = gray-level size zone matrix; HGZE = high gray-level zone emphasis; PET-CT = positron-emission tomography–computed tomography; SUV = standardized uptake value.

second-order features, whereas 3 papers included only conventional and histogram features (34, 36, 55). Tumors were mostly segmented using a threshold (n = 13) or manually (n = 10), and several teams compared the impact of different methods of delineation (32, 44, 50). In most studies, advanced radiomic features were not statistically significant predictors of treatment outcome in univariate analysis, and their added value was only suggested in multivariate models, leading to better results than SUV. Conclusions from these publications cannot always be tested on other data because of the absence of methodological information or the use of homemade software without any details about how the features were calculated (discretization, feature formulas). When homemade software is used, authors should give access to the code to ensure reproducibility of their work. The use of freely available radiomic software—for example, IBEX (63), LIFEx (64), or PyRadiomics (65)—could be an important step toward validation of the models in independent cohorts. A radiomic model will be fully validated only if it is successfully assessed by independent investigators. This might require some standardization of radiomic software so that features of a given name are calculated identically in different software and identical or similar values are obtained for each feature by different software when using the same input image. Furthermore, the availability of input

Table 1 List of clinical publications with keywords "(Radiomics or texture or textural) AND PET AND Radiotherapy"

Reference	Cancer type	Nb. of patients	Software	No. of scanners	Segmentation method	Features	Modality	Method of calculation	Stats	Outcome	Findings
Lohmann et al. (49)	Brain metastasis	47	MATLAB (CGIT)	1	Threshold (TBR = 1.6)	Histogram, GLCM, GLRLM, NGLDM, GLSZM	FET-PET	relative discretization, 64 levels	Fleiss's Kappa analysis, ROC analysis	Discrimination between brain metastasis recurrence and radiation injury	27 radiomic features had an AUC >0.7; TBR_{mean} alone had an accuracy of stratification of 81%; the highest accuracy (85%) was reached when combining TBR_{mean} and Coarseness_NIDM or SZE ($P < .01$); improvement in diagnostic accuracy was not significant.
Roman-Jimenez et al. (50)	Cervix	53	-	-	Region-growing with different thresholds (10%-90% SUV_{max})	Conventional, histogram, shape, GLCM, GLRLM	FDG-PET	Relative discretization, 64 levels	RF model	OS	36 out of the 1026 features were highly significant (AUC = 0.72-0.83; $P < .01$); AUC = 0.90 with a 9-feature model based on volumetric and shape features only after recursive elimination of the least important variables.
Yang et al. (38)	Cervix	20	-	1	Threshold (40% SUV_{max})	Conventional, GLRLM, GLSZM	FDG-PET	3D, relative discretization, 256 levels	Friedman's ANOVA, Mann-Whitney test	Evolution during treatment, assessment of treatment response	No feature was able to distinguish between CMR and PMR/NEW at baseline; 10 textural features showed significant changes during treatment in the CMR group ($P < .05$), as did 10 features in the PMR/NEW group ($P < .05$).
Reuzé et al. (42)	Cervix	118	LIFEx	2	Threshold (40% SUV_{max})	Conventional, GLCM, GLRLM, GLSZM	FDG-PET	3D, absolute discretization, 0.3 SUV step	Wilcoxon test, ROC analysis, AIC for feature selection	Local recurrence prediction	A 4-feature signature can predict local recurrence of locally advanced cervical cancer better than SUV_{max}. High variability observed in texture index between scanners corresponding to different generations.
Beukinga et al. (45)	Esophagus	97	MATLAB	1	Manual + Erosion	Clinical, geometry, conventional, histogram, GLCM, GLRLM, GLSZM	FDG-PET + CT	3D, bins of 0.5 g/mL	Univariate logistic model, multivariate logistic model, ROC	Complete vs incomplete response	A prediction model including clinical, geometry, CT, and PET features leads to an AUC of 0.78 compared to 0.58 with a simpler model including SUV_{max} only. This model was internally validated using a bootstrap approach (AUC = 0.74).

Desbordes et al. (51)	Esophagus	65	-	1	Contrast-based adaptive threshold algorithm	Clinical data, Conventional, GLCM, GLDM, GLSZM	FDG-PET	3D, absolute discretization, 0.5 SUV step	Spearman's rank correlation analysis, RF feature selection, ROC analysis, K-M curves	Response to therapy at 1 month, OS	Improvement of predictive and prognostic values by using the RF classifier compared to Mann-Whitney U Test and Univariate K-M survival analysis.
Nakajo et al. (43)	Esophagus	52	Home-made	1	Threshold (SUV >2.5)	Conventional, histogram, textural features (local and regional heterogeneity parameters)	FDG-PET	3D, absolute discretization, 64 levels	Spearman rank test, ROC analysis, K-M curves	Response to treatment	Higher MTV, TLG, intensity variability and size-zone variability shown in nonresponder patients. Limited added value of these parameters for the patient prognosis.
Tanet al. (52)	Esophagus	20	ITK	1	Connected-threshold method for SUV >2.5	Conventional, histogram, shape, GLCM	FDG-PET	Relative discretization	ROC analysis, Mann-Whitney test	Pathologic tumor response	3 textural features computed on postchemoradiation PET images were statistically predictive of treatment response (AUC >0.78, $P < .04$); pretreatment skewness, difference and ratio of SUV_{max}, and difference of SUV_{mean} were statistically significant (AUC >0.79 and $P < .05$).
Tixier et al. (37)	Esophagus	41	Home-made	1	FLAB	Conventional, histogram, GLCM, NGLDM, GLRLM, GLSZM	FDG-PET	3D, relative discretization	Kruskal-Wallis test, ROC analysis	Response to treatment	Textural features were highly predictive of treatment outcome (entropy: $P = .0006$; regional intensity variability: $P = .0002$).
van Rossum et al. (53)	Esophagus	217	IBEX	1	Gradient based	Conventional, histogram and textural features	FDG-PET	-	Univariate/ multivariate analyses with c-indices	Pathologic tumor response	Difference of run percentage and ICM entropy were predictive of higher chance of complete response (corrected c-index = 0.77 in combination with clinical and conventional parameters).
Van Dijk et al. (41)	HNSCC	161	-	1	-	Conventional, histogram, GLCM, GLRLM, GLSZM, NGLDM	FDG-PET+CT	3D, absolute discretization, bin width of 0.25 SUV	Univariable logistic regression, multivariate model using lasso regularization	Late radiation-induced xerostomia	11 histogram and 35 textural features were significantly associated with xerostomia at 12 months ($P < .05$); a basic model (SUV_{mean}, parotid dose, xerostomy at baseline) had an AUC of 0.77; the same AUC was reached when replacing SUV_{mean} by the 90th percentile of SUVs, or LRHGE.

(*continued on next page*)

Table 1 (*continued*)

Reference	Cancer type	Nb. of patients	Software	No. of scanners	Segmentation method	Features	Modality	Method of calculation	Stats	Outcome	Findings
Yu et al. (35)	HNSCC	40 (20 with tumor + 20 controls)	-	1	Manual	Conventional, histogram, GLCM, NGLDM, structural features	FDG-PET + CT	3D	K nearest neighbors and decision tree–based classifiers, ROC analysis	Discrimination between normal and tumor tissue	Textural analysis was able to discriminate between normal and abnormal tissues as well as a human expert; combination of PET and CT features improved results
Oh et al. (55)	HNSCC	27	home-made	2	Manual	Conventional and histogram	FDG-PET	-	Mann-Whitney U test, Cox proportional hazard models	DFS, OS	Baseline PET-based coarseness and busyness are highly predictive of chemoradiation response (best result with coarseness: AUC = 0.82, $P = .001$).
Cheng et al. (56)	HNSCC	70	MATLAB (CGIT)	2	Threshold (SUV >2.5)	Conventional, histogram, GLCM, NGLDM	FDG-PET	Relative discretization, 4-64 levels	Student *t*-test, K-M analysis, Cox models	PFS, OS, DSS	Uniformity from GLCM is an independent prognostic predictor (PFS, DSS: $P = .001$; OS: $P = .017$).
Chen et al. (44)	HNSCC	57	Home-made	1	Threshold (various methods)	Conventional, GLCM, GLRLM, NGLDM, GLSZM	FDG-PET	3D, absolute discretization between SUV_{min} and SUV_{max}, fixed bin size	Mann-Whitney U test, ROC curves, K-M curves	Correlation with genomic expression, DSS, PRFS	The overexpression of VEGF was associated with increased values of GLNUz and RLNU. Textural features combined with T-stage and immune-histochemical data can help to stratify treatment outcome.
Cook et al. (36)	NSCLC	53	-	2	Threshold 45% SUV_{max}, corrected by an expert	Conventional and histogram	FDG-PET	3D	K-M analysis, ROC curves, Mann-Whitney tests with CT RECIST	OS, PFS, local PFS	Coarseness, contrast, and busyness are associated with treatment failure by RECIST ($P = .004$, $P = .002$, $P = .027$ respectively) and poorer prognosis (PFS, LPFS: $P < .02$; OS: $P = .003$ for coarseness).
Dong et al. (39)	NSCLC	58	MATLAB	1	Threshold (SUV >3.0)	Conventional, histogram, GLCM	FDG-PET	3D, relative discretization, 64 levels	Wilcoxon test, ROC analysis, K-M analysis	Treatment response, OS, PFS	Baseline contrast and COV were the best predictors of treatment response (AUC = 0.804 and 0.781, respectively); change of radiomic features during treatment was more accurate (Δcontrast: AUC = 0.862, ΔCOV: AUC = 0.799); Δcontrast >70.3% was associated with improved PFS ($P =$

											.007); PFS was lower when ΔAUC-CSH was lower (P = .039); both baseline contrast and Δcontrast were associated with OS (P = .02 and .008, respectively); Δcontrast was the only textural feature with significant independent prognostic value for OS and PFS on multivariate analysis (P = .021 and .015, respectively).
Kirienko et al. (46)	NSCLC	295	LIFEx	2	Threshold (40% SUV_{max})	Conventional, histogram, shape, size, second and higher order features	FDG-PET + CT	3D, absolute discretization, 64 levels	Univariate analysis, multivariate Cox proportional K-M curves	DFS	Different Cox models were developed. The one including radiomic features from PET+CT images resulted in an AUC of 0.68
Lovinfosse et al. (57)	NSCLC	63	Home-made	2	FLAB	Conventional, histogram, GLCM, NGLDM, GLSZM	FDG-PET	Relative discretization, 64 levels	Logistic regression, K-M analysis	OS, DSS, DFS	On univariate analysis, only dissimilarity was associated with DSS (HR = 0.822, P = .037); 8 textural and histogram features were associated with DFS (P < .05); on multivariate analysis, 2 cut-off values were established for stratification with dissimilarity (DSS: P = 0.092; DFS: P = 0.0034).
Takeda et al. (58)	NSCLC	26	MATLAB (CGIT)	1	Manual	Conventional, histogram, GLCM, GLSZM	FDG-PET	Relative discretization, 64 levels	ICC, K-M analysis	Robustness to segmentation, LC, PFS, OS	For all features except ZP (ICC = 0.65), ICC was between 0.81 and 1.00 for 2 different delineations by 2 observers.
Vaidya et al. (33)	NSCLC	27	CERR	1	Manual	Conventional, histogram, GLCM	FDG-PET + CT	3D	Spearman's correlation and multivariable logistic regression	Local and locoregional recurrence	IVH had the highest univariate association with locoregional recurrence; a combination of CT-V70 and PET-V80 yielded the best correlation with loco-regional and local failures (r = 0.49 and r = 0.59, respectively).

(*continued on next page*)

Table 1 (*continued*)

Reference	Cancer type	Nb. of patients	Software	No. of scanners	Segmentation method	Features	Modality	Method of calculation	Stats	Outcome	Findings
Wu et al. (59)	NSCLC	101	MATLAB	1	Manual and FCM	Conventional, histogram, shape, GLCM, wavelets, Laws family	FDG-PET	-	35 robust, reproducible, nonredundant features in the clinical analysis; 70 training, 31 validation; *glmnet* for feature selection, Cox model for prediction	Robustness of features; prediction of distant metastasis	A signature combining SUV_{peak} and Gauss Cluster Shade (Laws filter group) was identified for prediction of distant metastases (training: c-index = 0.73, HR = 5.4, P = .0019; validation: c-index = 0.71, HR = 4.8, P = .0498); when histological subtype is added in the signature, c-index = 0.80 (P < .001), whereas histological type alone was not significant (P = .77).
Yip et al. (60)	NSCLC	26	MATLAB (CGIT)	1	Manual	NGLDM, GLCM, GLRLM	FDG-PET	Absolute discretization (32 levels) on rescaled images	Relative difference between 3D and 4D-based features; Kruskal-Wallis-test	Sensitivity of textural features to tumor motion (comparison 3D/4D images)	Significant differences were found between 3D and 4D PET imaging for 4 features (1%-10%, P < .05); variability between phases of 4D-PET was negligible; 4D-PET may have better prognostic value with radiomics analysis.
Yue et al. (40)	Pancreas	26	-	1	Manual except for SUV_{mean} calculation (threshold 40% SUV_{max})	Conventional, GLCM	FDG-PET	3D, relative discretization, 16 levels	Selection of feature using lasso/ elasticnet regression; Cox regression, ROC analysis, K-M analysis	OS	Any PET feature was prognostic at baseline; 8 features were significant on post-RT PET images; changes in SUV_{max}, entropy, variance, sum mean, cluster tendency, and maximum probability between pre-RT and post-RT images were highly predictive of OS (P < .01); after variable selection, a multivariate model gathering age, clinical N-stage, Δhomogeneity, Δvariance, and Δcluster tendency was identified (P = .001).

