European Journal of Nuclear Medicine and Molecular Imaging (2023) 50:2548–2560
https://doi.org/10.1007/s00259-023-06187-3

ORIGINAL ARTICLE

Role of [^{68}Ga]Ga-PSMA-11 PET radiomics to predict post-surgical ISUP grade in primary prostate cancer

Samuele Ghezzo[1,2] · Paola Mapelli[1,2] · Carolina Bezzi[1,2] · Ana Maria Samanes Gajate[2] · Giorgio Brembilla[1,3] · Irene Gotuzzo[4] · Tommaso Russo[1,3] · Erik Preza[2] · Vito Cucchiara[1,5,6] · Naghia Ahmed[7] · Ilaria Neri[1,2] · Sofia Mongardi[1] · Massimo Freschi[7] · Alberto Briganti[1,5,6] · Francesco De Cobelli[1,3] · Luigi Gianolli[2] · Paola Scifo[2] · Maria Picchio[1,2]

Received: 6 December 2022 / Accepted: 1 March 2023 / Published online: 18 March 2023

Abstract

Purpose The aim of this study is to investigate the role of [^{68}Ga]Ga-PSMA-11 PET radiomics for the prediction of post-surgical International Society of Urological Pathology ($_{PS}$ISUP) grade in primary prostate cancer (PCa).

Methods This retrospective study included 47 PCa patients who underwent [^{68}Ga]Ga-PSMA-11 PET at IRCCS San Raffaele Scientific Institute before radical prostatectomy. The whole prostate was manually contoured on PET images and 103 image biomarker standardization initiative (IBSI)-compliant radiomic features (RFs) were extracted. Features were then selected using the minimum redundancy maximum relevance algorithm and a combination of the 4 most relevant RFs was used to train 12 radiomics machine learning models for the prediction of $_{PS}$ISUP grade: ISUP ≥ 4 vs ISUP < 4. Machine learning models were validated by means of fivefold repeated cross-validation, and two control models were generated to assess that our findings were not surrogates of spurious associations. Balanced accuracy (bACC) was collected for all generated models and compared with Kruskal–Wallis and Mann–Whitney tests. Sensitivity, specificity, and positive and negative predictive values were also reported to provide a complete overview of models' performance. The predictions of the best performing model were compared against ISUP grade at biopsy.

Results ISUP grade at biopsy was upgraded in 9/47 patients after prostatectomy, resulting in a bACC = 85.9%, SN = 71.9%, SP = 100%, PPV = 100%, and NPV = 62.5%, while the best-performing radiomic model yielded a bACC = 87.6%, SN = 88.6%, SP = 86.7%, PPV = 94%, and NPV = 82.5%. All radiomic models trained with at least 2 RFs (GLSZM—Zone Entropy and Shape—Least Axis Length) outperformed the control models. Conversely, no significant differences were found for radiomic models trained with 2 or more RFs (Mann–Whitney $p > 0.05$).

Conclusion These findings support the role of [^{68}Ga]Ga-PSMA-11 PET radiomics for the accurate and non-invasive prediction of $_{PS}$ISUP grade.

Keywords Radiomics · PSMA · ISUP grade · PET · Prostate cancer

This article is part of the Topical Collection on Advanced Image Analyses (Radiomics and Artificial Intelligence)

✉ Maria Picchio
picchio.maria@hsr.it

1 Vita-Salute San Raffaele University, Milan, Italy

2 Nuclear Medicine Department, IRCCS San Raffaele Scientific Institute, Milan, Italy

3 Department of Radiology, IRCCS San Raffaele Scientific Institute, Milan, Italy

4 School of Medicine and Surgery, University of Milano Bicocca, Monza, Italy

5 Department of Urology, IRCCS San Raffaele Scientific Institute, Milan, Italy

6 Division of Experimental Oncology, URI, Urological Research Institute, Milan, Italy

7 Pathology Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy

Introduction

With more than 1.4 million new cases diagnosed in 2020, prostate cancer (PCa) is the second most commonly occurring cancer in men worldwide [1]. While digital rectal examination (DRE) and prostate specific antigen (PSA) levels are normally used to screen patients for PCa, histopathological verification in prostate biopsy cores is required for a definitive diagnosis [2].

International Society of Urological Pathology (ISUP) grade, derived from biopsy specimens, is one of the major determinants for PCa patient management and treatment planning [2]. However, as ISUP grade is measured on bioptic specimens, it is affected by sampling error and it cannot account for tumour heterogeneity [3]; moreover, bioptic sampling might entail the side effects of transrectal-biopsy, including haematuria, haematospermia, rectal bleeding, and pain [4].

Multiparametric magnetic resonance imaging (mp-MRI) targeted biopsy has shown to increase the detection of clinically significant PCa, with a concomitant reduction in the detection of insignificant PCa [2]. However, despite the introduction of the PI-RADS version 2 scoring system [5], mp-MRI inter-reader reproducibility is far from optimal, and any upgrading in ISUP grade after whole-mount histopathological analysis of surgical specimens is still reported in about 30% of mp-MRI targeted biopsies [2, 6].

Previous studies successfully established the utility of positron emission tomography (PET) with ^{68}Ga- prostate specific membrane antigen (PSMA)-11 for the characterization of primary PCa [7–12], reporting higher sensitivity of [^{68}Ga]Ga-PSMA-11 PET as compared to mp-MRI for tumour localization [10, 13].

Radiomics, defined as the high-throughput extraction of quantitative features from medical images, in combination with machine learning (ML) has been proven effective for the non-invasive characterization of several types of cancer, including PCa [14–19]. Nonetheless, the lack of standardized methodologies combined with the limited interpretability of the proposed models prevents radiomics to enter clinical practice.

The aim of the present study is to investigate the role of [^{68}Ga]Ga-PSMA-11 PET radiomics for the prediction of post-surgical ISUP grade ($_{PS}$ISUP). Our hypothesis is that radiomics bears the potential to predict $_{PS}$ISUP grade. Models' interpretability will be ensured by thorough controls and by the comparison of the generated signature with that presented in a recently published state-of-the-art work [18]. Furthermore, the potential clinical utility of [^{68}Ga]Ga-PSMA-11 PET radiomics will be assessed by confronting its performance to that of biopsy for $_{PS}$ISUP grade prediction.

Material and methods

Patients

We retrieved all the consecutive patients who underwent [^{68}Ga]Ga-PSMA-11 PET for PCa staging at IRCCS San Raffaele Scientific Institute between 09/09/2020 and 10/01/2022. Inclusion criteria were as follows: (i) presence of at least one intraprostatic pathological finding on [^{68}Ga]Ga-PSMA-11 PET; (ii) radical prostatectomy (RP) after PET imaging; (iii) availability of ISUP grade both at biopsy and RP; and (iv) absence of neoadjuvant treatments. Among the 144 patients screened, 47 met the criteria and were used for analysis. See Fig. 1 for a flowchart showing the patients' selection process.

Clinical and histopathological information were collected and post-surgical ISUP grades were used to stratify the cohort into low-grade ($_{PS}$ISUP < 4) and high-grade PCa ($_{PS}$ISUP ≥ 4). Extensive information on patients' characteristics is reported in Table 1.

This retrospective study was approved by the Institutional Ethics Committee of IRCCS San Raffaele Scientific Institute. Informed consent was waived due to the retrospective nature of the study.

PET imaging

PET scans were acquired using either Signa PET/MRI 3 Tesla system ($N = 25$) or PET/CT ($N = 22$), Discovery-STE or Discovery-690 (GE Healthcare, Waukesha, WI, USA). [^{68}Ga]Ga-PSMA-11 PET scans were acquired according to the normal clinical practice, following the joint EANM and SNMMI procedure guideline for PCa imaging [20].

