Combined ^{18}F-Fluciclovine PET/MRI shows potential for detection and characterization of high-risk prostate cancer

Mattijs Elschot[1], Kirsten M. Selnæs[1,2], Elise Sandsmark[1], Brage Krüger-Stokke[1,3], Øystein Størkersen[4], Guro F. Giskeødegård[1,2], May-Britt Tessem[1], Siver A. Moestue[5], Helena Bertilsson[6,7], Tone F. Bathen[1,2]

[1]Dept. Circulation and Medical Imaging, NTNU, Norwegian University of Science and Technology, Trondheim, Norway
[2]St. Olavs Hospital, Trondheim, Norway
[3]Dept. Radiology, St. Olavs Hospital, Trondheim, Norway
[4]Dept. Pathology, St. Olavs Hospital, Trondheim, Norway
[5]Dept. Laboratory Medicine, Children's and Women's Health, NTNU, Norwegian University of Science and Technology, Trondheim, Norway
[6]Dept. Urology, St. Olavs Hospital, Trondheim, Norway
[7]Dept. Cancer Research and Molecular Medicine, NTNU, Norwegian University of Science and Technology, Trondheim, Norway

Corresponding and first author:
Mattijs Elschot
Postdoctoral fellow
NTNU
Mail MTFS*3.1313
POBox 8905
N-7491 Trondheim
T +4773598634
F +31887555491
mattijs.elschot@ntnu.no

Word count: 4999
Short running title: Fluciclovine PET/MRI for high-risk PCa
Keywords: Fluciclovine, FACBC, Prostate Cancer, PET/MRI, PLS-DA

ABSTRACT

The objective of this study is to investigate if quantitative imaging features derived from combined ^{18}F-Fluciclovine Positron Emission Tomograpy (PET) / multiparametric Magnetic Resonance Imaging (MRI) show potential for detection and characterization of primary prostate cancer. **Methods:** Twenty-eight (28) patients diagnosed with high-risk prostate cancer underwent simultaneous ^{18}F-Fluciclovine PET/MRI before radical prostatectomy. Volumes-of-interest (VOIs) of prostate tumors, benign prostatic hyperplasia (BPH) nodules, prostatitis, and healthy tissue were delineated on T2-weighted images using histology as a reference. Tumor VOIs were marked as high-grade (≥ Gleason Grade group 3) or not. MRI and PET features were extracted on the voxel and VOI-level. Partial least-squared discriminant analysis (PLS-DA) with double leave-one-patient-out cross validation was performed to classify tumor from benign tissue (BPH, prostatitis, healthy tissue) and high-grade tumor from other tissue (low-grade tumor, benign tissue). The performances of PET, MRI, and combined PET/MRI features were compared using the area under the receiver operating characteristic curve (AUC). **Results:** Voxel and VOI features were extracted from 40 tumor (26 high-grade), 36 BPH, 6 prostatitis, and 37 healthy tissue VOIs. PET/MRI performed better than MRI and PET for classification of tumor vs benign tissue (voxel: AUC 87%, 81%, and 83%; VOI: AUC 96%, 93%, and 93%, respectively) and high-grade tumor vs other tissue (voxel: AUC 85%, 79%, and 81%; VOI: AUC 93%, 93%, and 91%, respectively). T2-weighted MRI, diffusion-weighted MRI and PET features were most important for classification. **Conclusion:** Combined ^{18}F-Fluciclovine PET/multiparametric MRI

shows potential for improving detection and characterization of high-risk prostate cancer, in comparison to MRI and PET alone.

INTRODUCTION

Prostate cancer is the most frequently detected type of cancer in men and consititutes a major healthcare problem in developed countries (*1*). Medical imaging plays an increasingly important role in the management of prostate cancer and is, amongst others, used for diagnosis and stratification of indolent and clinically significant disease (*2*), guiding prostate biopsies (*3*), and targeting localized therapy (*4*). In Norway, multiparametric MRI, i.e. the combination of MR images with distinct contrasts, is currently the first diagnostic tool for patients suspected for prostate cancer based on digital rectal examination and/or prostate specific antigen blood testing (*5*). However, multiparametric MRI leaves ample room for improvement; despite the Prostate Imaging – Reporting and Data System guidelines (*6*), there is a highly variable accuracy for detection of clinically significant disease (*7*), while grading of tumor aggressiveness suffers from overlapping values between Gleason scores (*8*).

PET imaging with the synthetic amino acid analog radiotracer anti-1-amino-3-^{18}F-fluorocyclobutane-1-carboxylic acid (^{18}F-Fluciclovine, also known as ^{18}F-FACBC) has recently been approved by the Food and Drug Administration and European Medicines Agency for detection of recurrent prostate cancer in patients with biochemical relapse following initial treatment with curative intent (*9-11*). For primary prostate cancer, the metabolic information derived from ^{18}F-Fluciclovine PET images could be useful for detection and characterization of localized tumors (*12-14*) and may be complementary to multiparametric MRI (*12*).