Bang et al. (61)	Rectum	74	MaZda	1	Different threshold-based contours for MTV and TLG; manual segmentation for texture analysis	Conventional, histogram, GLCM, GLRLM	FDG-PET	Relative discretization, 64 levels in mean ± 3σ	Univariate/ multivariate binary logistic regression and Cox regression analyses	Treatment response, 3-year DFS	MTV, histogram, and textural features from GLCM were significantly associated with response to chemoradiation ($P <$ .05 for MTV extracted with liver SUV_{mean} ± 3σ threshold and histogram-based features, $P <$ .01 for GLCM sum entropy and entropy); textural features combined with MTV were predictive of recurrence.
Bundschuh et al. (34)	Rectum	27	Interview Fusion Workstation	1	Manual	Conventional and histogram	FDG-PET	3D	ROC analysis, Youden index, Pearson correlation analysis, K-M analysis	Histopathologic response, prediction of survival	COV statistically significant to assess histopathologic response during and after chemoradiation; AUC of COV was higher than conventional features for early and late response assessment.
Lapa et al. (62)	thyroid	12	Interview Fusion Workstation	3	Manual	Conventional, histogram, GLCM, GLSZM	^{68}Ga-PET	-	Pearson correlation, t-test, ROC analysis, K-M analysis	OS, PFS	Contrast (AUC = 0.89, P = .0006), GLNU (AUC = 0.93, P = .007), intensity variation (AUC = 0.89, P = .001), SRE (AUC = 0.85, P = .018) correlated significantly with PFS; a cut-off between responders and nonresponders was identified for these features (P = .04, P = .04, P = .01, P = .02, respectively); no significant correlation was found between OS and textural or conventional features.

(*continued on next page*)

Table 1 (*continued*)

Reference	Cancer type	Nb. of patients	Software	No. of scanners	Segmentation method	Features	Modality	Method of calculation	Stats	Outcome	Findings
Hatt et al. (32)	Multisite: breast, cervix, esophagus; NSCLC; head and neck	555	Home-made	2	FLAB and manual	Conventional, GLCM, GLSZM	FDG-PET	3D (2 methods) / different methods of relative discretization	Spearman's correlation, Cox models, K-M curves	Methodology + OS in NSCLC and esophageal cancer	Dissimilarity and entropy were found to have prognostic value in esophageal cancer (HR = 1.92; $P = .0052$) and NSCLC (HR = 2.82; $P < .0001$), respectively; in NSCLC, heterogeneity, volume, and stage were independent prognostic factors that allowed stratification of patients when combined ($P < .0001$).

Abbreviations: AIC = Akaike Information Criterion; AUC = area under the curve; CMR = complete metabolic response; COV = coefficient of variation; CSH = area under the cumulative SUV-volume histogram; DFS = disease-free survival; DSS = disease-specific survival; FET-PET = ^{18}F-fluoro-ethyl-tyrosine; FLAB = fuzzy locally adaptive bayesian segmentation; GLNU = gray-level non-uniformity; GLRLM = gray-level run-length matrix; GLSZM = gray-level size zone matrix; HNSCC = head and neck squamous cell carcinoma; ICC = interclass correlation coefficient; ITK = insight toolkit platform for image analysis; K-M = Kaplan-Meier; IVH = intensity volume histogram; LC = local control; LRHGE = ; MTV = metabolic tumor volume; NEW = progression; NGLDM = neighborhood gray-level difference matrix; NSCLC = non-small cell lung cancer; OS = overall survival; PFS = progression-free survival; PMR = partial metabolic response; PRFS = primary relapse-free survival; RECIST = Response Evaluation Criteria In Solid Tumors; RF = Random Forest; ROC = receiver operating characteristic; SRE = short run emphasis; SZE = short zone emphasis; TBR = tumor to brain ratio; TLG = total lesion glycolysis; VEGF = vascular endothelial growth factor.

data (anonymized segmented images) could facilitate multicenter studies and verification of results. To our knowledge, only 4 nonrandomized clinical trials combining PET radiomics and radiation therapy have been proposed (Table E1; available online at www.redjournal.org). These studies aim to develop a decision support system, correlate radiomic features with immune modulation, or evaluate the change in tumor heterogeneity during treatment.

Biological target volume and biologically guided dose painting

Tumor heterogeneity is known to be a determinant factor in tumor progression. There is currently a lack of noninvasive methods for depicting tumor heterogeneity, and radiation therapy is generally prescribed assuming that the target volume is homogeneous (66). The use of radiomic features such as textural features could be of great interest to describe precisely cellular heterogeneity and guide radiation therapy treatment. Orlhac *et al.* compared autoradiography and PET images of tumors in mice to demonstrate that textural analysis reflects cellular arrangement (67). Further work is needed to better understand the relationship among cell biology, genomics, and tumor-based radiomics features computed from medical images (68).

Dose painting aims at delivering a heterogeneous dose pattern to boost areas suspected to be more radioresistant that could be identified using functional imaging (69). Numerous studies involving various histology and tumor localizations used PET imaging for dose painting, investigating the benefit of multitracer images (70-75). Two dose painting methods have been described in the literature:

1. Dose painting by numbers, in which the dose to each voxel is determined as a function of the functional image intensity;
2. Dose painting by contours, in which homogeneous doses are delivered in subregions of the tumor contoured semiautomatically using functional images and specific intensity thresholds (76).

The use of baseline PET-CT to define biological radiation therapy boost volumes has been investigated for pancreatic, prostate, head and neck, cervical, lung, and esophageal cancers (71, 77-82). Most ongoing clinical trials are focused on head and neck cancer (72, 83-87) and use a fixed SUV threshold or a manual delineation for determining subvolumes in which the dose has to be escalated (Table E2; available online at www.redjournal.org). Several clinical trials involving PET to guide dose planning are still recruiting. These clinical trials will be necessary to validate the clinical added value of PET-based dose painting, which has not yet been fully demonstrated. The first long-term analysis of patients with prostate cancer treated with a

PET-based simultaneous integrated boost recently showed a higher but not significant difference in biochemical recurrence-free survival (92% vs 85% without simultaneous integrated boost, $P = .17$) (88). No retrospective studies or ongoing clinical trials have used radiomic features other than SUV for dose planning. More sophisticated radiomic models involving additional features might enhance biological target volume and biologically guided dose painting. Furthermore, the potential added value of radiomic features calculated in images acquired during treatment for adaptive planning remains to be investigated.

Beyond ^{18}F-FDG

Even if FDG is the most common radiotracer used in PET imaging, there is increasing interest in other tracers in oncology, targeting different metabolic pathways or receptor expressions at preclinical or clinical stage (89-91).

Texture analysis studies involving tracers other than FDG are not common, but the concept developed for FDG can be extended to other tracers that are especially relevant in oncology. As an example, preliminary studies on ^{18}F-fluorothymidine images showed a high correlation between 2 histogram-based features and survival in patients with newly diagnosed glioma (92) or between entropy of the histogram and survival in patients with colorectal cancer (43). Moreover, ^{18}F-dihydroxyphenylalanine PET has been reported as a highly accurate tool for differentiating radiation necrosis from progressive brain metastases after stereotactic radiosurgery (93). Limitations of these studies include the low number of patients and the lack of assessment of the radiomic model on data that differed from the data used to design the model.

Prospective radiomic studies combining images acquired with different tracers could help the community to assess their added value compared to FDG only.

Challenges in PET Radiomic Pipeline

As illustrated in Figure 2, several sources of variability can be identified at each step of the radiomic pipeline. The fact that the robustness of radiomic models with respect to these sources of variability is almost never studied considerably limits the dissemination of radiomics to the clinical radiation therapy workflow.

Continuous technological advances in PET imaging

Since the first PET scanner was developed by Phelps et al in 1975 (94) and the first hybrid PET-CT scanner was developed in 2000 by Beyer et al (95, 96), tremendous advances have occurred in PET imaging with the advent of new detectors, computing power, and electronic miniaturization. PET imaging currently plays a key role in oncology, and image quality has been considerably improved in recent years. New detector geometry (now typically with 20 cm axial field of view [FOV] and 70 cm transaxial FOV), better detector performance, and more sophisticated image reconstruction algorithms have significantly increased PET sensitivity and enhanced spatial resolution (~4-5 mm around the FOV center) (97) and contrast-to-noise ratio with the use of time-of-flight (98), while reducing scanning time. The enhanced spatial resolution allows for smaller voxel size, resulting in possible finer assessment of tumor heterogeneity. Other developments regarding motion correction, metal artefact reduction on the CT component for more accurate PET attenuation correction, and continuous bed motion (99) are ongoing to further enhance routine image quality in the near future.

Although these improvements in PET image quality are of course beneficial, they result in large variability in PET image quality as a function of the scanner and acquisition and reconstruction protocols used in different departments. Indeed, improved spatial resolution yields more accurate tracer activity recoveries and hence produces significant changes in SUV (100-103). As an example, in lymph nodes, sophisticated reconstruction including a model of the detector response increased SUV_{max} by 43% (4.5-6.5) and SUV_{mean} by 28% (3.6-4.6) with respect to values measured when the detector response was not modeled ($P < .05$) (103). Such differences in SUV will affect dose planning based on PET if the boost region is defined based on an SUV threshold.

The variability of PET textural indices as a function of scanner and acquisition and reconstruction protocols can be assessed using phantom acquisitions and patient data and has been found to be highly feature dependent (104-109). A comparison of 5 studies is presented in Table 2, with a selection of common radiomic features (104-108). Some results for several features vary among studies because of the use of different radiotracers, PET scanners, and methods of calculation. Methods for characterizing feature variability were also different and used either coefficients of variation or percentages of difference. However, some trends can be extracted from this table. Entropy from the co-occurrence matrix is one of the most robust textural features, and voxel size is the most important source of variability influencing all parameters, independent of imaging modality (67, 110, 111). A limitation of these studies is the use of a single PET device in each study so that the variability across scanners could not be investigated.

The impact of respiratory motion on textural features was highlighted using a heterogeneous phantom (112) and patients with lung cancer (60, 113). In particular, long run low gray, busyness, maximal correlation coefficient, and coarseness were significantly different ($P < .05$) between 3D and 4D PET images in 26 patients with lung cancer (60), suggesting that respiratory motion blur significantly changes the textural feature values. Similarly, among 56 features extracted from 23 patients with lung cancer in another study, only 26% demonstrated less than 5% difference between 3D and 4D PET images (113). Comparison of textural features between 4D-PET phases shows, however, that similar values

Table 2 Robustness of selected textural features (TF), using coefficient of variation (COV_{TF}) or percent difference (%Diff) as a robustness classifier (−, +, and ++ for low variability, intermediate variability, and high variability, respectively) (104-108)

Matrix	Feature	Source of variability													
		All parameters		*Acquisition & Reconstruction*				*Post-filtering level*			*Iteration number*		*Matrix size*		
		Shiri et al.	Galavis et al.	Shiri et al.	Cortes-Rodicio et al.	Bailly et al.	Yan et al.	Shiri et al.	Bailly et al.	Yan et al.	Bailly et al.	Yan et al.	Shiri et al.	Bailly et al.	Yan et al.
Conventional	SUV_{mean}	-	++	-	-		-	-		-		-	+		-
	SUV_{max}	-	++	+	+		-	+		+		-	-		++
	SUV_{peak}	-		-	+		-	-		-		-	+		-
Histogram	SkewnessH	+	++	++	++		++	+		++		++	+		++
	KurtosisH	-	++	-	-		-	-		-		-	-		+
	EntropyH	-	-	-			-	-		-		-	-		-
	EnergyH		++				-			-		-			+
GLCM	Homogeneity	-	++	-	-	-	-	-	-	-	-	-	++	+	++
	Energy					-	-		-	-	-	-		++	++
	ContrastCM	-	++	+	++	++	-	+	-	+	+	-	++	++	++
	Correlation	-	++	-	-	++	-	+	+	+	++	+	-	++	++
	Entropy	-	+	-	-	-	-	-	-	-	-	-	-	-	+
	Dissimilarity	-		-	+	-	-	-	-	-	-	-	++	++	++
GLRLM	SRE	-	++	-	-		+	-		+		-	-		++
	LRE	-	++	-	-		+	-		+		-	++		++
	LGRE	++	-	++	++	++	-	++	+	-	++	-	++	++	-
	HGRE	-	+	-	-	+	-	+	+	-	+	-	++	+	-
	SRLGE	++	++	++	++		+	++		+		-	++		++
	SRHGE	-	++	-	+		+	+		+		-	++		++
	LRLGE	++	++	++	++		+	++		+		-	++		++
	LRHGE	+	++	+	+		+	++		+		-	++		++
	GLNUr		++				-			-		-			++
	RLNU		++		-		+			-		-			++
	RP	-	++	-	-	-	-	-	-	-	-	-	++	-	++
NGLDM	Coarseness	-	++	-	+		-	+		+		+	++		++
	ContrastDM	-	++	+	++		-	-		+		-	++		++
	Busyness	+	++	+	+		-	+		+		+	++		++
GLSZM	SZE	-		-	-		+	-		++		-	-		+
	LZE	+		+	-		+	+		+		-	++		+
	LGZE	+		++		++	-	++	+	-	++	-	++	++	-
	HGZE	-		-		+	-	+	+	-	+	-	+	+	+
	SZLGE	++		++	++	++	+	++	++	+	++	-	++	++	+
	SZHGE	-		-	+	+	+	-	+	+	+	-	-	+	++
	LZLGE	++		++	++		+	++		++		-	++		+
	LZHGE	+		+	+		+	++		++		+	++		+
	GLNUz						-			-		-			++
	ZLNU				-	+	+		+	-	+	-		++	++
	ZP	-		-	-	-	++	-	-	++	-	++	++	+	++