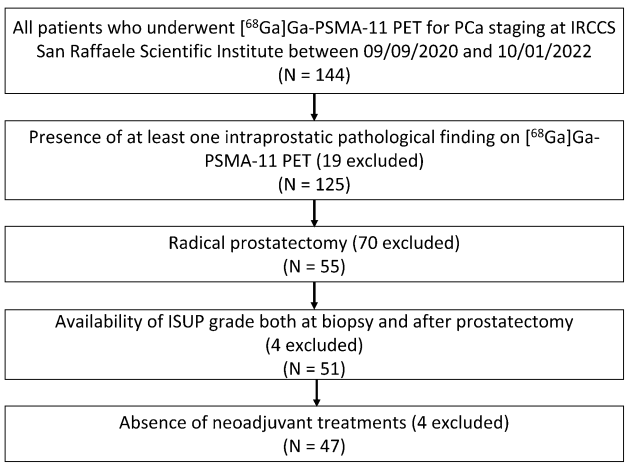

Fig. 1 Flowchart showing the patients' selection process

Table 1 Patients' characteristics

Patients' characteristics ($N=47$)	Value
Age (years), median (IQR)	65 (60 – 71)
iPSA (ng/ml), median (IQR)	6.32 (4.79 – 7.85)
Pathologic T staging, *n* (%)	
2	3 (6.4)
2a	2 (4.2)
2c	6 (12.8)
3a	23 (48.9)
3b	13 (27.7)
ISUP grade (biopsy), *n* (%)	
1	2 (4.3)
2	7 (14.9)
3	15 (31.9)
4	12 (25.5)
5	11 (23.4)
ISUP grade (RP), *n* (%)	
1	0 (0)
2	6 (12.8)
3	9 (19.1)
4	4 (8.5)
5	28 (59.6)

IQR, interquartile range

Specifically, [^{68}Ga]Ga-PSMA-11 PET/CT and PET/MRI scans started approximately 60 min after the injection of approximately 2 MBq per kilogram body weight, 4 min per bed position.

[^{68}Ga]Ga PSMA-11 PET image read-out was performed by two Nuclear Medicine physicians on the Advantage Workstation (AW, General Electric Healthcare, Waukesha, WI, USA) and the presence of [^{68}Ga]Ga-PSMA-11 intraprostatic focal and intense increased uptake, either with or without alterations visible on the co-registered CT or MR images, was considered positive for malignancy. SUV_{max} and SUV_{mean} were calculated for the dominant intraprostatic lesion in all patients showing at least one intraprostatic [^{68}Ga]Ga-PSMA-11 positive finding. Specifically, SUV_{mean} was calculated by using the 41% threshold of SUV_{max}, adapted by the reading physicians in those cases where the fixed threshold did not allow an optimal contouring.

Surgical procedure

All patients received radical prostatectomy plus extended pelvic lymph node dissection (ePLND) using either a robotic-assisted (RARP) or an open approach. Extended lymph node dissection was performed in all cases given a pre-surgical risk of lymph node invasion > 5% and it included the removal of fibrofatty tissue along the external iliac vein, the lateral limit being the genitofemoral nerve and the distal limit being the deep circumflex vein. Proximally, the cranial limit was represented by the crossing between the ureter and common iliac vessels. All fibrofatty tissue within the obturator fossa was removed. The dissection was performed from lateral to medial up to the umbilical artery and the bladder wall represented the medial limit. Lymph nodes along as well as medially and laterally to the internal iliac vessels were also removed.

Histopathological analysis

Radical prostatectomy fresh specimen was received from the operating room. The specimen was covered on surface with India ink, and fixed for 24 h in formalin, embedded, and examined with the whole-mount technique [21]. The prostate body was step-sectioned at 4-mm intervals, perpendicular to the gland's long axis (apical-basal). The 2014 ISUP/WHO modified Gleason score and Grade Group were provided [22]. The index lesion was considered as the cancer focus with the highest Gleason score, or the largest focus as measured by the volume in the case of more than one lesion with the same Gleason score. The index lesion site, pT, grading group, extra prostatic extension (yes/no), positive surgical margins (yes/no), and metastatic lymph nodes (number, maximum diameter, and site) were gathered.

Image segmentation

Two radiologists manually segmented the whole prostate on co-registered PET/CT or PET/MR images. CT or MRI was used for anatomical localization on 3D Slicer (revision 29,402) [23] because of its high anatomical accuracy. As the co-registered PET/CT or PET/MR images were used for contouring, it follows that the VOIs have been drawn directly on PET.

The whole prostate was used as volume of interest (VOI) as it provides a more robust inter-reader reproducibility [24] and better accounts for tumour heterogeneity, while at the same time avoiding sampling error, issues of radiomics for small lesions (less than 64 voxels) [25], and the difficulties of multi-lesion characterization.

Feature extraction and selection

All VOIs were normalized, discretized using fixed bin width (FBW = 0.3), and then resampled to $2.0 \times 2.0 \times 2.0$ mm^3 voxels via sitkBspline on PyRadiomics [26]. Afterwards, 103 3D Image Biomarker Standardization Initiative (IBSI) [27] compliant radiomic features (RFs) were extracted from the resampled VOIs on PET images using PyRadiomics. See supplementary Table 1 for a detailed description of the extracted RFs.

Since previous studies showed the additional utility bore by the inclusion of conventional [^{68}Ga]Ga-PSMA-11 PET parameters [18, 19], SUV_{max} and SUV_{mean} were considered for analysis, resulting in a 105-long feature vector for each patient.

ComBat software (https://forlhac.shinyapps.io/Shiny_ComBat/) was used to harmonize radiomic features extracted from images acquired with different scanners. Features' harmonization was especially relevant in this study because analyzed images were acquired on either PET/MR or PET/CT scanners [28]. The 4 most relevant, non-redundant RFs were then identified by means of minimum redundancy maximum relevance (mRMR) algorithm [29] and kept for further analysis.

Finally, to provide further insight into the relevance and robustness of the identified RFs, these were correlated with the RFs exploited by Solari et al. [18]. This recent work, at present, represents the most recent and biggest study currently available using the whole prostate as VOI for ^{68} Ga-PSMA-11 PET radiomics in combination with machine learning for the prediction of post-surgical Gleason Grade ($_{PS}$GG).

Machine learning models

Three separate machine learning classifiers, namely logistic regression (LR), support vector machine (SVM), and K nearest neighbour (KNN), were trained to predict $_{PS}$ISUP grade. The sample was split into training and validation sets using fivefold stratified cross-validation, ensuring the presence of 3 patients with low-grade PCa in each fold. Four sets were used as training and the remaining one as validation in each fold.

Since only 15 patients had $_{PS}$ISUP < 4 (~ 32%), oversample using Adaptive Synthetic (ADASYN) algorithm was applied to oversample all training features of the less prevalent class, resulting in a 1:1 proportion with the most prevalent class ($_{PS}$ISUP ≥ 4) [30]. For each of the 5 cross-validation folds, the training data was augmented and used to train the 3 different classifiers. The performance of trained models for the prediction of $_{PS}$ISUP grade was then tested on the non-augmented validation data. The average performance of the classifiers over the five folds in the validation cohort represents the estimate of the cross-validation estimators.

Cross-validation is nearly unbiased, but it is known to be highly variable when applied to small samples [31]. To reduce its variability, the whole process was repeated 10 times, changing every time the split into the five folds, and the mean of the repeated cross-validation estimators was used as final estimate of models' performance [31]. The whole pipeline for training and validation of the implemented models is depicted in Fig. 2.

To generate the most informative, yet simplest, model, the machine learning classifiers were trained using the 4 RFs identified via mRMR, and then analyses were repeated using only the top 3 RFs, 2 RFs, and finally using only the most relevant RF according to mRMR. Altogether, 12 radiomics machine learning models for the prediction of $_{PS}$ISUP grade were generated.

Additionally, to show that the performance of the generated radiomics models does not depend on spurious associations, two control models were created.