For Bailly et al (105): high = with COV_{TF} not statistically different from $COV_{SUVmean}$ and COV_{TF} statistically different from COV_{SUVmax} with $COV_{TF} < COV_{SUVmax}$; intermediate = COV_{TF} not statistically different from COV_{SUVmax}; low = COV_{TF} statistically different from COV_{SUVmax} with $COV_{TF} > COV_{SUVmax}$. For Yan et al. (107), Shiri et al (108), and Cortes-Rodicio et al. (104): high = $COV_{TF} > 20\%$; intermediate = $10\% < COV_{TF} \leq 20\%$; low = $COV_{TF} \leq 10\%$. For Galavis et al. (106): high = $\%Diff \geq 30\%$; intermediate = $10\% \leq \%Diff \leq 25\%$; low = $\%Diff \leq 5\%$ with $\%Diff = 100 \times (TF - TF_{mean})/TF_{mean}$ where TF corresponds to texture feature value for each reconstructed image, and TF_{mean} is the average value.

could be obtained from all phases (60, 113, 114). Half of the features computed by Oliver et al (113) showed less than 5% variability among 10 phases of respiratory-gated PET images.

The impact of noise on image features has also been investigated in ungated and respiratory-gated PET images of patients with lung cancer (115). In this study, Gaussian noise of varying standard deviations was added to images, and 81 radiomic features were calculated. Features computed from GLSZM were the most sensitive to noise (90% average difference between feature values extracted from ungated and respiratory-gated images). Difference increased with increases in noise level for intensity and GLCM features, highlighting the strong effect of noise on most radiomic features.

Because of variations in device technology and image quality, most previously published clinical radiomic studies have been performed using a single scanner and a single protocol to avoid biases. This approach therefore does not provide any guarantee regarding the validity of the results in different settings, and it limits the number of patients included in each study. In 2 studies (42, 55), 2 scanners were used. In one study (42), the radiomic model was designed based on images acquired on a single PET/CT scanner and then was evaluated by using images acquired from a different scanner, thus testing the robustness of the model, which was less performant for predicting cervical cancer recurrence on the second scanner. Oh et al (55) applied a smoothing kernel and spatial resampling to PET reconstructed images. The kernel characteristics were determined so that the recovery coefficients deduced from American College of Radiology phantom images were matched between the 2 devices, suggesting that the smoothed images from the 2 scanners had similar spatial resolution.

The issue of radiomic model robustness with respect to different scanners was highlighted in a recent study in which a signature for predicting the locoregional recurrence in cervical cancer was derived based on textural features extracted from images acquired using 2 different PET scanners (Fig. 5A) (42). In a healthy liver region, 4 of 9 radiomic features were found to be significantly different ($P < .05$ for SUV_{max}, SUV_{peak}, homogeneity and entropy from GLCM) between the 2 image sets associated with the 2 scanners, demonstrating a scanner effect. In univariate analysis, a similar trend in textural feature values was observed for images coming from the 2 scanners (eg, higher entropy values for relapsing patients than for nonrelapsing patients; Fig. 5D). However, all entropy values were higher in the most recent PET scanner equipped with time-of-flight than in the oldest scanner, leading to a very different cutoff value for predicting recurrence in images from the 2 scanners. The 4-feature signature predicting tumor recurrence derived from images from the most recent PET scanner with good discrimination performance (area under the curve [AUC] = 0.86) performed more poorly when used on images from the oldest scanner (AUC = 0.76) (Fig. 5E). As illustrated using this example, a radiomic model derived from data acquired in certain conditions can thus not be straightforwardly extended to data acquired differently, which considerably limits the dissemination of radiomic models. Harmonization of data is absolutely needed so that radiomic models can be disseminated.

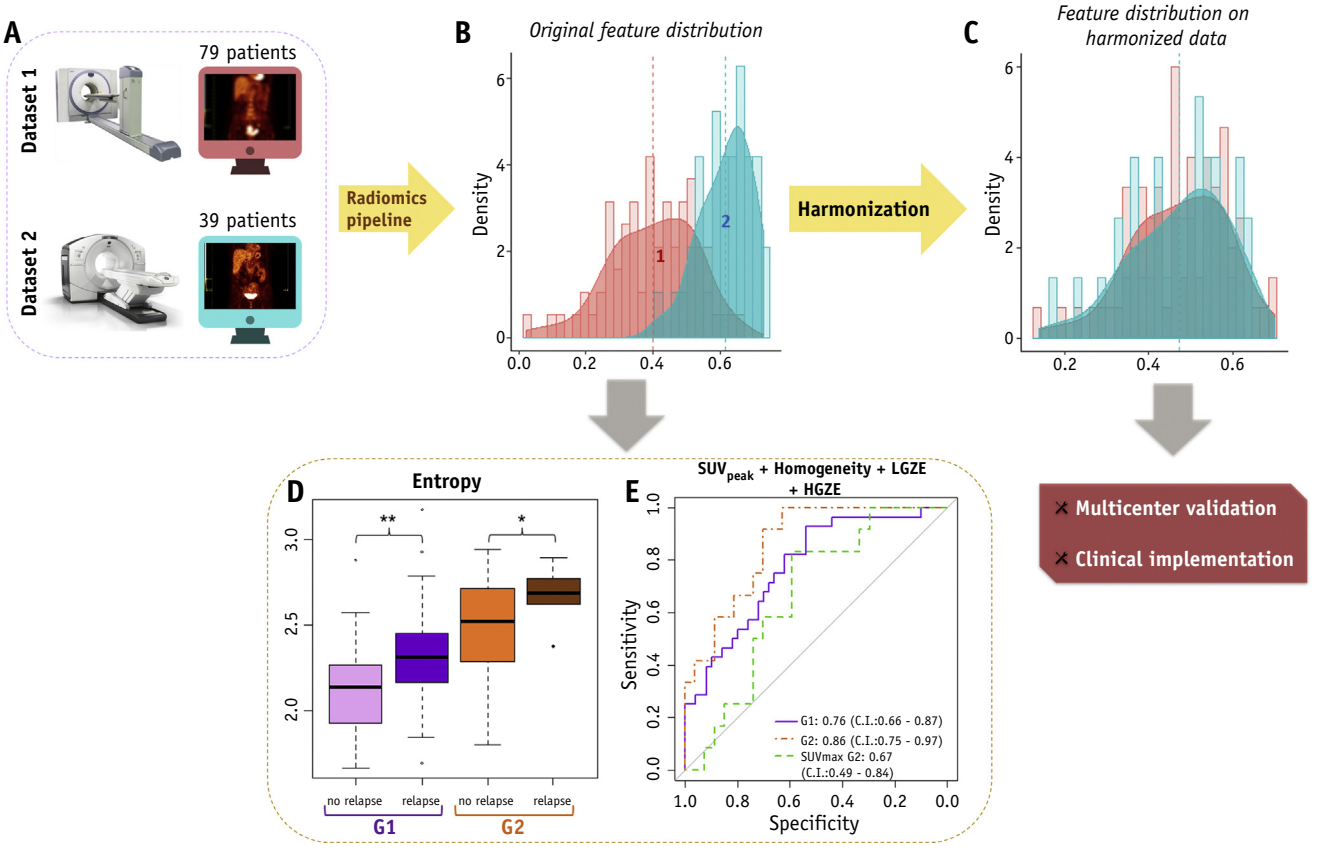

Fig. 5. Multiscanner radiomic study in patients with cervical cancer (42). Two patient cohorts were retrospectively included and gathered depending on the PET scanner used for image acquisition (A). Distribution of features were significantly different between groups, highlighting the fact that feature values greatly depend on the PET device, and making a predictive model of recurrence difficult to export (B). Using a statistical method of feature standardization (ComBat), distributions are similar, and multicenter validation could be possible (C).

The issue of varying values as a function of the imaging device and protocol was identified before radiomics development and was already a concern in interpreting SUV in FDG PET. Several organizations have thus developed standardization procedures for more reproducible quantification in PET for clinical trials (American College of Radiology, European Association of Nuclear Medicine, Society of Nuclear Medicine and Molecular Imaging−Clinical Trials Network). These procedures are all based on phantom acquisitions and aim at reducing multicenter variability in spite of different image acquisition and reconstruction protocols. Even if harmonization programs give some indication regarding patient preparation (116), discrepancies are still present between centers in terms of biases related to the patient's declared weight (117), the risk of deviation in postinjection uptake time (118−120), or the optimal injected activity (121).

Variations associated with reconstruction protocols can be reduced using postfiltering steps (122, 123) so that all images come close to a target spatial resolution value. A drawback of this approach is that images with the highest initial resolution see their resolution degraded, which is especially adverse for an analysis of image texture (124).

A harmonization method applicable a posteriori to the extracted features has been recently proposed (125). This method derived from genomic data analysis and called ComBat can identify a categorical confounding variable such as the model of the PET scanner and estimate the feature values devoid of this confounding variable effect. The method has been validated in healthy liver tissue and in breast cancer tumors; none of the 9 radiomic feature distributions that were investigated were significantly different between 2 nuclear medicine departments after the harmonization process ($P < .0001$ for more than 4 features before harmonization vs $P > .1$ for harmonized features). The accuracy of identification of triple-negative lesions based on radiomic feature cut-offs derived from the dataset of 1 nuclear medicine department was significantly enhanced when applied on another dataset after harmonization. When applying this approach to the data of patients with cervical cancer coming from the 2 PET scanners shown in Figure 5, the method was also quite effective (Figs. 5B and 5C). Using this harmonization method, retrospective analyses gathering data from different PET centers becomes feasible, and the implementation of radiomic analysis in a radiation therapy workflow is conceivable.

Effects of postprocessing pipeline on extracted values

Delineation method

As PET images are now considered useful to improve target volume delineation in radiation oncology and to define regions in which the dose could be escalated, a consensus on PET delineation is crucial. Automatic contouring methods have been described as inappropriate to distinguish between high uptakes associated with tumor tissues and high physiological uptake or uptake due to inflammation (126). Many semiautomatic segmentation methods have been described, and an overview of the available segmentation methods is given in Table E3 (available online at www.redjournal.org) (127-139). A comprehensive description of segmentation methods used in PET imaging has recently been published in a report of the AAPM (American Association of Physicists in Medicine) task group (140). Advantages, drawbacks, and impacts of these methods on the resulting volume of interest (VOI) have already been widely discussed in the literature (126, 140-149), and the overall conclusion is that there is no such thing as a unique method that performs best in all circumstances. The method to be chosen highly depends on the method availability, ease of parameterization, image features, tumor type, and presence of neighboring structures with high uptake. SUV thresholding methods have been shown to have a significant impact on dose painting in patients with lung cancer (150). For dose painting by contours, the variability in dose prescription between reconstruction methods was lower when target volumes were delineated using an SUV_{peak}-based threshold than with SUV_{max}-based thresholding (150).

In the context of radiomics for outcome prediction, what matters is the difference in radiomic features as a function of the delineation method used. Indeed, for a radiomic model to be widely applicable, the radiomic features on which is based should be robust to segmentation, or all images should be segmented with the same method after ensuring that images are of similar quality in terms of spatial resolution and signal to noise ratio, which is currently quite difficult to achieve. If not available in the literature, an analysis of the robustness of radiomic features should be performed for each tumor type and segmentation method when user input is required.

The impact of segmentation method on radiomic feature values has been discussed in the literature. It has been reported that only 5 textural features (Entropy, SRE: short run emphasis, RP: run percentage, SZE: short zone emphasis, LRE: long run emphasis) were less dependent on segmentation method than SUV_{mean} with $sd_{\delta}(feature) < sd_{\delta}(SUV_{mean})$, and 12 features were highly dependent on the segmentation method, with $sd_{\delta}(feature) > sd_{\delta}(MTV)$ (151). SUV_{max} thresholding at 40% was shown to be the most robust segmentation method in patients with non-small cell lung cancer, considering interobserver variability (compared to fuzzy locally adaptive bayesian (FLAB) and free-hand segmentation) and the association of radiomic features with patient survival (146). The impact of threshold value has been investigated in several publications, showing substantial differences in stratification power (22, 147). Using phantom images, the FLAB method has been shown to be more accurate than thresholding compared with ground-truth extracted from CT images because of the process of probabilistic classification for the delineation of small or

highly heterogeneous lesions (137, 152). The AUC values for response prediction were higher for most radiomic features calculated on volumes delineated using FLAB compared with a fixed 42% threshold. The difference was statistically significant for homogeneity, dissimilarity, high intensity emphasis, and the AUC of the cumulative histogram (153). In a study including 88 patients with cervical cancer, Altazi et al compared manual segmentation by 2 observers with a semiautomated graphical-based method (154). Only 10 of 79 features were reported as highly reproducible among segmentation methods in this study. However, cervical cancer segmentation is challenging because of the presence of adjacent organs with a very high uptake, such as the bladder, and the rectum.