The first control model ("radiomics baseline") was developed using VOI volume and SUV_{max} as only inputs, to confirm that our models do not rely on surrogates of volume and intensity. The second control model ("PET zeros") was built to demonstrate the complementary value of different types of features; specifically, all voxel values within the investigated VOIs were set to 0, so that only shape features remained preserved while all the others became meaningless. If shape features are relevant for the prediction of $_{PS}$ISUP grade, this model will be accurate; otherwise, it will yield random predictions.

These two baseline models were trained and tested following the same workflow used for the radiomics models (2 models, 3 classifiers; resulting in 6 control machine learning models).

Since the cohort was imbalanced, the top performing approach was selected based on balanced accuracy (bACC) defined by the formula:

$$bACC = \frac{sensitivity + specificity}{2}$$

and 95% binomial proportion confidence intervals calculated with the Clopper-Pearson exact method were reported [32]. To provide a complete overview of the models' performance, also sensitivity (SN), specificity (SP), positive predictive value (PPV), and negative predictive value (NPV) are reported.

Statistical analysis

The performance of the control and radiomics models (with 4, 3, 2, and 1 RFs) obtained with LR, SVM, and KNN was compared by the Kruskal–Wallis test. Then, post hoc comparisons using the Mann–Whitney test were performed to study more in detail the differences between the generated models. The best models generated with the 3 different classifiers were also compared to assess whether results are affected by the choice of the classifier. The higher probability of type I errors due to familywise error rates was controlled by false discovery rate. p-values < 0.05 were considered statistically significant.

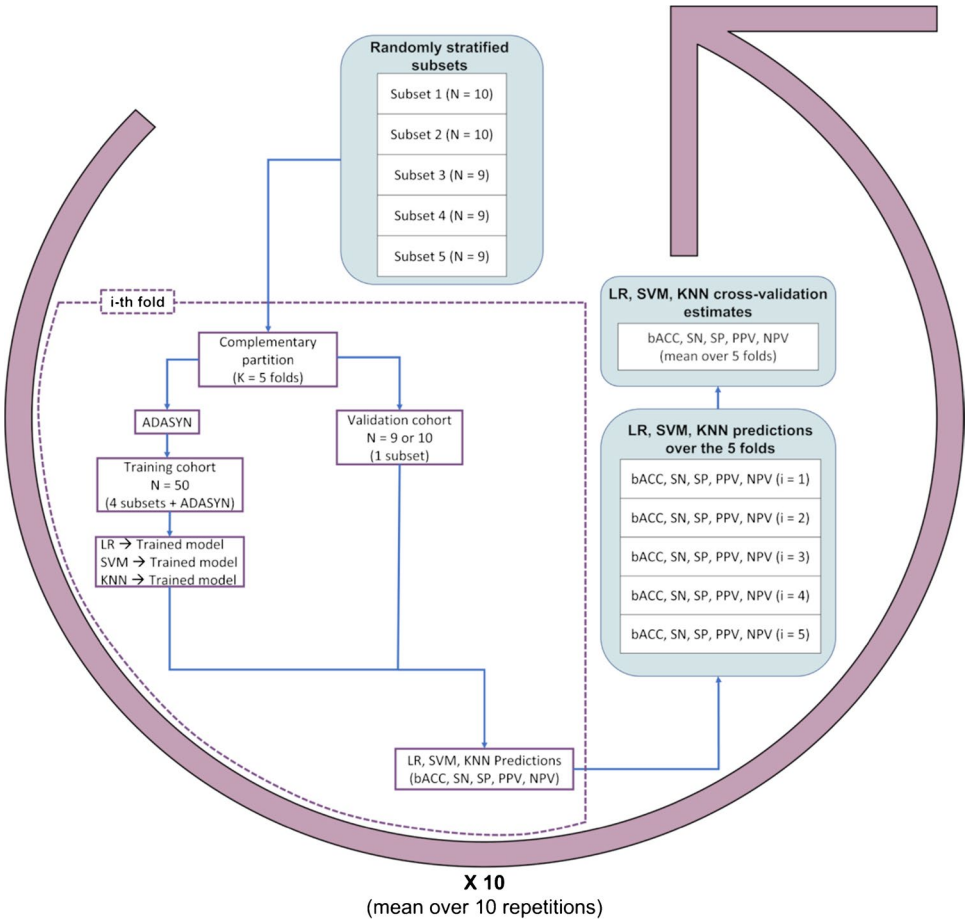

Fig. 2 Training and validation workflow of machine learning models

To assess the potential clinical utility of [^{68}Ga]Ga-PSMA-11 PET radiomics for the prediction of $_{PS}$ISUP grade, the performance of the best radiomic model developed in this study, evaluated using bACC, was compared to that of biopsy ISUP grade ($_{B}$ISUP) that, at present, represents the state of the art for PCa management.

Pearson correlation was used to assess the association between the RFs exploited in this study and those identified by Solari et al. [18]

Statistical analyses were performed with R statistical software [33], and machine learning models were generated with Python version 3.7, Python Software Foundation.

Results

Patients

The 47 patients enrolled in this study (median age: 65 years, interquartile range: 60–71 years) had median iPSA = 6.32 ng/ml (interquartile range: 4.79–7.85 ng/ml) and were grouped in 2 categories: low-grade ($N = 15$) and high-grade PCa ($N = 32$) based on $_{PS}$ISUP grade.

The scans of patients examined with [^{68}Ga]Ga-PSMA-11 PET/CT started 62 min ± 5 min (mean ± SD) after the injection of 1.94 ± 0.24 MBq kilogram body weight (mean ± SD), while the scans of patients that underwent [^{68}Ga]Ga-PSMA-11 PET/MRI started 60 min ± 7 min (mean ± SD) after the injection of 2.17 ± 0.28 MBq kilogram body weight (mean ± SD).

Radiomic features

The whole prostate was segmented as shown in Fig. 3 and 103 IBSI-compliant RFs were extracted. The 4 most relevant RFs according to mRMR algorithm (1. GLSZM—Zone Entropy, 2. Shape—Least Axis Length, 3. First Order – Minimum, 4. GLSZM—Small Area Low Grey Level Emphasis) were used for analysis.

Pearson r correlation coefficients between the RFs analyzed in this study and those identified by Solari et al. are shown in Table 2. GLSZM – Zone Entropy is strongly correlated with GLDM – Dependence Entropy, GLCM – Contrast, SUV_{mean}, 90th percentile, and GLDM – High Grey Level Emphasis ($r = 0.66$, 0.69, 0.70, 0.68, 0.75, respectively). The only shape feature employed by us, namely

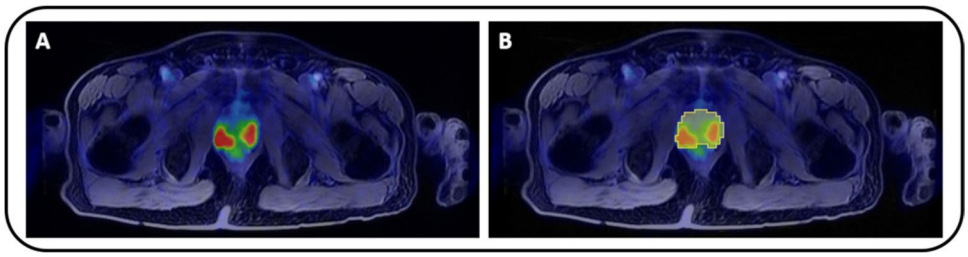

Fig. 3 A 70-year-old patient with biopsy-proven PCa. Post-surgical ISUP grade 5, with iPSA = 25 ng/ml. **A** Exemplar image of ^{68} Ga-PSMA-11 PET/MRI intraprostatic bilateral positive findings. **B** Manual segmentation of the prostate

Table 2 Comparison with the literature

	Pearson correlation, *r* (*p* value)								
GLSZM. Zone Entr	0.66 (5.25e − 07)	0.416 (0.004)	0.687 (9.96e − 08)	0.627 (2.33e − 06)	0.702 (3.78e − 08)	0.029 (0.845)	0.682 (1.29e − 07)	0.747 (1.69e − 09)	NA
Least axis len	0.274 (0.062)	− 0.009 (0.952)	0.176 (0.237)	0.224 (0.129)	0.147 (0.323)	0.63 (1.08e − 06)	0.155 (0.299)	0.172 (0.248)	NA
Min	0.162 (0.276)	0.169 (0.255)	0.025 (0.866)	− 0.007 (0.964)	− 0.144 (0.334)	− 0.285 (0.052)	0.323 (0.027)	0.217 (0.143)	NA
GLSZM. salgle	− 0.574 (2.46e − 05)	− 0.284 (0.053)	− 0.413 (0.004)	− 0.417 (0.004)	− 0.393 (0.006)	0.019 (0.897)	− 0.493 (4.4e − 04)	− 0.513 (2.2e − 04)	NA
	GLDM.dep entr	MGTDM.con	GLCM.con	NGTDM.com	SUV_{mean}	Max. 2D diam	90^{th}%	GLDM.hgle	Tot SUV

Correlation between RFs exploited in this study (first column) and those identified in a similar work [18] using RFs extracted from the whole prostate on ^{68} Ga-PSMA-11 PET images for the prediction of post-surgical Gleason grade. This analysis allows to assess whether the two radiomic signatures are associated, despite the fact they employ different features. *GLSZM*, grey level size zone matrix; *GLDM*, grey level dependence matrix; *NGTDM*, neighbouring grey tone difference matrix

Least Axis Length, correlated with the only shape RF used by Solari and colleagues (Maximum 2D diameter, $r = 0.63$).

Machine learning models

Characteristics of the implemented machine learning models after hyperparameter optimization are summarized in Supplementary Table 2, while their performance, measured with bACC (95% CI), SN, SP, PPV, and NPV, is reported in Table 3. The Kruskal–Wallis test highlighted different performances in the generated models regardless of the classifier used for analyses ($p < 0.001$ for results obtained with logistic regression, support vector machine, and KNN). Correlation matrices showing *p*-values for post hoc comparisons using the Mann–Whitney test for all model combinations are reported in Table 4.

The "radiomics baseline" model performed significantly worse than all the other investigated approaches, resulting in bACC = 70.6%, SN = 67.2%, and SP = 74% (Tables 3 and 4).

Conversely, "PET zeros" yielded bACC comparable to that of the radiomics models employing 1 RF (Tables 3 and 4). However, when Shape – Least Axis Length was removed from the analysis, its performance dropped to chance level (bACC = 60%, SN = 33.3%, SP = 86.7%). See Supplementary Table 4 for sub-analysis regarding the "PET zeros" model.

Regardless of the employed classifier, using the top 4, 3, or 2 RFs did not significantly alter the performance of the generated models ($p > 0.05$, Table 4 and Supplementary Fig. 1). On the other hand, when only GLSZM—Zone Entropy was exploited, the performance dropped resulting in the best bACC overall of 79.2%, SN = 77.7%, and SP = 80.6%.

The best model obtained using LR exploited just 2 RFs and yielded bACC = 87.5%, SN = 88.3%, and SP = 86.7%. A similar performance was found with SVM employing 3 RFs (bACC = 87.6%, SN = 88.6%, SP = 86.7%, $p = 0.889$). KNN, at most, reached bACC = 81.3%, SN = 82.6%, and SP = 80%, achieving the worst results among the considered classifiers ($p = 0.018$ as compared with the best model generated using LR, and $p = 0.008$ compared with SVM). A graphical representation of the comparison between the best models generated with the 3 different classifiers is reported in Fig. 4.

Comparison with biopsy

Twenty-four out of the 47 investigated patients were classified as low-grade ISUP (ISUP < 4) at biopsy, with the remaining 23/47 being labelled high grade. Specifically, 9/47 patients were graded ISUP 2 or less, and 15/47 ISUP grade 3, 12/47 were defined ISUP grade 4 and 11/47 ISUP, grade 5. Post-surgical examination of histopathological specimens confirmed low-grade $_{B}$ISUP in 15

Table 3 Models' performance

	Model	Balanced accuracy, % (CI)	Sensitivity, %	Specificity, %	Positive predicted value, %	Negative predicted value, %
Logistic regression	Radiomics baseline	70.6 (57.4 – 82.7)	67.2	74	85.8	55.8
	PET zeros	77.4 (63.3 – 87.7)	72.2	82.7	90.9	61.9
	1 RF	79.2 (66.7 – 89.3)	77.7	80.6	90.5	66.7
	2 RFs	87.5 (76.9 – 95.2)	88.3	86.7	93.8	81.1
	3 RFs	86.5 (74.3 – 93.8)	89.7	83.3	92.9	82.3
	4 RFs	86.3 (74.3 – 93.8)	85.9	86.7	94.4	77.6
Support vector machine	Radiomics baseline	69.6 (55.1 – 80.9)	60.5	78.7	87.3	51.8
	PET zeros	76.3 (61.9 – 86.1)	69.3	83.3	91.3	60.3
	1 RF	77.9 (64.3 – 87.7)	73.7	82	91	64.2
	2 RFs	87.3 (76.9 – 95.2)	87.9	86.7	94	80.8
	3 RFs	87.6 (76.9 – 95.2)	88.6	86.7	94	82.5
	4 RFs	87.1 (74.3 – 93.8)	88.3	86	93.8	80.9
K nearest neighbour	Radiomics baseline	64.8 (50.7 – 77.3)	61.6	67.9	82.4	48
	PET zeros	72.7 (59.7 – 84.4)	68.7	76.7	87.6	56.6
	1 RF	74.1 (59.7 – 84.4)	65.3	82.7	90.6	54.8
	2 RFs	81.3 (69.2 – 90.9)	82.6	80	91.1	69.8
	3 RFs	81.2 (69.2 – 90.9)	82.5	80	90.4	70.3
	4 RFs	81.2 (69.2 – 90.9)	82.3	80	91.5	73.4

Performance of the machine learning models on the validation sets reported using balanced accuracy (and 95% confidence interval), sensitivity, specificity, positive predictive value, and negative predictive value. The same control models (radiomics baseline and PET zeros) and radiomics machine learning models trained with 1, 2, 3, and 4 RFs were generated using three different classifiers

patients and high grade in 23 patients. All the subjects graded as ISUP 4 or more at biopsy were confirmed to be high grade at RP, but in 9 patients the biopsy underestimated the aggressiveness of PCa, reporting $_{B}$ISUP < 4 that was later revised to be high-grade PCa after surgery. Overall, the bACC of biopsy in this cohort was 85.9% (CI: 72–94%), SN = 71.9%, SP = 100%, PPV = 100%, and NPV = 62.5%, while the best radiomics model resulted in bACC = 87.6% (CI: 77–95%), SN = 88.6%, SP = 86.7%, PPV = 94%, and NPV = 82.5% (SVM, 3 RFs). See Fig. 5 for a visualization of the comparison between SVM trained with 3 RFs and biopsy for the prediction of $_{PS}$ISUP grade.

Discussion

This study investigated the role of [^{68}Ga]Ga-PSMA-11 PET radiomics for the non-invasive characterization of PCa during staging phase, demonstrating an excellent performance of radiomics machine learning models for the prediction of $_{PS}$ISUP grade (balanced accuracy up to 87.6%).

At present, histopathological analysis of prostate biopsy represents the cornerstone for the management of patients affected by PCa. However, this procedure is highly invasive, and it is known to possibly underestimate the post-surgical ISUP grade in up to 30% of cases. Therefore, there is an urgent need for novel, accurate, and non-invasive tools to characterize primary PCa.