To overcome the problem of segmentation method choice, Foley et al performed a radiomic analysis on esophageal cancer (155) using an automatic decision-tree-based pipeline for segmentation (156). The advantage of this method is that the best-fitting automatic segmentation method is selected for each individual case among various methods (eg, adaptive iterative thresholding, region-growing, k-means iterative clustering, fuzzy-c-means iterative clustering, Gaussian mixture models–based clustering, watershed transform–based algorithm using Sobel filters). This strategy could be of interest for reducing variability induced by the segmentation step.

Method of discretization

To reduce the impact of noise and the size of textural matrices, voxels of similar intensities are gathered in a defined number of gray levels (Fig. 6C) before second-order feature calculation (Fig. 6B) (157, 158). SUV values can be discretized with a fixed bin size between fixed bounds (absolute method AD, Eq. 1, Fig. 6B) or with a fixed number of bins between SUV_{min} and SUV_{max} of the VOI (relative method RD, Eq. 2 [158], Fig. 6B or Eq. 3 [37, 38]). Even if these methods are the most commonly used, other methods such as the Max-Lloyd clustering algorithm (159) or histogram equalization (114) have also been described in the literature.

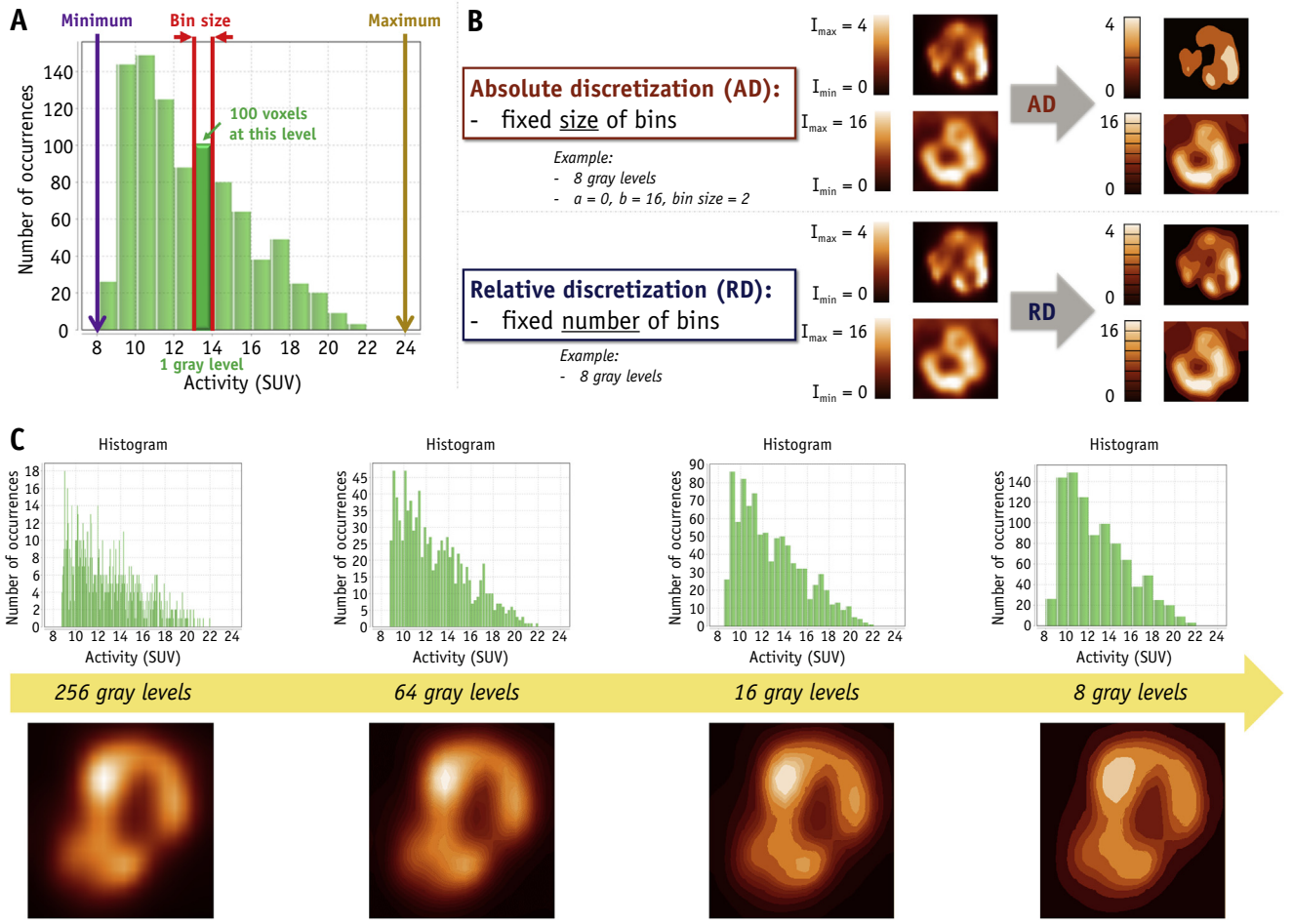

Fig. 6. Intensity discretization. (A) Histogram of intensities displays the minimum, maximum, bin size, and gray level quantities. (B) Two methods of gray level discretization are shown, absolute discretization (AD: fixed bin size between fixed bounds) and relative discretization (RD: fixed bin number between minimum and maximum intensity in the tumor). (C) Impact of gray level number on a real patient (with cervical cancer), between 8 and 256 levels.

$$AD(x) = \text{round}\left(D \times \frac{I(x) - a}{b - a}\right) \quad \text{(Eq.1)}$$

$$RD_1(x) = \text{round}\left(D \times \frac{I(x) - SUV_{min}}{SUV_{max} - SUV_{min}}\right) \quad \text{(Eq.2)}$$

$$RD_2(x) = \text{round}\left(D \times \frac{I(x) - SUV_{min}}{SUV_{max} - SUV_{min} + 1}\right) \quad \text{(Eq.3)}$$

In Equations 1, 2 and 3, *AD(x)* and *RD(x)* are the resampled intensity values corresponding to voxel *x*, I(x) is the original value in voxel *x* before resampling, *D* is the number of discrete values (also called bins), and *(a, b)* are the minimum and maximum bounds in SUV units for absolute discretization.

Robustness of radiomic features as a function of the discretization method has been investigated, and feature values can be compared between studies only if the same discretization process was used (67, 151, 157, 158, 160). It has been shown that many PET textural features are highly correlated to MTV with relative discretization (153, 158). Despite a lower correlation with MTV, a higher correlation with SUV_{max} has been reported for AD (157). Absolute discretization has been shown as more appropriate in clinical cases (157), and this can be intuitively understood. With RD, a tumor in which SUV varies with a Gaussian distribution between 2 and 4 will see its voxel values discretized similarly to a tumor in which SUV varies with a Gaussian distribution between 2 and 16; it is likely that the later tumor is more heterogeneous than the former because it includes voxels with SUV = 2 and voxels with SUV = 16 (Fig. 6A). The bounds of absolute discretization should be carefully chosen to encompass all SUV values encountered in the entire cohort and should be set to a value greater than the maximum SUV_{max} observed over all VOIs (67). For Leijenaar et al (157), the optimal value of intensity resolution (bin width) could not be identified because textural features computed with different bin sizes could carry complementary information. Absolute discretization with a fixed bin size of 0.3 SUV unit leading to 64 discrete values between 0 and 20 SUV (158, 160) or 128 values between 0 and 40 SUV because of an $SUV_{max} > 20$ in several VOI has been proposed (42, 161).

All these results emphasize the need for precise description of the method used so that a study design can be replicated using different data. As for segmentation, each discretization method has advantages and drawbacks and can lead to substantially different results (158, 162).

Model building and validation

Usually, each patient (ie, each tumor in most cases) is characterized by a vector of radiomic feature values. Most radiomic studies aim at establishing a link with a clinical parameter, which can be categorical (eg, stage, cancer subtype, relapse) or continuous (eg, censored like survival, risk score), as a decision support system, using statistical learning approaches (Figs. 2D and 7).

Statistical learning methods are either supervised or unsupervised. Supervised methods learn to predict an explicit variable from labeled data (called a "training set" or "discovery set") using optimization (classification tasks for categorical variables, regression tasks for continuous variables) (Fig. 5). They are based on optimization procedures that tend to minimize the prediction error (loss function). The model then is evaluated on new data (called the "validation set"), which ideally is independent of the

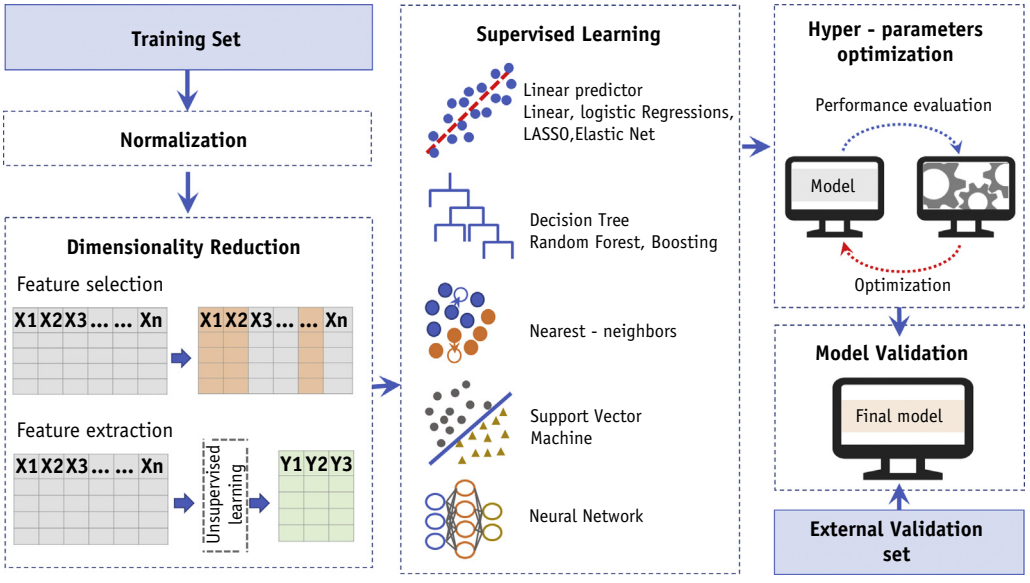

Fig. 7. Model building and validation. After feature normalization and dimensionality reduction, different methods of supervised learning can be used to establish a model. This model must be tested on a second set, ideally acquired in an external center with another scanner, to be clinically validated.

training set to ensure the generalizability of the model ("external validation set") (163). In the current literature, validation is more frequently performed on a subset of the whole cohort, as in 2 of the studies presented in Table 1 (42, 46). In 93% of the studies, external validation was not performed. Results therefore should be interpreted with caution because the performance of the model on independent data remains unknown.

There is a wide range of supervised learning methods. These include, for example, linear predictors (e, linear or logistic regressions) (164), decision tree algorithms (eg, random forests) (50, 51), and support vector machine (165, 166). For time-to-event data (eg, overall survival or disease-free survival), dedicated algorithms can deal with continuous censored data (ie, Cox proportional hazards regressions, random survival forest), allowing for a patient stratification based on the predicted risk.

As opposed to supervised learning methods, unsupervised methods do not use a predefined output variable. They are used to obtain a new representation of the data, either by grouping similar data together (clustering) (167, 168) or by finding a lower-dimensional representation of the data (dimensionality reduction).

In radiomic studies, as in other omics fields (eg, genomics, proteomics), the number of radiomic features often far exceeds the number of patients. This situation leads to a high probability of overfitting, resulting in a perfect model derived from the learning data but without any ability to extend to other datasets. This risk can be estimated by cross-validation methods during the learning step and controlled by dimension reduction methods (164). Simpler models involving fewer radiomic features are therefore more robust, and reducing dimensionality allows for reducing computational complexity.

The first approach to reduce the number of features, also called feature selection methods, retains only a subset of the initial variables, ranked according to a performance criterion, to discard irrelevant and redundant variables. This can be carried out using 3 approaches:

1. "Filter" method: Relevant variables are selected in a first preprocessing step, independently of the learning method, depending on their relevance, redundancy, or robustness over time (169, 170).

2. "Wrapper" method: An approach similar to the filter method, but the performance of the variables is quantified depending on the learning algorithm subsequently used, potentially leading to a different selection if another learning method is chosen (171).

3. "Embedded" method: The variable selection and the learning algorithm are combined into a single mathematical problem (172).

Statistical procedures can also be used to reduce dimension, starting from all variables available in the initial database and summarizing them into a smaller number of new variables using unsupervised learning. These new, constructed variables combine information from all initial features but are not interpretable on their own.