Previous studies investigated the association between quantitative MRI parameters and $_{PS}$ISUP grade, showing a negative correlation between ADC values and PCa aggressiveness [34, 35]. Recently, also semi-quantitative PET parameters have been studied in relation to clinical and histopathological outcomes in PCa, and a positive association was found between [^{68}Ga]Ga-PSMA-11 SUV_{max} and ISUP grade [36].

In the last years, an increased interest in radiomics for PCa diagnosis and management has been observed. While the majority of the first studies were based on morphological imaging [37–40], recently, the potential of [^{68}Ga]Ga-PSMA-11 PET radiomics has strongly emerged [18, 41–43]. Solari et al. developed a SVM capable of discriminating between $_{PS}$GG 1–3, $_{PS}$GG 4, and $_{PS}$GG 5 with a bACC = 74.9% [18], while the model generated by Zamboglou et al. reached an area under the curve (AUC) = 0.84 for the prediction of post-surgical Gleason Score [42].

The combination of RFs extracted from [^{68}Ga]Ga-PSMA-11 PET images with apparent diffusion coefficient (ADC) maps obtained by MRI seems to yield even better results, as reported by Solari et al. that showed an improvement of bACC from 74.9 to 81.57% [18]. Similar results were also obtained by Papp et al.

Table 4 Comparisons between models' performances

		Mann-Whitney test p-value					
Logistic regression	**radiomics baseline**	<0.001***	<0.001***	<0.001***	0.003**	0.014*	
	PET zeros	0.001**	0.001**	0.001**	0.903		
	1 RF	0.002**	0.001**	0.001**			
	2 RFs	0.903	0.903				
	3 RFs	0.992					
	4 RFs						
	Model	**4 RFs**	**3 RFs**	**2 RFs**	**1 RF**	**PET zeros**	**radiomics baseline**
Support vector machine	**radiomics baseline**	<0.001***	<0.001***	<0.001***	0.013*	0.017*	
	PET zeros	<0.001***	<0.001***	<0.001***	0.963		
	1 RF	<0.001***	<0.001***	<0.001***			
	2 RFs	0.837	0.963				
	3 RFs	0.717					
	4 RFs						
	Model	**4 RFs**	**3 RFs**	**2 RFs**	**1 RF**	**PET zeros**	**radiomics baseline**
K nearest neighbor	**radiomics baseline**	<0.001***	<0.001***	<0.001***	<0.001***	0.001**	
	PET zeros	0.007**	0.014*	0.013*	0.903		
	1 RF	0.011*	0.018*	0.017*			
	2 RFs	0.963	0.903				
	3 RFs	0.903					
	4 RFs						
	Model	**4 RFs**	**3 RFs**	**2 RFs**	**1 RF**	**PET zeros**	**radiomics baseline**

Correlation matrices showing p value forMann-Whitney tests between all models performances(for each classifier); * p <0.05, ** $p < 0.01$, *** $p < 0.001$

that implemented a mixed ensemble learning scheme built on random forest classifiers validated by means of Monte Carlo cross-validation (accuracy = 81%, AUC = 0.86) [19].

Despite these promising results, the lack of interpretability and standardization that affects the field of radiomics are issues that currently hamper the possibility to translate the application of radiomics into clinical practice [14]. For this reason, the methodology of the present study has been developed to guarantee robust and explicable results, based on previously acquired knowledge.

In this work, radiomic features were extracted from the whole prostate; this approach is less prone to inter-reader variability and provides more robust segmentations [24]. Moreover, the risk of sampling error is avoided, as it is the possibility to analyze small lesions that are likely to be surrogates of VOI volume [44, 45]. The choice to segment the whole organ for RF extraction and subsequent analysis is also supported by practical reasons, including the fact that emerging tools for automatic segmentation of prostate on MRI, CT, and PET images are achieving excellent accuracies [46–48]. The nearly instantaneous definition of a VOI is a substantial step towards the translation of radiomics into clinical practice, as the manual contouring of VOIs is extremely time-consuming and hardly compatible with the clinical routine. Furthermore, the automatic segmentation of the prostate, paired with the extraction and analysis of a set of predefined RFs, will certainly improve the reproducibility and stability of radiomic models, thus facilitating the translation ability of radiomics.

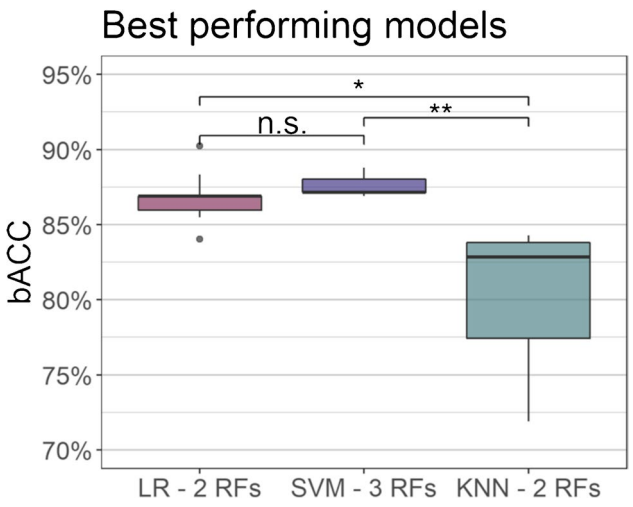

Fig. 4 Boxplots showing the balanced accuracy of the best performing models for each of the three investigated classifiers; *$p<0.05$, **$p<0.01$

In the present work, minimum redundancy maximum relevance approach was used to identify meaningful RFs for the prediction of $_{PS}$ISUP grade, and only the 4 most relevant RFs were selected to avoid overfitting. One of the main differences between this study and those previously investigating the role of [^{68}Ga]Ga-PSMA-11 PET radiomics for the characterization of primary PCa is that all analyses were repeated by employing only the best 3, 2, and finally using just the most significant RF to identify the simplest, yet most informative, model.

Interestingly, regardless of the adopted classifier, using the top 4 RFs or just a combination of GLSZM—Zone Entropy and shape—Least Axis Length gave comparable results, suggesting that all information used by the machine learning models was stored in these two features.

Furthermore, RFs used in this study were correlated with those identified as highly informative in a previous work by using the whole prostate as VOI for [^{68}Ga]Ga-PSMA-11 PET radiomics in a similar setting [18]. This analysis allowed to assess the potential association between the signatures identified in these two studies, even though they employ different features, thus increasing the interpretability of the generated signatures.

Correlation of GLSZM—Zone Entropy and shape—Least Axis Length with the features identified by Solari et al. [18] showed that GLSZM—Zone Entropy strongly correlates with GLDM – Dependence Entropy, GLCM – Contrast, NGTDM – Complexity, SUV_{mean}, 90th Percentile, and GLDM – High Gray Level Emphasis. The first three RFs describe non-uniformity within the VOI which indicates high heterogeneity and the latter higher intensities of voxel values, both very well-known markers of tumour aggressiveness [49, 50]. Maximum 2D Diameter Row was the most often selected PET feature by Solari et al. and correlated significantly with Least Axis Length used in this study ($r=0.63$, $p<0.001$).

These findings support the idea that machine learning models trained with an indicator of non-uniformity (i.e. GLSZM—Zone Entropy or GLDM – Dependence Entropy) and a shape feature representing the measure of the VOI might be enough to accurately characterize PCa in the staging phase. Furthermore, SUV_{mean}, an acknowledged [^{68}Ga]Ga-PSMA-11 PET parameter that has been largely investigated in the past [36, 50, 51], was not exploited in the present work as it was discarded by mRMR algorithm, but was found to be informative by both Papp and Solari [18, 19] and it correlated with GLSZM—Zone Entropy, the most important RF here identified. Therefore, in future works it would be interesting to also analyze the utility of SUV_{mean} combined with RFs for the prediction of $_{PS}$ISUP grade.