Another important concept in statistical learning is feature manipulation or normalization, which corresponds to simple transformations applied to the original data. For instance, standardization makes all features have a zero mean and unit variance. This allows values of different scales to be adjusted to a common scale to avoid overweighting one variable over another when the model is sensitive to magnitude, as for regression models. Feature normalization can also help a learning algorithm to converge faster.

In PET, there is no consensus on which feature selection or learning methods should be used (Table 2). Given the state of the art, an analysis of the impact of several learning approaches on the stability and robustness of the proposed models is needed (164, 173).

Recently, deep learning has emerged as a new, accessible area of machine learning with the widespread availability of fast computers, although it is not yet used in the specific field of PET radiomics, as shown in Table 1. This technique is typically based on neural networks that include many layers for learning multiple levels of hierarchical representations, using supervised and unsupervised feature extraction and transformations (174). These models can be difficult to train and optimize because a classical network can produce tens of thousands of features (175). A large amount of data is therefore recommended, although some promising methods of computer vision, such as data augmentation and transfer learning, may help to solve the problem of small datasets (176).

Interpretation of radiomic feature values

Except for conventional indices (eg, SUV, MTV), most radiomic features are dimensionless quantities. Examples of distributions of 3 features (SkewnessH from the histogram, entropy from GLCM, and high gray-level zone emphasis (HGZE) from GLSZM) computed in cervix tumors are shown in Figure 4A, with some extreme cases to illustrate what they reflect. In addition, a selection of common radiomic feature values is presented in Figure 4B for a patient case. The relationship between the activity pattern shown in the image and textural feature values has been investigated (177) using patient data and simulations. From 6 textural features, 2 groups were identified: (1) homogeneity, entropy, SRE, and LRE reflect uptake heterogeneity without distinguishing between hypersignal and hyposignal within a uniform distribution; and (2) HGZE and LGZE correlate with the average uptake rather than local heterogeneity. Another study compared a visual assessment of tumor heterogeneity as seen on PET images with textural features for non-small cell lung cancer (178). A 3-level scale of heterogeneity was used by 2 nuclear

medicine physicians to score the visual heterogeneity of the tumors. Moderate correlations ($0.4 <$ Spearman's rank coefficient < 0.6) were observed between visual scores and textural features, and an added value was shown for textural feature analysis over visual assessment for overall survival prediction. This is a first step toward the definition of guidelines to help physicians in interpreting textural feature values in clinical practice. Another important step will be the determination of expected textural index values for nonpathologic organs after acquisition, reconstruction, or feature extraction standardization.

Perspectives

Multimodal imaging, radiation therapy, and radiomics

Although PET-CT images and radiation therapy planning CT can be merged, registration errors are frequent, and several studies have been published in favor of performing PET-CT acquisitions in treatment position for radiation therapy planning, with recommendations and workflow implications (179-181). With the recent advent of PET-MRI scanners and despite issues related to the lack of directly available electronic density maps from PET-MRI scanners (182), several promising studies have also focused on the implementation of PET-MRI for radiation therapy planning (183-185). Deformable image registration is also a promising option to integrate multimodal imaging in biologically guided radiation therapy (186).

The accurate fusion of anatomic and molecular imaging is also a great added value for outcome prediction. Features may be extracted separately from functional and anatomic images, using modality-specific or the same contours, and then combined in statistical analyses. Vaidya et al (33) reported that a model combining 2 features from PET and CT performed better than models using a single PET or a single CT feature for predicting radiation therapy tumor response in lung cancer ($\rho = 0.49$, $P = .007$ for locoregional failure with a 2-feature PET+CT model vs $\rho = 0.34$, $P = .04$ for the best PET feature and $\rho = -0.27$, $P = .09$ for the best CT feature). Beukinga et al (45) reported that a model combining PET and CT features in addition to clinical parameters was more accurate than response prediction based only on the SUV_{max} for prediction of chemoradiation therapy response in esophageal cancer (AUC = 0.74 for the multivariate signature vs 0.54 for SUV_{max} only). Lartizien et al (166) evaluated the performance of different classifiers based on either multimodal models combining PET and CT radiomic features or PET features only in staging lymphoma. The best performance was achieved by the multimodal support vector machine model trained on a

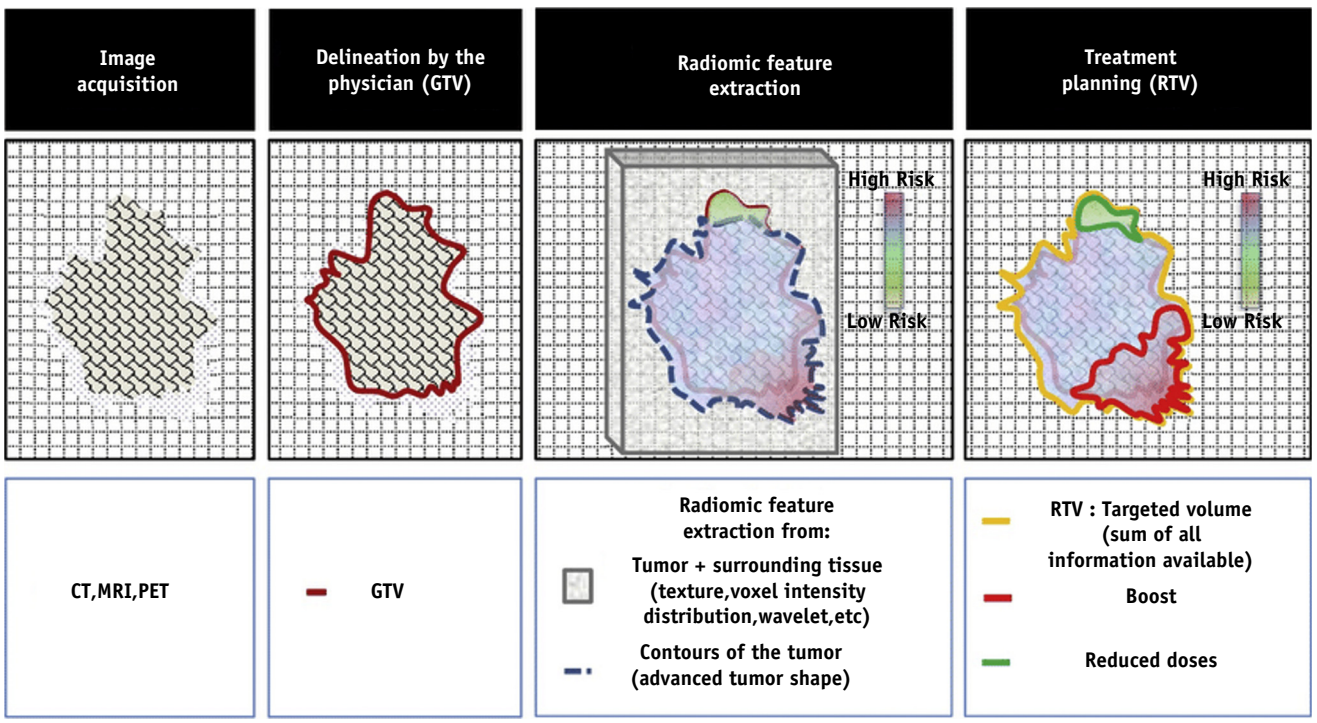

Fig. 8. Determination of the radiomic target volume (193). Conventionally, gross tumor volume is segmented by the physician from PET, MRI, and CT images. This step being time-consuming, textural analysis can be seen as a useful tool to automatically determine a radiomic target volume. From this volume, high-risk and low-risk regions can be identified via machine learning algorithms to allow dose painting in radiation therapy. *Abbreviations:* CT = computed tomography; PET = positron-emission tomography.

combination of 12 first- and second-order features from both PET and CT (AUC = 0.91 vs 0.89 for PET features only). Other approaches such as nomograms have also been evaluated in this multimodal context, showing the complementarity of PET entropy and CT zone percentage with functional volume for lung cancer staging (187). Another multimodal study was performed on cervical cancer, gathering features computed from PET and MRI (188). Lucia et al proved that entropy from diffusion weighted imaging (DWI) MRI and gray-level non-uniformity (GLNU) from PET images were independent predictors of recurrence (HR = 34.18 and 5.39, $P < .00001$, respectively) and locoregional control (HR = 28.10 and 33.11, $P < .00001$ respectively), with significantly higher prognostic power than clinical parameters.

Despite most retrospective studies showing better results when combining PET and CT or MRI in radiomic models (33, 159, 166, 188-192), further work remains necessary to assess the clinical added value of multimodality in radiomics models using larger cohorts and external validation data.

Dose painting using PET radiomics

An evolution of biological target volume is the concept of radiomic target volume (193). From an automatic segmentation validated by the physician, local radiomic features can be automatically calculated to produce parametric maps of radiomic features in a box encompassing the segmented volume (Fig. 8). Based on radiomic feature values computed in each voxel, it might be possible to identify high-risk and low-risk regions in the tumor. Treatment planning could then be modulated as a function of the risk, with increased dose to high-risk volumes and lower dose to low-risk volumes to limit toxicities. Yet, the applicability of such a concept requires correlating parametric feature maps with tumor aggressiveness and addressing the limited spatial resolution of PET imaging (~4 to ~8 mm as a function of the image reconstruction algorithm setting). Prospective validations will also be necessary to assess the added value of such dose painting on tumor control. This concept was also described by Soufi et al, who proposed a pipeline of automated segmentation and labeling of heterogeneous subvolumes in lung tumors using textural features on FDG-PET imaging (194).

Preliminary work in SPECT imaging

There are few reports on the use of SPECT imaging for radiation therapy (195-197). In recent years, limitations that could confound the ability of SPECT to provide quantitative information (photon attenuation, scatter, partial volume effect, and motion artefacts) have been addressed through the use of SPECT/CT acquisitions and quantitative image reconstruction (196), making SPECT images quantitatively reliable. For instance, ^{99m}Tc-mebrofenin-SPECT was evaluated in dose planning adaptation for liver cancer or metastases to protect healthy hepatic tissue (195). SPECT may also be used to improve dose planning for thoracic lesions, depending on pulmonary function (197), or to develop predictive dose response models such as for kidney preservation during abdominal radiation therapy (198). To our knowledge, the only large SPECT radiomic study included 141 patients and pertained to neurodegenerative disease, monitoring the progression of Parkinson syndrome (199). This study concluded that imaging textural features could be promising for follow-up of patients with Parkinson disease, correlated with neurocognitive scores.

Conclusion

Radiomics is a promising domain to assist physicians in decision-making and to implement personalized radiation therapy with individualized doses and volumes. Physicians and physicists should pay attention to the possible confounding factors along the radiomic pipeline. From image acquisition to statistical analysis and implementation for radiation therapy treatment planning, there is still a lack of consensus. In the era of big data and machine learning analysis, there is also a crucial need for standardization, from written guidelines to closer collaboration among image-processing researchers, physicians, physicists, and data scientists. From this perspective, approaches from the Quantitative Imaging Network in the United States (200) or the German National Cohort Consortium (201) aim to gather researchers in quantitative imaging to standardize practices and facilitate clinical validation. Methodology should be critically reviewed, with the inclusion of essential elements in all publications so that results can be reproduced. Precise descriptions of the cohort, acquisition parameters, method of computation, and statistical pipeline are absolutely necessary to replicate studies. Validation of the radiomic workflow on an independent dataset should be systematically required to allow for the identification of robust and clinically useful radiomic models in radiation therapy. The next step toward implementation of radiomics in healthcare is the development of randomized clinical trials assessing the added value of such quantitative imaging features in prospective cohorts.

References

1. Limkin EJ, Sun R, Dercle L, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol* 2017;28:1191-1206.
2. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;48:441-446.
3. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. *Radiology* 2016;278:563-577.
4. O'Connor JPB, Aboagye EO, Adams JE, et al. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol* 2016;14:169-186.