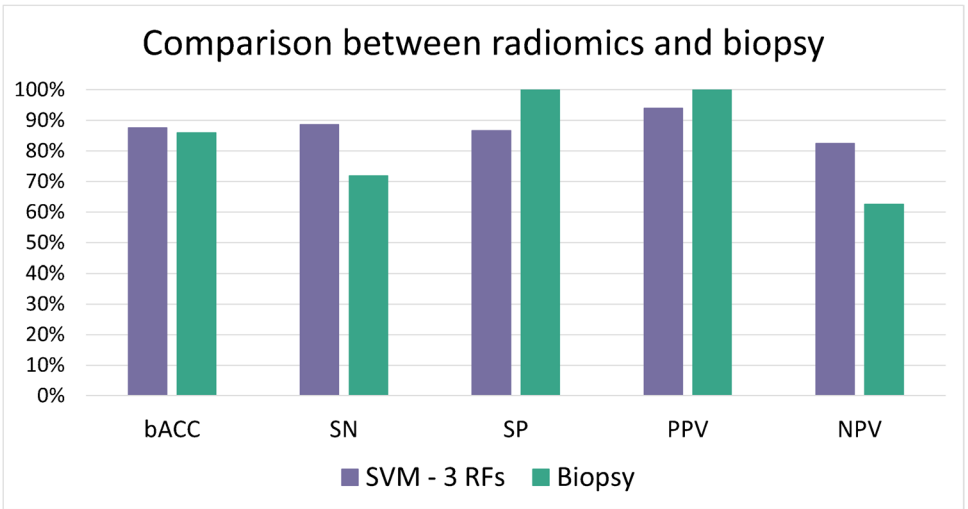

Fig. 5 Bar plot showing the performance of the best-performing radiomic model compared to biopsy for the prediction of post-surgical ISUP grade

According to a recent review, approximately thirty studies investigated the role of radiomics based on MR, CT, and PET imaging for the assessment of PCa aggressiveness by employing several different classifiers [14]. Logistic regression, support vector machine, and random forest are the most represented and the results of this study confirmed their utility (SVM and LR bACC = 87.6% and 87.5%, respectively), while the K nearest neighbour seems to underperform, reaching the highest bACC of 81.3%.

The interpretability of radiomic findings is at least as important as the results themselves, as the translation ability of a radiomic signature largely depends on its interpretability. In fact, the correct interpretation of a radiomic model allows to (1) test it; (2) simplify it, ensuring the generation of the simplest, yet most informative model; and (3) gain insights into the underlying pathophysiological mechanisms. However, an accurate study design may not be enough to ensure model interpretability in radiomics. Indeed, it has been previously reported that well-conducted studies with high radiomics quality scores [52] generated complex radiomic signatures that were later found to be surrogates of simple features, such as volume [53, 54]. For this reason, control models should always be employed in radiomics studies as recommended by the recently published joint EANM/SNMMI guidelines on radiomics in nuclear medicine [55].

To ensure that the findings of this study are indeed relying on the investigated features, two control models were generated. A "radiomics baseline" model composed by VOI volume and SUV_{max} yielded the worst results overall, resulting, at best, in a bACC of 70.6% (CI: 57.4–82.7). Another model was developed by setting to 0 all voxel values within the VOI and produced similar results to the radiomics model exploiting only GLSZM—Zone Entropy. However, as expected, when Least Axis Length was removed from the analysis, its performance dropped to chance level. These findings support the hypothesis that the radiomic signature here presented relies on more complex features than VOI volume and SUV_{max} and that shape and texture features play a complementary role in the prediction of $_{PS}$ISUP grade.

To be applied in clinical practice, radiomic signatures not only have to be interpretable and robust, but also should outperform the current gold standard for PCa management. All patients enrolled in the present study had availability of ISUP grade assessed both at biopsy and after radical prostatectomy, thus allowing a thorough comparison for the accuracy of radiomic models and biopsy. While examination of prostatic biopsy cores resulted in higher specificity and positive predictive value (both 100%), the best developed radiomic model achieved slightly higher bACC, sensitivity, and negative predictive values (87.6 (CI: 77–95%), 88.6, and 82.5% respectively, compared to 85.9 (CI: 72–94), 71.9, and 62.5% found with biopsy).

The correct identification of high-risk PCa patients (ISUP 4 and 5) is of the utmost importance as these patients have a higher risk of developing biochemical recurrence and often require a combination of multiple modalities of treatment [2]. Therefore, the higher sensitivity shown by the radiomic models as compared to biopsy potentially bears important clinical implications.

Since radiomics is a novel discipline, these first promising results have to be confirmed in large, multicentric cohorts. Then, [^{68}Ga]Ga-PSMA-11 PET radiomics could be proposed in clinical trials, paired with the analysis of histopathological specimen gathered from biopsy, for driving the clinical management of PCa patients, tailoring their treatment planning, and potentially improving their prognosis.

Some limitations should be pointed out. Firstly, this is a single-centre study based on a relatively small cohort. However, the investigated cohort was highly homogeneous and the sample was split into train and validation sets using fivefold stratified cross-validation, ensuring the presence of low-grade patients in each fold, and the whole process was repeated 10 times to provide a more robust estimate of the classifiers' performance [31].

Secondly, recent evidence suggests that a combination of RFs extracted from [^{68}Ga]Ga-PSMA-11 PET and ADC maps represents the best approach to characterize PCa and its course [18, 19, 56]. Unfortunately, this approach was not applicable to our cohort as 22 patients were examined with PET/CT and a sample of 25 patients would have not been enough to train machine learning models.

Moreover, PET images analyzed in this study were acquired with either PET/MR or PET/CT tomographs, therefore potentially increasing the inter-scanner variability. However, the use of ComBat software to harmonize the extracted RFs reduced as much as possible the variability due to the use of different scanners.

Finally, a selection bias might have occurred as only patients with biopsy-proven PCa that underwent RP were included in this study. Nonetheless, this selection was necessary to have stable ground truth for models' generation and validation.

Conclusion

In the present study, we found that radiomics, combined with machine learning, achieved slightly higher bACC compared to the histological analysis of prostate biopsy cores for the prediction of post-surgical ISUP grade, thus supporting its use for an accurate and non-invasive characterization of PCa in the staging phase. Future, multicentric studies will be needed to establish with certainty the accuracy and reproducibility of the radiomic signature here proposed.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00259-023-06187-3.

Author contribution Conceptualization and study design: SG; formal analysis: SG, CB, image acquisition and interpretation: PM, AMSG, GB, IG, TR, EP; surgical procedures: VC, AB; pathology analysis: NA, MF; data curation: SG, IN, SM; writing – original draft preparation: SG; writing – review and editing: PM, PS, LG, AB, FDC, MP; supervision: PS, FDC, MP. All the authors read and approved the submitted version of the manuscript.

Funding This research was funded by the Italian Association for Cancer Research (grant IG 2017 Id.20571) and by the Italian Ministry of Health (PE-2016–02361273); EUDRACT number: 2018–001034-18. Signa PET/MRI system (GEMS, Wakesha, WI, USA) used in the present work was purchased with funding from the Italian Ministry of Health.

Data availability The datasets generated during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval This retrospective study was approved by the Institutional Ethics Committee of IRCCS San Raffaele Scientific Institute. Informed consent was waived due to the retrospective nature of the study.

Competing interests The authors declare no competing interests.