5. Gillies RJ, Anderson AR, Gatenby RA, et al. The biology underlying molecular imaging in oncology: From genome to anatome and back again. *Clin Radiol* 2010;65:517-521.
6. Jin J-Y, Kong F-MS. Personalized radiation therapy (PRT) for lung cancer. *Adv Exp Med Biol* 2016;890:175-202.
7. Farwell MD, Pryma DA, Mankoff DA. PET/CT imaging in cancer: Current applications and future directions. *Cancer* 2014;120:3433-3445.
8. Gill BS, Pai SS, McKenzie S, et al. Utility of PET for radiotherapy treatment planning. *PET Clin* 2015;10:541-554.
9. Calvo FA, Sole CV, de la Mata D, et al. ^{18}F-FDG PET/CT-based treatment response evaluation in locally advanced rectal cancer: A prospective validation of long-term outcomes. *Eur J Nucl Med Mol Imaging* 2013;40:657-667.
10. Bahce I, Vos CG, Dickhoff C, et al. Metabolic activity measured by FDG PET predicts pathological response in locally advanced superior sulcus NSCLC. *Lung Cancer Amst Neth* 2014;85:205-212.
11. Higgins KA, Hoang JK, Roach MC, et al. Analysis of pretreatment FDG-PET SUV parameters in head-and-neck cancer: Tumor SUV_{mean} has superior prognostic value. *Int J Radiat Oncol Biol Phys* 2012;82:548-553.
12. Hu S-L, Yang Z-Y, Zhou Z-R, et al. Role of $SUV_{(max)}$ obtained by 18F-FDG PET/CT in patients with a solitary pancreatic lesion: Predicting malignant potential and proliferation. *Nucl Med Commun* 2013;34:533-539.
13. Baba S, Isoda T, Maruoka Y, et al. Diagnostic and prognostic value of pretreatment SUV in 18F-FDG/PET in breast cancer: Comparison with apparent diffusion coefficient from diffusion-weighted MR imaging. *J Nucl Med Off Publ Soc Nucl Med* 2014;55:736-742.
14. Khiewvan B, Macapinlac HA, Lev D, et al. The value of ^{18}F-FDG PET/CT in the management of malignant peripheral nerve sheath tumors. *Eur J Nucl Med Mol Imaging* 2014;41:1756-1766.
15. Vargas HA, Burger IA, Goldman DA, et al. Volume-based quantitative FDG PET/CT metrics and their association with optimal debulking and progression-free survival in patients with recurrent ovarian cancer undergoing secondary cytoreductive surgery. *Eur Radiol* 2015;25:3348-3353.
16. Klabatsa A, Chicklore S, Barrington SF, et al. The association of ^{18}F-FDG PET/CT parameters with survival in malignant pleural mesothelioma. *Eur J Nucl Med Mol Imaging* 2014;41:276-282.
17. Foley KG, Fielding P, Lewis WG, et al. Prognostic significance of novel ^{18}F-FDG PET/CT defined tumour variables in patients with oesophageal cancer. *Eur J Radiol* 2014;83:1069-1073.
18. Herrera FG, Breuneval T, Prior JO, et al. [(18)F]FDG-PET/CT metabolic parameters as useful prognostic factors in cervical cancer patients treated with chemo-radiotherapy. *Radiat Oncol Lond Engl* 2016;11:43.
19. Gomez CR. Editorial: Tumor hypoxia: Impact in tumorigenesis, diagnosis, prognosis, and therapeutics. *Front Oncol* 2016;6:229.
20. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern* 1973;610-621; SMC-3.
21. El Naqa I, Grigsby P, Apte A, et al. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. *Pattern Recognit* 2009;42:1162-1171.
22. Hatt M, Laurent B, Fayad H, et al. Tumour functional sphericity from PET images: Prognostic value in NSCLC and impact of delineation method. *Eur J Nucl Med Mol Imaging* 2018;45:630-641.
23. Sollini M, Cozzi L, Antunovic L, et al. PET radiomics in NSCLC: State of the art and a proposal for harmonization of methodology. *Sci Rep* 2017;7:358.
24. Heukelom J, Hamming O, Bartelink H, et al. Adaptive and innovative Radiation Treatment FOR improving Cancer treatment outcomE (ARTFORCE); a randomized controlled phase II trial for individualized treatment of head and neck cancer. *BMC Cancer* 2013;13:84.
25. Thie JA. Understanding the standardized uptake value, its methods, and implications for usage. *J Nucl Med Off Publ Soc Nucl Med* 2004;45:1431-1434.
26. Dercle L, Ammari S, Bateson M, et al. Limits of radiomic-based entropy as a surrogate of tumor heterogeneity: ROI-area, acquisition protocol and tissue site exert substantial influence. *Sci Rep* 2017;7:7952.
27. Mallat SG. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans Pattern Anal Mach Intell* 1989;11:674-693.
28. Laws KI. Rapid Texture Identification, Proc. SPIE 0238, Image Processing for Missile Guidance, (23 December 1980); https://doi.org/10.1117/12.959169.
29. Xu D-H, Kurani AS, Furst JD, et al. Run-length encoding for volumetric texture. *4th IASTED Int Conf Vis Imaging Image Process VIP* 2004;27.
30. Thibault G, Fertil B, Navarro C, et al. Texture indexes and gray level size zone matrix. Application to cell nuclei classification. 10th International Conference on Pattern Recognition and Information Processing, PRIP 2009, 2009. Minsk: Belarus; 2009; pp.140-145.
31. Amadasun M, King R. Textural features corresponding to textural properties. *IEEE Trans Syst Man Cybern* 1989;19:1264-1274.
32. Hatt M, Majdoub M, Vallières M, 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.
33. Vaidya M, Creach KM, Frye J, et al. Combined PET/CT image characteristics for radiotherapy tumor response in lung cancer. *Radiother Oncol J Eur Soc Ther Radiol Oncol* 2012;102:239-245.
34. Bundschuh RA, Dinges J, Neumann L, et al. Textural parameters of tumor heterogeneity in ^{18}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;55:891-897.
35. Yu H, Caldwell C, Mah K, et al. Coregistered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning. *IEEE Trans Med Imaging* 2009;28:374-383.
36. Cook GJR, Yip C, Siddique M, et al. Are pretreatment 18F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? *J Nucl Med Off Publ Soc Nucl Med* 2013;54:19-26.
37. Tixier F, Le Rest CC, Hatt M, et al. Intratumor heterogeneity characterized by textural features on baseline ^{18}F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med Off Publ Soc Nucl Med* 2011;52:369-378.
38. Yang F, Thomas MA, Dehdashti F, et al. Temporal analysis of intratumoral metabolic heterogeneity characterized by textural features in cervical cancer. *Eur J Nucl Med Mol Imaging* 2013;40:716-727.
39. Dong X, Sun X, Sun L, et al. Early change in metabolic tumor heterogeneity during chemoradiotherapy and its prognostic value for patients with locally advanced non-small cell lung cancer. *PloS One* 2016;11:e0157836.
40. Yue Y, Osipov A, Fraass B, et al. Identifying prognostic intratumor heterogeneity using pre- and post-radiotherapy 18F-FDG PET images for pancreatic cancer patients. *J Gastrointest Oncol* 2017;8:127-138.
41. van Dijk LV, Noordzij W, Brouwer CL, et al. 18F-FDG PET image biomarkers improve prediction of late radiation-induced xerostomia. *Radiother Oncol* 2018;126:89-95.
42. Reuzé S, Orlhac F, Chargari C, et al. Prediction of cervical cancer recurrence using textural features extracted from ^{18}F-FDG PET images acquired with different scanners. *Oncotarget* 2017.
43. Nakajo M, Kajiya Y, Tani A, et al. A pilot study for texture analysis of (18)F-FDG and (18)F-FLT-PET/CT to predict tumor recurrence of patients with colorectal cancer who received surgery. *Eur J Nucl Med Mol Imaging* 2017;44:2158-2168.
44. Chen S-W, Shen W-C, Lin Y-C, et al. Correlation of pretreatment 18F-FDG PET tumor textural features with gene expression in pharyngeal cancer and implications for radiotherapy-based treatment outcomes. *Eur J Nucl Med Mol Imaging* 2017;44:567-580.

45. Beukinga RJ, Hulshoff JB, van Dijk LV, et al. Predicting response to neoadjuvant chemoradiotherapy in esophageal cancer with textural features derived from pretreatment 18F-FDG PET/CT imaging. *J Nucl Med Off Publ Soc Nucl Med* 2017;58:723-729.
46. Kirienko M, Cozzi L, Antunovic L, et al. Prediction of disease-free survival by the PET/CT radiomic signature in non-small cell lung cancer patients undergoing surgery. *Eur J Nucl Med Mol Imaging* 2018;45:207-217.
47. Buvat I, Orlhac F, Soussan M. Tumor. Texture Analysis in PET: Where Do We Stand? *J Nucl Med* 2015;56:1642-1644.
48. Zwanenburg A, Leger S, Vallières M, et al. Image biomarker standardisation initiative-feature definitions. Available at: https://arxiv.org/abs/1612.07003. Accessed April 15 2018.
49. Lohmann P, Stoffels G, Ceccon G, et al. Radiation injury vs. recurrent brain metastasis: Combining textural feature radiomics analysis and standard parameters may increase 18F-FET PET accuracy without dynamic scans. *Eur Radiol* 2017;27:2916-2927.
50. Roman-Jimenez G, Acosta O, Leseur J, et al. Random forests to predict tumor recurrence following cervical cancer therapy using pre- and per-treatment 18F-FDG PET parameters. *Conf Proc IEEE Eng Med Biol Soc* 2016;2016:2444-2447.
51. Desbordes P, Ruan S, Modzelewski R, et al. Predictive value of initial FDG-PET features for treatment response and survival in esophageal cancer patients treated with chemo-radiation therapy using a random forest classifier. *PloS One* 2017;12:e0173208.
52. Tan S, Kligerman S, Chen W, et al. Spatial-temporal [^{18}F]FDG-PET features for predicting pathologic response of esophageal cancer to neoadjuvant chemoradiation therapy. *Int J Radiat Oncol Biol Phys* 2013;85:1375-1382.
53. van Rossum PSN, Fried DV, Zhang L, et al. The incremental value of subjective and quantitative assessment of 18F-FDG PET for the prediction of pathologic complete response to preoperative chemoradiotherapy in esophageal cancer. *J Nucl Med Off Publ Soc Nucl Med* 2016;57:691-700.
54. Yu H, Caldwell C, Mah K, et al. Automated radiation targeting in head-and-neck cancer using region-based texture analysis of PET and CT images. *Int J Radiat Oncol* 2009;75:618-625.
55. Oh JS, Kang BC, Roh J-L, et al. Intratumor textural heterogeneity on pretreatment (18)F-FDG PET images predicts response and survival after chemoradiotherapy for hypopharyngeal cancer. *Ann Surg Oncol* 2015;22:2746-2754.
56. Cheng N-M, Fang Y-HD, Chang JT-C, et al. Textural features of pretreatment 18F-FDG PET/CT images: Prognostic significance in patients with advanced T-stage oropharyngeal squamous cell carcinoma. *J Nucl Med Off Publ Soc Nucl Med* 2013;54:1703-1709.
57. Lovinfosse P, Janvary ZL, Coucke P, et al. FDG PET/CT texture analysis for predicting the outcome of lung cancer treated by stereotactic body radiation therapy. *Eur J Nucl Med Mol Imaging* 2016;43:1453-1460.
58. Takeda K, Takanami K, Shirata Y, et al. Clinical utility of texture analysis of 18F-FDG PET/CT in patients with Stage I lung cancer treated with stereotactic body radiotherapy. *J Radiat Res (Tokyo)* 2017;58:862-869.
59. Wu J, Aguilera T, Shultz D, et al. Early-stage non-small cell lung cancer: Quantitative imaging characteristics of (18)F fluorodeoxyglucose PET/CT allow prediction of distant metastasis. *Radiology* 2016;281:270-278.
60. Yip S, McCall K, Aristophanous M, et al. Comparison of texture features derived from static and respiratory-gated PET images in non-small cell lung cancer. *PloS One* 2014;9:e115510.
61. Bang J-I, Ha S, Kang S-B, et al. Prediction of neoadjuvant radiation chemotherapy response and survival using pretreatment [(18)F]FDG PET/CT scans in locally advanced rectal cancer. *Eur J Nucl Med Mol Imaging* 2016;43:422-431.
62. Lapa C, Werner RA, Schmid J-S, et al. Prognostic value of positron emission tomography-assessed tumor heterogeneity in patients with thyroid cancer undergoing treatment with radiopeptide therapy. *Nucl Med Biol* 2015;42:349-354.
63. Zhang L, Fried DV, Fave XJ, et al. IBEX: An open infrastructure software platform to facilitate collaborative work in radiomics. *Med Phys* 2015;42:1341-1353.
64. Nioche C, Orlhac F, Boughdad S, et al. LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res* 2018. canres.0125.2018.
65. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017;77:e104-e107.
66. Alfonso JCL, Jagiella N, Núñez L, et al. Estimating dose painting effects in radiotherapy: A mathematical model. *PloS One* 2014;9:e89380.
67. Orlhac F, Thézé B, Soussan M, et al. Multiscale Texture Analysis: From 18F-FDG PET Images to Histologic Images. *J Nucl Med Off Publ Soc Nucl Med* 2016;57:1823-1828.
68. Panth KM, Leijenaar RTH, Carvalho S, et al. Is there a causal relationship between genetic changes and radiomics-based image features? An in vivo preclinical experiment with doxycycline inducible GADD34 tumor cells. *Radiother Oncol* 2015;116:462-466.
69. Ling CC, Humm J, Larson S, et al. Towards multidimensional radiotherapy (MD-CRT): Biological imaging and biological conformality. *Int J Radiat Oncol* 2000;47:551-560.
70. Lelandais B, Ruan S, Denœux T, et al. Fusion of multi-tracer PET images for dose painting. *Med Image Anal* 2014;18:1247-1259.
71. Even AJG, van der Stoep J, Zegers CML, et al. PET-based dose painting in non-small cell lung cancer: Comparing uniform dose escalation with boosting hypoxic and metabolically active sub-volumes. *Radiother Oncol* 2015;116:281-286.
72. Berwouts D, De Wolf K, Lambert B, et al. Biological ^{18}[F]-FDG-PET image-guided dose painting by numbers for painful uncomplicated bone metastases: A 3-arm randomized phase II trial. *Radiother Oncol J Eur Soc Ther Radiol Oncol* 2015;115:272-278.
73. Trani D, Yaromina A, Dubois L, et al. Preclinical assessment of efficacy of radiation dose painting based on intratumoral FDG-PET uptake. *Clin Cancer Res Off J Am Assoc Cancer Res* 2015;21:5511-5518.
74. Chang JH, Wada M, Anderson NJ, et al. Hypoxia-targeted radiotherapy dose painting for head and neck cancer using (18)F-FMISO PET: A biological modeling study. *Acta Oncol Stockh Swed* 2013;52:1723-1729.
75. Differding S, Sterpin E, Janssens G, et al. Methodology for adaptive and robust FDG-PET escalated dose painting by numbers in head and neck tumors. *Acta Oncol Stockh Swed* 2016;55:217-225.
76. Shi X, Meng X, Sun X, et al. PET/CT imaging-guided dose painting in radiation therapy. *Cancer Lett* 2014;355:169-175.
77. Wilson JM, Mukherjee S, Chu K-Y, et al. Challenges in using ^{18}F-fluorodeoxyglucose-PET-CT to define a biological radiotherapy boost volume in locally advanced pancreatic cancer. *Radiat Oncol Lond Engl* 2014;9:146.
78. von Eyben FE, Kairemo K, Kiljunen T, et al. Planning of external beam radiotherapy for prostate cancer guided by PET/CT. *Curr Radiopharm* 2015;8:19-31.
79. Nkhali L, Thureau S, Edet-Sanson A, et al. FDG-PET/CT during concomitant chemo radiotherapy for esophageal cancer: Reducing target volumes to deliver higher radiotherapy doses. *Acta Oncol Stockh Swed* 2015;54:909-915.
80. Cihoric N, Tapia C, Krüger K, et al. IMRT with ^{18}FDG-PET⧹CT based simultaneous integrated boost for treatment of nodal positive cervical cancer. *Radiat Oncol Lond Engl* 2014;9:83.
81. Kolarova I, Vanasek J, Kandrnal V, et al. PET/CT significance for planning radiotherapy of head and neck cancer. *Neoplasma* 2012;59:536-540.
82. Badakhshi H, Graf R, Prasad V, et al. The impact of ^{18} F-FET PET-CT on target definition in image-guided stereotactic radiotherapy in patients with skull base lesions. *Cancer Imaging* 2014;14:25.
83. Berwouts D, Madani I, Duprez F, et al. Long-term outcome of ^{18}F-fluorodeoxyglucose-positron emission tomography-guided dose