References

1. WCRF. Prostate cancer statistics. 2020. Available from:https://www.wcrf.org/cancer-trends/prostate-cancer-statistics/. Accessed 1 Jul 2022
2. Mottet N, van den Bergh RCN, Briers E, Van den Broeck T, Cumberbatch MG, De Santis M, et al. EAU-EANM-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer—2020 Update. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. Eur Urol. 2021;79:243–62. https://linkinghub.elsevier.com/retrieve/pii/S0302283820307697.
3. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278:563–77. https://doi.org/10.1148/radiol.2015151169.
4. Wenzel M, Theissen L, Preisser F, Lauer B, Wittler C, Humke C, et al. Complication rates after TRUS guided transrectal systematic and MRI-targeted prostate biopsies in a high-risk region for antibiotic resistances. Front Surg. 2020;7. https://doi.org/10.3389/fsurg.2020.00007/full.
5. Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, et al. PI-RADS prostate imaging – reporting and data system: 2015, Version 2. Eur Urol. 2016;69:16–40. https://linkinghub.elsevier.com/retrieve/pii/S0302283815008489.
6. Ahdoot M, Wilbur AR, Reese SE, Lebastchi AH, Mehralivand S, Gomella PT, et al. MRI-targeted, systematic, and combined biopsy for prostate cancer diagnosis. N Engl J Med. 2020;382:917–28. https://doi.org/10.1056/NEJMoa1910038.
7. Ferraro DA, Becker AS, Kranzbühler B, Mebert I, Baltensperger A, Zeimpekis KG, et al. Diagnostic performance of 68Ga-PSMA-11 PET/MRI-guided biopsy in patients with suspected prostate cancer: a prospective single-center study. Eur J Nucl Med Mol Imaging. 2021;48:3315–24. https://doi.org/10.1007/s00259-021-05261-y.
8. Maurer T, Gschwend JE, Rauscher I, Souvatzoglou M, Haller B, Weirich G, et al. Diagnostic efficacy of 68 gallium-PSMA positron emission tomography compared to conventional imaging for lymph node staging of 130 consecutive patients with intermediate to high risk prostate cancer. J Urol. 2016;195:1436–43. https://doi.org/10.1016/j.juro.2015.12.025.
9. von Eyben FE, Picchio M, von Eyben R, Rhee H, Bauman G. 68Ga-Labeled prostate-specific membrane antigen ligand positron emission tomography/computed tomography for prostate cancer: a systematic review and meta-analysis. Eur Urol Focus. 2018;4:686–93. https://linkinghub.elsevier.com/retrieve/pii/S2405456916301602.
10. Donato P, Morton A, Yaxley J, Ranasinghe S, Teloken PE, Kyle S, et al. 68Ga-PSMA PET/CT better characterises localised prostate cancer after MRI and transperineal prostate biopsy: is 68Ga-PSMA PET/CT guided biopsy the future? Eur J Nucl Med Mol Imaging. 2020;47:1843–51. https://doi.org/10.1007/s00259-019-04620-0.
11. Lopci E, Saita A, Lazzeri M, Lughezzani G, Colombo P, Buffi NM, et al. 68 Ga-PSMA positron emission tomography/computerized tomography for primary diagnosis of prostate cancer in men with contraindications to or negative multiparametric magnetic resonance imaging: a prospective observational study. J Urol. 2018;200:95–103. https://doi.org/10.1016/j.juro.2018.01.079.
12. Hofman MS, Lawrentschuk N, Francis RJ, Tang C, Vela I, Thomas P, et al. Prostate-specific membrane antigen PET-CT in patients with high-risk prostate cancer before curative-intent surgery or radiotherapy (proPSMA): a prospective, randomised, multicentre study. Lancet. 2020;395:1208–16. https://linkinghub.elsevier.com/retrieve/pii/S0140673620303147.
13. Rhee H, Thomas P, Shepherd B, Gustafson S, Vela I, Russell PJ, et al. Prostate specific membrane antigen positron emission tomography may improve the diagnostic accuracy of multiparametric magnetic resonance imaging in localized prostate cancer. J Urol. 2016;196:1261–7. https://doi.org/10.1016/j.juro.2016.02.3000.
14. Ghezzo S, Bezzi C, Presotto L, Mapelli P, Bettinardi V, Savi A, et al. State of the art of radiomic analysis in the clinical management of prostate cancer: a systematic review. Crit Rev Oncol Hematol. 2022;169:103544. https://linkinghub.elsevier.com/retrieve/pii/S1040842821003310.
15. Bezzi C, Mapelli P, Presotto L, Neri I, Scifo P, Savi A, et al. Radiomics in pancreatic neuroendocrine tumors: methodological issues and clinical significance. Eur J Nucl Med Mol Imaging. 2021;48:4002–15. https://doi.org/10.1007/s00259-021-05338-8.
16. Tunali I, Gillies RJ, Schabath MB. Application of radiomics and artificial intelligence for lung cancer precision medicine. Cold Spring Harb Perspect Med. 2021;11:a039537. https://doi.org/10.1101/cshperspect.a039537.
17. Lohmann P, Galldiks N, Kocher M, Heinzel A, Filss CP, Stegmayr C, et al. Radiomics in neuro-oncology: basics, workflow, and applications. Methods. 2021;188:112–21. https://linkinghub.elsevier.com/retrieve/pii/S1046202319303172.
18. Solari EL, Gafita A, Schachoff S, Bogdanović B, Villagrán Asiares A, Amiel T, et al. The added value of PSMA PET/MR radiomics for prostate cancer staging. Eur J Nucl Med Mol Imaging. 2022;49:527–38. https://doi.org/10.1007/s00259-021-05430-z.
19. Papp L, Spielvogel CP, Grubmüller B, Grahovac M, Krajnc D, Ecsedi B, et al. Supervised machine learning enables non-invasive lesion characterization in primary prostate cancer with [68Ga]Ga-PSMA-11 PET/MRI. Eur J Nucl Med Mol Imaging. 2021;48:1795–805. https://link.springer.com/https://doi.org/10.1007/s00259-020-05140-y.
20. Fendler WP, Eiber M, Beheshti M, Bomanji J, Ceci F, Cho S, et al. 68Ga-PSMA PET/CT: Joint EANM and SNMMI procedure guideline for prostate cancer imaging: version 1.0. Eur J Nucl Med