painting for head and neck cancer: Matched case-control study. *Head Neck* 2017;39:2264-2275.
84. Differding S, Sterpin E, Hermand N, et al. Radiation dose escalation based on FDG-PET driven dose painting by numbers in oropharyngeal squamous cell carcinoma: A dosimetric comparison between TomoTherapy-HA and RapidArc. *Radiat Oncol Lond Engl* 2017;12:59.
85. Rasmussen JH, Håkansson K, Vogelius IR, et al. Phase I trial of 18F-fludeoxyglucose based radiation dose painting with concomitant cisplatin in head and neck cancer. *Radiother Oncol* 2016;120:76-80.
86. Berwouts D, Olteanu LAM, Duprez F, et al. Three-phase adaptive dose-painting-by-numbers for head-and-neck cancer: Initial results of the phase I clinical trial. *Radiother Oncol* 2013;107:310-316.
87. Dirix P, Vandecaveye V, De Keyzer F, et al. Dose painting in radiotherapy for head and neck squamous cell carcinoma: Value of repeated functional imaging with (18)F-FDG PET, (18)F-fluoromisonidazole PET, diffusion-weighted MRI, and dynamic contrast-enhanced MRI. *J Nucl Med* 2009;50:1020-1027.
88. Schlenter M, Berneking V, Krenkel B, et al. Intensity-modulated radiotherapy of prostate cancer with simultaneous integrated boost after molecular imaging with 18F-choline-PET/CT: Clinical results and quality of life. *Strahlenther Onkol* 2018;194:638-645.
89. Treglia G, Sadeghi R, Del Sole A, et al. Diagnostic performance of PET/CT with tracers other than F-18-FDG in oncology: An evidence-based review. *Clin Transl Oncol* 2014;16:770-775.
90. Provost C, Prignon A, Cazes A, et al. 68Ga-DOTATOC and FDG PET imaging of preclinical neuroblastoma models. *Anticancer Res* 2016;36:4459-4466.
91. van Kruchten M, de Vries EFJ, Arts HJG, et al. Assessment of estrogen receptor expression in epithelial ovarian cancer patients using 16α-18F-fluoro-17β-estradiol PET/CT. *J Nucl Med* 2015;56:50-55.
92. Mitamura K, Yamamoto Y, Kudomi N, et al. Intratumoral heterogeneity of (18)F-FLT uptake predicts proliferation and survival in patients with newly diagnosed gliomas. *Ann Nucl Med* 2017;31:46-52.
93. Cicone F, Minniti G, Romano A, et al. Accuracy of F-DOPA PET and perfusion-MRI for differentiating radionecrotic from progressive brain metastases after radiosurgery. *Eur J Nucl Med Mol Imaging* 2015;42:103-111.
94. Phelps ME, Hoffman EJ, Mullani NA, et al. Application of annihilation coincidence detection to transaxial reconstruction tomography. *J Nucl Med* 1975;16:210-224.
95. Beyer T, Townsend DW, Brun T, et al. A combined PET/CT scanner for clinical oncology. *J Nucl Med* 2000;41:1369-1379.
96. Charron M, Beyer T, Bohnen NN, et al. Image analysis in patients with cancer studied with a combined PET and CT scanner. *Clin Nucl Med* 2000;25:905-910.
97. Karlberg AM, Sæther O, Eikenes L, et al. Quantitative comparison of PET performance—Siemens Biograph mCT and mMR. *EJNMMI Phys* 2016;3:5.
98. Surti S, Karp JS. Advances in time-of-flight PET. *Phys Med* 2016;32:12-22.
99. van der Vos CS, Koopman D, Rijnsdorp S, et al. Quantification, improvement, and harmonization of small lesion detection with state-of-the-art PET. *Eur J Nucl Med Mol Imaging* 2017;44:4-16.
100. Quak E, Le Roux P-Y, Lasnon C, et al. Does PET SUV harmonization affect PERCIST response classification? *J Nucl Med* 2016;57:1699-1706.
101. Lasnon C, Hicks RJ, Beauregard J-M, et al. Impact of point spread function reconstruction on thoracic lymph node staging with 18F-FDG PET/CT in non-small cell lung cancer. *Clin Nucl Med* 2012;37:971-976.
102. Andersen FL, Klausen TL, Loft A, et al. Clinical evaluation of PET image reconstruction using a spatial resolution model. *Eur J Radiol* 2013;82:862-869.
103. Akamatsu G, Mitsumoto K, Taniguchi T, et al. Influences of point-spread function and time-of-flight reconstructions on standardized uptake value of lymph node metastases in FDG-PET. *Eur J Radiol* 2014;83:226-230.
104. Cortes-Rodicio J, Sanchez-Merino G, Garcia-Fidalgo MA, et al. Identification of low variability textural features for heterogeneity quantification of (18)F-FDG PET/CT imaging. *Rev Esp Med Nucl Imagen Mol* 2016;35:379-384.
105. Bailly C, Bodet-Milin C, Couespel S, et al. Revisiting the robustness of PET-based textural features in the context of multi-centric trials. *PloS One* 2016;11:e0159984.
106. Galavis PE, Hollensen C, Jallow N, et al. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. *Acta Oncol* 2010;49:1012-1016.
107. Yan J, Chu-Shern JL, Loi HY, et al. Impact of image reconstruction settings on texture features in 18F-FDG PET. *J Nucl Med* 2015;56:1667-1673.
108. Shiri I, Rahmim A, Ghaffarian P, et al. The impact of image reconstruction settings on 18F-FDG PET radiomic features: Multi-scanner phantom and patient studies. *Eur Radiol* 2017;27:4498-4509.
109. Nyflot MJ, Yang F, Byrd D, et al. Quantitative radiomics: Impact of stochastic effects on textural feature analysis implies the need for standards. *J Med Imaging (Bellingham)* 2015;2:041002.
110. Shafiq-Ul-Hassan M, Zhang GG, Latifi K, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. *Med Phys* 2017;44:1050-1062.
111. Mackin D, Fave X, Zhang L, et al. Harmonizing the pixel size in retrospective computed tomography radiomics studies. *PLoS One* 2017;12:e0178524.
112. Carles M, Torres-Espallardo I, Alberich-Bayarri A, et al. Evaluation of PET texture features with heterogeneous phantoms: Complementarity and effect of motion and segmentation method. *Phys Med Biol* 2017;62:652-668.
113. Oliver JA, Budzevich M, Zhang GG, et al. Variability of image features computed from conventional and respiratory-gated PET/CT images of lung cancer. *Transl Oncol* 2015;8:524-534.
114. Carles M, Bach T, Torres-Espallardo I, et al. Significance of the impact of motion compensation on the variability of PET image features. *Phys Med Biol* 2018;63:065013.
115. Oliver JA, Budzevich M, Hunt D, et al. Sensitivity of image features to noise in conventional and respiratory-gated PET/CT images of lung cancer: Uncorrelated noise effects. *Technol Cancer Res Treat* 2017;16:595-608.
116. Surasi DS, Bhambhvani P, Baldwin JA, et al. 18F-FDG PET and PET/CT patient preparation: A review of the literature. *J Nucl Med Technol* 2014;42:5-13.
117. Lasnon C, Houdu B, Kammerer E, et al. Patient's weight: A neglected cause of variability in SUV measurements? A survey from an EARL accredited PET centre in 513 patients. *Eur J Nucl Med Mol Imaging* 2016;43:197-199.
118. Beaulieu S, Kinahan P, Tseng J, et al. SUV varies with time after injection in (18)F-FDG PET of breast cancer: Characterization and method to adjust for time differences. *J Nucl Med* 2003;44:1044-1050.
119. Kurland BF, Muzi M, Peterson LM, et al. Multicenter clinical trials using 18F-FDG PET to measure early response to oncologic therapy: Effects of injection-to-acquisition time variability on required sample size. *J Nucl Med* 2016;57:226-230.
120. Wangerin KA, Muzi M, Peterson LM, et al. Effect of (18)F-FDG uptake time on lesion detectability in PET imaging of early stage breast cancer. *Tomogr J Imaging Res* 2015;1:53-60.
121. de Groot EH, Post N, Boellaard R, et al. Optimized dose regimen for whole-body FDG-PET imaging. *EJNMMI Res* 2013;3:63.
122. Quak E, Le Roux P-Y, Hofman MS, et al. Harmonizing FDG PET quantification while maintaining optimal lesion detection: Prospective multicentre validation in 517 oncology patients. *Eur J Nucl Med Mol Imaging* 2015;42:2072-2082.
123. Boellaard R. The engagement of FDG PET/CT image quality and harmonized quantification: From competitive to complementary. *Eur J Nucl Med Mol Imaging* 2016;43:1-4.