Mol Imaging. 2017;44:1014–24. https://link.springer.com/https://doi.org/10.1007/s00259-017-3670-z.
21. Montironi R, Mazzucchelli R, van der Kwast T. Morphological assessment of radical prostatectomy specimens. A protocol with clinical relevance. Virchows Arch. 2003;442:211–7. https://doi.org/10.1007/s00428-002-0741-7.
22. Epstein JI, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. Am J Surg Pathol. 2016;40:244–52. https://journals.lww.com/00000478-201602000-00010.
23. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging. 2012;30:1323–41. https://linkinghub.elsevier.com/retrieve/pii/S0730725X12001816.
24. Han C, Ma S, Liu X, Liu Y, Li C, Zhang Y, et al. Radiomics models based on apparent diffusion coefficient maps for the prediction of high-grade prostate cancer at radical prostatectomy: comparison with preoperative biopsy. J Magn Reson Imaging. 2021;54:1892–901. https://onlinelibrary.wiley.com/doi/https://doi.org/10.1002/jmri.27565.
25. Hatt M, Majdoub M, Vallières M, Tixier F, Le Rest CC, Groheux D, et al. 18 F-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. 2015;56:38–44. https://jnm.snmjournals.org/lookup/doi/https://doi.org/10.2967/jnumed.114.144055.
26. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res. 2017;77:e104–7. https://aacrjournals.org/cancerres/article/77/21/e104/662617/Computational-Radiomics-System-to-Decode-the.
27. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology. 2020;295:328–38. https://pubs.rsna.org/doi/https://doi.org/10.1148/radiol.2020191145.
28. Orlhac F, Eertink JJ, Cottereau A-S, Zijlstra JM, Thieblemont C, Meignan M, et al. A guide to combat harmonization of imaging biomarkers in multicenter studies. J Nucl Med. 2022;63:172–9. https://jnm.snmjournals.org/lookup/doi/https://doi.org/10.2967/jnumed.121.262464.
29. Ding C, Peng H. Minimum redundancy feature selection from microarray gene expression data. J Bioinform Comput Biol. 2005;03:185–205. https://www.worldscientific.com/doi/abs/https://doi.org/10.1142/S0219720005001004.
30. Haibo He, Yang Bai, Garcia EA, Shutao Li. ADASYN: adaptive synthetic sampling approach for imbalanced learning. 2008 IEEE Int Jt Conf Neural Networks (IEEE World Congr Comput Intell. IEEE; 2008. p. 1322–8. https://ieeexplore.ieee.org/document/4633969/.
31. Kim J-H. Estimating classification error rate: repeated cross-validation, repeated hold-out and bootstrap. Comput Stat Data Anal. 2009;53:3735–45. https://linkinghub.elsevier.com/retrieve/pii/S0167947309001601.
32. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika. 1934;26:404. https://www.jstor.org/stable/2331986?origin=crossref.
33. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria; 2022. https://www.R-project.org/. Accessed 10 Jul 2022.
34. Gaur S, Harmon S, Rosenblum L, Greer MD, Mehralivand S, Coskun M, et al. Can apparent diffusion coefficient values assist PI-RADS Version 2 DWI Scoring? A correlation study using the PI-RADSv2 and International Society of Urological Pathology Systems. Am J Roentgenol. 2018;211:W33–41. https://www.ajronline.org/doi/https://doi.org/10.2214/AJR.17.18702.
35. Sahin M, Kizilay F, Guler E, Sarsik B, Harman M, Kalemci S, et al. Multıparametric prostate magnetic resonance imaging before radical prostatectomy: can it predict histopathology? Urol J. 2020;18:417–21. https://www.ncbi.nlm.nih.gov/pubmed/33037604.
36. Kwan TN, Spremo S, Teh AYM, McHarg D, Thangasamy I, Woo HH. Performance of Ga-68 PSMA PET/CT for diagnosis and grading of local prostate cancer. Prostate Int. 2021;9:107–12. https://linkinghub.elsevier.com/retrieve/pii/S2287888220300544.
37. Chaddad A, Niazi T, Probst S, Bladou F, Anidjar M, Bahoric B. Predicting Gleason score of prostate cancer patients using radiomic analysis. Front Oncol. 2018;8. https://www.frontiersin.org/article/https://doi.org/10.3389/fonc.2018.00630/full.
38. McGarry SD, Bukowy JD, Iczkowski KA, Unteriner JG, Duvnjak P, Lowman AK, et al. Gleason probability maps: a radiomics tool for mapping prostate cancer likelihood in MRI space. Tomography. 2019;5:127–34. https://www.mdpi.com/2379-139X/5/1/127.
39. Toivonen J, Montoya Perez I, Movahedi P, Merisaari H, Pesola M, Taimen P, et al. Radiomics and machine learning of multisequence multiparametric prostate MRI: towards improved non-invasive prostate cancer characterization. Penney KL, editor. PLoS One. 2019;14:e0217702. https://dx.plos.org/https://doi.org/10.1371/journal.pone.0217702.
40. Hectors SJ, Cherny M, Yadav KK, Beksaç AT, Thulasidass H, Lewis S, et al. Radiomics features measured with multiparametric magnetic resonance imaging predict prostate cancer aggressiveness. J Urol. 2019;202:498–505. https://www.jurology.com/doi/https://doi.org/10.1097/JU.0000000000000272.
41. Aksu A, Vural Topuz Ö, Yılmaz G, Çapa Kaya G, Yılmaz B. Dual time point imaging of staging PSMA PET/CT quantification; spread and radiomic analyses. Ann Nucl Med. 2022;36:310–8. https://link.springer.com/https://doi.org/10.1007/s12149-021-01705-5.
42. Zamboglou C, Carles M, Fechter T, Kiefer S, Reichel K, Fassbender TF, et al. Radiomic features from PSMA PET for non-invasive intraprostatic tumor discrimination and characterization in patients with intermediate- and high-risk prostate cancer - a comparison study with histology reference. Theranostics. 2019;9:2595–605. https://www.thno.org/v09p2595.htm.
43. Guglielmo P, Marturano F, Bettinelli A, Gregianin M, Paiusco M, Evangelista L. Additional value of PET radiomic features for the initial staging of prostate cancer: a systematic review from the literature. Cancers (Basel). 2021;13:6026. https://www.mdpi.com/2072-6694/13/23/6026.
44. Orlhac F, Soussan M, Maisonobe J-A, Garcia CA, Vanderlinden B, Buvat I. Tumor texture analysis in 18 F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. J Nucl Med. 2014;55:414–22. https://jnm.snmjournals.org/lookup/doi/https://doi.org/10.2967/jnumed.113.129858.
45. Orlhac F, Nioche C, Klyuzhin I, Rahmim A, Buvat I. Radiomics in PET imaging. PET Clin. 2021;16:597–612. https://linkinghub.elsevier.com/retrieve/pii/S1556859821000468.
46. Hatt M, Cheze le Rest C, Turzo A, Roux C, Visvikis D. A Fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. IEEE Trans Med Imaging. 2009;28:881–93. https://ieeexplore.ieee.org/document/4749328/.
47. MIM. MIM software [Internet]. https://www.mimsoftware.com. Accessed 4 Jul 2022.
48. Liu C, Gardner SJ, Wen N, Elshaikh MA, Siddiqui F, Movsas B, et al. Automatic segmentation of the prostate on CT images using deep neural networks (DNN). Int J Radiat Oncol. 2019;104:924–32. https://linkinghub.elsevier.com/retrieve/pii/S0360301619303761.
49. Andor N, Graham TA, Jansen M, Xia LC, Aktipis CA, Petritsch C, et al. Pan-cancer analysis of the extent and consequences of

intratumor heterogeneity. Nat Med. 2016;22:105–13. https://www.nature.com/articles/nm.3984.
50. Roberts MJ, Morton A, Donato P, Kyle S, Pattison DA, Thomas P, et al. 68Ga-PSMA PET/CT tumour intensity pre-operatively predicts adverse pathological outcomes and progression-free survival in localised prostate cancer. Eur J Nucl Med Mol Imaging. 2021;48:477–82. https://doi.org/10.1007/s00259-020-04944-2.
51. Koerber SA, Utzinger MT, Kratochwil C, Kesch C, Haefner MF, Katayama S, et al. 68 Ga-PSMA-11 PET/CT in newly diagnosed carcinoma of the prostate: correlation of intraprostatic PSMA uptake with several clinical parameters. J Nucl Med. 2017;58:1943–8. https://doi.org/10.2967/jnumed.117.190314.
52. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017;14:749–62. https://www.nature.com/articles/nrclinonc.2017.141.
53. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;5:4006. https://www.nature.com/articles/ncomms5006.
54. Welch ML, McIntosh C, Haibe-Kains B, Milosevic MF, Wee L, Dekker A, et al. Vulnerabilities of radiomic signature development: the need for safeguards. Radiother Oncol. 2019;130:2–9. https://linkinghub.elsevier.com/retrieve/pii/S0167814018335515.
55. Hatt M, Krizsan AK, Rahmim A, Bradshaw TJ, Costa PF, Forgacs A, et al. Joint EANM/SNMMI guideline on radiomics in nuclear medicine. Eur J Nucl Med Mol Imaging. 2023;50:352–75. https://doi.org/10.1007/s00259-022-06001-6.
56. Feliciani G, Celli M, Ferroni F, Menghi E, Azzali I, Caroli P, et al. Radiomics analysis on [68Ga]Ga-PSMA-11 PET and MRI-ADC for the prediction of prostate cancer ISUP grades: preliminary results of the BIOPSTAGE Trial. Cancers (Basel). 2022;14:1888. https://www.mdpi.com/2072-6694/14/8/1888.