124. Lasnon C, Majdoub M, Lavigne B, et al. ^{18}F-FDG PET/CT heterogeneity quantification through textural features in the era of harmonisation programs: A focus on lung cancer. *Eur J Nucl Med Mol Imaging* 2016;43:2324-2335.
125. Orlhac F, Boughdad S, Philippe C, et al. A post-reconstruction harmonization method for multicenter radiomic studies in PET. *J Nucl Med* 2018. jnumed.117.199935.
126. MacManus M, Nestle U, Rosenzweig KE, et al. Use of PET and PET/CT for radiation therapy planning: IAEA expert report 2006-2007. *Radiother Oncol* 2009;91:85-94.
127. Drever L, Roa W, McEwan A, et al. Iterative threshold segmentation for PET target volume delineation. *Med Phys* 2007;34:1253-1265.
128. Foster B, Bagci U, Mansoor A, et al. A review on segmentation of positron emission tomography images. *Comput Biol Med* 2014;50:76-96.
129. Adams R, Bischof L. Seeded region growing. *IEEE Trans Pattern Anal Mach Intell* 1994;16:641-647.
130. Bağci U, Yao J, Caban J, et al. A graph-theoretic approach for segmentation of PET images. *Conf Proc IEEE Eng Med Biol Soc* 2011;2011:8479-8482.
131. Grady L. Random walks for image segmentation. *IEEE Trans Pattern Anal Mach Intell* 2006;28:1768-1783.
132. Onoma DP, Ruan S, Thureau S, et al. Segmentation of heterogeneous or small FDG PET positive tissue based on a 3D-locally adaptive random walk algorithm. *Comput Med Imaging Graph* 2014;38:753-763.
133. Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. *Int J Comput Vis* 1988;1:321-331.
134. Geets X, Lee JA, Bol A, et al. A gradient-based method for segmenting FDG-PET images: Methodology and validation. *Eur J Nucl Med Mol Imaging* 2007;34:1427-1438.
135. Aristophanous M, Penney BC, Martel MK, et al. A Gaussian mixture model for definition of lung tumor volumes in positron emission tomography. *Med Phys* 2007;34:4223-4235.
136. Hatt M, Cheze le Rest C, Descourt P, et al. Accurate automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications. *Int J Radiat Oncol Biol Phys* 2010;77:301-308.
137. Hatt M, Cheze le Rest C, Turzo A, et al. A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. *IEEE Trans Med Imaging* 2009;28:881-893.
138. Pham DL, Xu C, Prince JL. Current methods in medical image segmentation. *Annu Rev Biomed Eng* 2000;2:315-337.
139. Ma Z, Tavares JMRS, Jorge RN. Segmentation of structures in medical images: Review and a new computational framework. *8th Int Symp Comput Methods Biomech Biomed Eng* 2008. https://web.fe.up.pt/~tavares/downloads/publications/artigos/ZM_CMBBE2008_paper.pdf.
140. Hatt M, Lee JA, Schmidtlein CR, et al. Classification and evaluation strategies of auto-segmentation approaches for PET: Report of AAPM task group No. 211. *Med Phys* 2017;44:e1-e42.
141. Devic S, Tomic N, Faria S, et al. Defining radiotherapy target volumes using 18F-fluoro-deoxy-glucose positron emission tomography/computed tomography: Still a Pandora's box? *Int J Radiat Oncol Biol Phys* 2010;78:1555-1562.
142. Nestle U, Kremp S, Schaefer-Schuler A, et al. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med Off Publ Soc Nucl Med* 2005;46:1342-1348.
143. Dewalle-Vignion A-S, Yeni N, Petyt G, et al. Evaluation of PET volume segmentation methods: Comparisons with expert manual delineations. *Nucl Med Commun* 2012;33:34-42.
144. Withofs N, Bernard C, Van der Rest C, et al. FDG PET/CT for rectal carcinoma radiotherapy treatment planning: Comparison of functional volume delineation algorithms and clinical challenges. *J Appl Clin Med Phys* 2014;15:4696.
145. Cheebsumon P, Yaqub M, van Velden FHP, et al. Impact of [^{18}F]FDG PET imaging parameters on automatic tumour delineation: Need for improved tumour delineation methodology. *Eur J Nucl Med Mol Imaging* 2011;38:2136-2144.
146. Bashir U, Azad G, Siddique MM, et al. The effects of segmentation algorithms on the measurement of 18F-FDG PET texture parameters in non-small cell lung cancer. *EJNMMI Res* 2017;7:60.
147. Ben Bouallègue F, Tabaa YA, Kafrouni M, et al. Association between textural and morphological tumor indices on baseline PET-CT and early metabolic response on interim PET-CT in bulky malignant lymphomas. *Med Phys* 2017;44:4608-4619.
148. Johnson PB, Young LA, Lamichhane N, et al. Quantitative imaging: Correlating image features with the segmentation accuracy of PET based tumor contours in the lung. *Radiother Oncol* 2017;123:257-262.
149. van Velden FHP, Kramer GM, Frings V, et al. Repeatability of radiomic features in non-small-cell lung cancer [(18)F]FDG-PET/CT studies: Impact of reconstruction and delineation. *Mol Imaging Biol* 2016;18:788-795.
150. Knudtsen IS, van Elmpt W, Ollers M, et al. Impact of PET reconstruction algorithm and threshold on dose painting of non-small cell lung cancer. *Radiother Oncol* 2014;113:210-214.
151. Orlhac F, Soussan M, Maisonobe J-A, et al. 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* 2014;55:414-422.
152. Hatt M, Dekker A, De Ruysscher D, et al. Accurate functional volume definition in PET for radiotherapy treatment planning, 2008 IEEE Nuclear Science Symposium Conference Record, Dresden, Germany, 2008, pp. 5567-5571. https://doi.org/10.1109/NSSMIC.2008.4774509.
153. Hatt M, Tixier F, Cheze Le Rest C, et al. Robustness of intratumour ^{18}F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. *Eur J Nucl Med Mol Imaging* 2013;40:1662-1671.
154. Altazi BA, Zhang GG, Fernandez DC, et al. Reproducibility of F18-FDG PET radiomic features for different cervical tumor segmentation methods, gray-level discretization, and reconstruction algorithms. *J Appl Clin Med Phys* 2017;18:32-48.
155. Foley KG, Hills RK, Berthon B, et al. Development and validation of a prognostic model incorporating texture analysis derived from standardised segmentation of PET in patients with oesophageal cancer. *Eur Radiol* 2018;28:428-436.
156. Berthon B, Marshall C, Evans M, et al. ATLAAS: An automatic decision tree-based learning algorithm for advanced image segmentation in positron emission tomography. *Phys Med Biol* 2016;61:4855-4869.
157. Leijenaar RTH, Nalbantov G, Carvalho S, et al. The effect of SUV discretization in quantitative FDG-PET radiomics: The need for standardized methodology in tumor texture analysis. *Sci Rep* 2015;5:11075.
158. Orlhac F, Soussan M, Chouahnia K, et al. 18F-FDG PET-Derived Textural Indices Reflect Tissue-Specific Uptake Pattern in Non-Small Cell Lung Cancer. *PloS One* 2015;10:e0145063.
159. Vallières M, Freeman CR, Skamene SR, et al. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Phys Med Biol* 2015;60:5471-5496.
160. Orlhac F, Nioche C, Soussan M, et al. Understanding changes in tumor textural indices in PET: a comparison between visual assessment and index values in simulated and patient data. *J Nucl Med Off Publ Soc Nucl Med* 2016.
161. Schernberg A, Reuze S, Orlhac F, et al. A score combining baseline neutrophilia and primary tumor SUVpeak measured from FDG PET is associated with outcome in locally advanced cervical cancer. *Eur J Nucl Med Mol Imaging* 2017.

162. Hatt M, Tixier F, Pierce L, et al. Characterization of PET/CT images using texture analysis: The past, the present … any future? *Eur J Nucl Med Mol Imaging* 2017;44:151-165.
163. Paragios N, Duncan J, Ayache N, editors. Handbook of Biomedical Imaging − Methodologies and Clinical Research. New York: Springer; 2015.
164. Parmar C, Grossmann P, Bussink J, et al. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep* 2015;5:13087.
165. Gao X, Chu C, Li Y, et al. The method and efficacy of support vector machine classifiers based on texture features and multi-resolution histogram from (18)F-FDG PET-CT images for the evaluation of mediastinal lymph nodes in patients with lung cancer. *Eur J Radiol* 2015;84:312-317.
166. Lartizien C, Rogez M, Niaf E, et al. Computer-aided staging of lymphoma patients with FDG PET/CT imaging based on textural information. *IEEE J Biomed Health Inform* 2014;18:946-955.
167. Antunovic L, Gallivanone F, Sollini M, et al. [18F]FDG PET/CT features for the molecular characterization of primary breast tumors. *Eur J Nucl Med Mol Imaging* 2017;44:1945-1954.
168. Gallivanone F, Panzeri MM, Canevari C, et al. Biomarkers from in vivo molecular imaging of breast cancer: Pretreatment 18F-FDG PET predicts patient prognosis, and pretreatment DWI-MR predicts response to neoadjuvant chemotherapy. *Magma N Y N* 2017;30:359-373.
169. Balagurunathan Y, Kumar V, Gu Y, et al. Test−retest reproducibility analysis of lung CT image features. *J Digit Imaging* 2014;27:805-823.
170. Armato SG, Meyer CR, Mcnitt-Gray MF, et al. The Reference Image Database to Evaluate Response to therapy in lung cancer (RIDER) project: A resource for the development of change-analysis software. *Clin Pharmacol Ther* 2008;84:448-456.
171. Kohavi R, John GH. Wrappers for feature subset selection. *Artif Intell* 1997;97:273-324.
172. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc Ser B Stat Methodol* 2005;67:301-320.
173. Leger S, Zwanenburg A, Pilz K, et al. A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. *Sci Rep* 2017;7:13206.
174. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436-444.
175. Lao J, Chen Y, Li Z-C, et al. A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme. *Sci Rep* 2017;7:10353.
176. Pan SJ, Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng* 2010;22:1345-1359.
177. Orlhac F, Nioche C, Soussan M, et al. Understanding Changes in Tumor Texture Indices in PET: A Comparison Between Visual Assessment and Index Values in Simulated and Patient Data. *J Nucl Med Off Publ Soc Nucl Med* 2017;58:387-392.
178. Tixier F, Hatt M, Valla C, et al. Visual versus quantitative assessment of intratumor 18F-FDG PET uptake heterogeneity: Prognostic value in non-small cell lung cancer. *J Nucl Med* 2014;55:1235-1241.
179. Thomas CM, Pike LC, Hartill CE, et al. Specific recommendations for accurate and direct use of PET-CT in PET guided radiotherapy for head and neck sites. *Med Phys* 2014;41:041710.
180. Geiger GA, Kim MB, Xanthopoulos EP, et al. Stage migration in planning PET/CT scans in patients due to receive radiotherapy for non-small-cell lung cancer. *Clin Lung Cancer* 2014;15:79-85.
181. Sam S, Shon IH, Vinod SK, et al. Workflow and radiation safety implications of (18)F-FDG PET/CT scans for radiotherapy planning. *J Nucl Med Technol* 2012;40:175-177.
182. Berker Y, Li Y. Attenuation correction in emission tomography using the emission data − A review. *Med Phys* 2016;43:807-832.
183. Sjölund J, Forsberg D, Andersson M, et al. Generating patient specific pseudo-CT of the head from MR using atlas-based regression. *Phys Med Biol* 2015;60:825-839.
184. Paulus DH, Oehmigen M, Grüneisen J, et al. Whole-body hybrid imaging concept for the integration of PET/MR into radiation therapy treatment planning. *Phys Med Biol* 2016;61:3504-3520.
185. Wang H, Chandarana H, Block KT, et al. Dosimetric evaluation of synthetic CT for magnetic resonance-only based radiotherapy planning of lung cancer. *Radiat Oncol Lond Engl* 2017;12:108.
186. Leibfarth S, Mönnich D, Welz S, et al. A strategy for multimodal deformable image registration to integrate PET/MR into radiotherapy treatment planning. *Acta Oncol Stockh Swed* 2013;52:1353-1359.
187. Desseroit M-C, Visvikis D, Tixier F, et al. Development of a nomogram combining clinical staging with (18)F-FDG PET/CT image features in non-small-cell lung cancer stage I-III. *Eur J Nucl Med Mol Imaging* 2016;43:1477-1485.
188. Lucia F, Visvikis D, Desseroit M-C, et al. Prediction of outcome using pretreatment18F-FDG PET/CT and MRI radiomics in locally advanced cervical cancer treated with chemoradiotherapy. *Eur J Nucl Med Mol Imaging* 2018;45:768-786.
189. Zhou Z, Folkert M, Iyengar P, et al. Multi-objective radiomics model for predicting distant failure in lung SBRT. *Phys Med Biol* 2017;62:4460-4478.
190. Desseroit M-C, Tixier F, Weber WA, et al. Reliability of PET/CT shape and heterogeneity features in functional and morphologic components of non-small cell lung cancer tumors: A repeatability analysis in a prospective multicenter cohort. *J Nucl Med* 2017;58:406-411.
191. Bogowicz M, Riesterer O, Stark LS, et al. Comparison of PET and CT radiomics for prediction of local tumor control in head and neck squamous cell carcinoma. *Acta Oncol Stockh Swed* 2017;1-6.
192. Bowen SR, Yuh WTC, Hippe DS, et al. Tumor radiomic heterogeneity: Multiparametric functional imaging to characterize variability and predict response following cervical cancer radiation therapy. *J Magn Reson Imaging JMRI* 2018;47:1388-1396.
193. Sun R, Orlhac F, Robert C, et al. In Regard to Mattonen et al. *Int J Radiat Oncol Biol Phys* 2016;95:1544-1545.
194. Soufi M, Kamali-Asl A, Geramifar P, et al. A novel framework for automated segmentation and labeling of homogeneous versus heterogeneous lung tumors in [(18)F]FDG-PET imaging. *Mol Imaging Biol* 2017;19:456-468.
195. Shen S, Jacob R, Bender LW, et al. A technique using 99mTc-mebrofenin SPECT for radiotherapy treatment planning for liver cancers or metastases. *Med Dosim* 2014;39:7-11.
196. Ritt P, Vija H, Hornegger J, et al. Absolute quantification in SPECT. *Eur J Nucl Med Mol Imaging* 2011;38(Suppl 1):S69-S77.
197. De Bari B, Deantonio L, Bourhis J, et al. Should we include SPECT lung perfusion in radiotherapy treatment plans of thoracic targets? Evidences from the literature. *Crit Rev Oncol Hematol* 2016;102:111-117.
198. Lopez-Gaitan J, Ebert MA, Robins P, et al. Radiotherapy of abdomen with precise renal assessment with SPECT/CT imaging (RAPRASI): Design and methodology of a prospective trial to improve the understanding of kidney radiation dose response. *BMC Cancer* 2013;13:381.
199. Rahmim A, Salimpour Y, Jain S, et al. Application of texture analysis to DAT SPECT imaging: Relationship to clinical assessments. *Neuroimage Clin* 2016;12:e1-e9.
200. Kalpathy-Cramer J, Freymann JB, Kirby JS, et al. Quantitative imaging network: Data sharing and competitive algorithm validation leveraging. The Cancer Imaging Archive. *Transl Oncol* 2014;7:147-152.
201. Schlett CL, Hendel T, Weckbach S, et al. Population-based imaging and radiomics: Rationale and perspective of the German national cohort MRI study. *Rofo* 2016;188:652-661.