In comparison to sequential multiparametric MRI and PET/CT examinations, simultaneous PET/MRI has the advantages of shorter (cumulative) scan times, simpler patient logistics, reduced radiation exposure from omitting the CT, and an intrinsic alignment of the PET and MR images. We have previously presented an optimized imaging protocol for simultaneous ^{18}F-Fluciclovine PET/MRI that maximizes the diagnostic information obtained from the PET images (*14*). The objective of this study is to investigate if simultaneous ^{18}F-Fluciclovine PET/multiparametric MRI has the potential to improve detection and characterization of primary prostate cancer. For this purpose, we built and evaluated statistical models to assess which combination of MRI and PET-derived imaging features best discriminates between histologically verified malignant and benign tissue. We also show that these models can be used to create cancer probability maps, which could be regarded a 'visual summary' of the combined PET and MRI information to support image interpretation in future clinical practice.

MATERIALS AND METHODS

Patients

Patients classified as high-risk according to modified D'Amico criteria (prostate specific antigen > 20 ng/ml and/or clinical stage ≥ cT3a and/or Gleason score ≥ 8) scheduled for robot-assisted radical prostatectomy with extended pelvic lymph node dissection were recruited for a prospective study investigating the merit of combined ^{18}F-Fluciclovine PET/MRI for loco-regional staging of primary prostate cancer (ClinicalTrials.gov; identifier NCT02076503). The study was approved by our institution (St. Olavs Hospital, Trondheim University Hospital) and the Regional Committee for Medical and Health Research Ethics, Central Norway. All patients gave written informed consent before enrollment. In this work we performed a retrospective analysis of these prospectively collected data, focusing on the detection and characterization of prostate tumors.

Imaging

Patients underwent a PET/MRI exam on a 3 T Biograph mMR scanner (Siemens Medical Systems, Erlangen, Germany) prior to surgery, as previously described in (*14*). In this work, we used the T2-weighted (T2W), diffusion-weighted (DW) and dynamic contrast enhanced (DCE) MR images, which together constitute the clinical multiparametric MRI examination for T-staging, as well as the simultaneously acquired ^{18}F-Fluciclovine PET images. All images were post-processed to obtain parametric maps from which quantitative imaging features were obtained. For this purpose, the T2W images were intensity normalized (nT2W)

to the levator ani muscle. The DW images were corrected for geometric distortion (*15*), after which apparent diffusion coefficient (ADC) maps were calculated using a mono-exponential decay model (including b=50/400/800 s/mm^2). DW images at b=800 s/mm^2 (b800) were also used for further analysis. Maps of the volume transfer constant (K^{trans}), fractional volume of the extravascular extracellular space (v_e), and the fractional plasma volume (v_p) were calculated from motion-corrected DCE images using the extended Tofts' model with a population-based arterial input function (*16,17*). PET data from 5-10, 18-23, and 33-38 minutes post-injection were reconstructed to standardized uptake value (SUV) maps (SUV^{5-10}, SUV^{18-23}, and SUV^{33-38}, respectively) using a manufacturer-provided algorithm (Siemens HDPET, 3 iterations, 21 subsets, 4 mm full width at half maximum Gaussian filter). These time windows were previously shown to have potential for prostate cancer diagnosis (*14*). All images were co-registered and resampled to T2W image space using a multi-resolution rigid registration scheme based on mutual information in elastix (*18*).

VOIs and Feature Extraction

Histopathology-matched VOIs formed the basis for image feature extraction. In short, a pathologist specialized in uropathology delineated cancer foci, benign prostatic hyperplasia (BPH) nodules, and regions of prostatitis on hematoxylin and eosin stained whole mount histology slides of the excised prostate gland (*14*). Cancer grade was described according to the Gleason Scoring system (*19*). The histology slides were then spatially matched to the T2W images based on

anatomical landmarks and served as a reference for the delineation of three-dimensional VOIs of tumor, BPH, prostatitis and healthy tissue. The latter VOIs were outlined in both the peripheral zone and central gland (transition and central zone), if possible, and placed as far away as possible from the lesions to minimize partial volume effects. Tumor VOIs were assigned to be high-grade (≥ Gleason Score 4+3, i.e. ≥ Gleason Grade group 3) or low-grade (*20*). Furthermore, the volume (mL) was recorded for each VOI and an ordinal confidence score (1 not confident; 2 reasonably confident; 3 confident) was assigned, which expressed how well the delineated VOI resembled histology. VOIs with volume < 0.5 mL and/or confidence score < 2 were excluded from further analysis.

From each of the remaining VOIs, imaging features were extracted on the voxel and VOI-level. The voxel-level features were collected in a 9 x N_{voxels} matrix, where each of the N_{voxels} rows contained the image intensities of the 9 parametric maps (nT2W, b800, ADC, K^{trans}, v_e, v_p, SUV^{5-10}, SUV^{18-23}, and SUV^{33-38}) for a single voxel, with N_{voxels} the total number of voxels in the VOIs. The VOI-level features were collected in a 99 x N_{vois} matrix, where each of the N_{vois} rows contained 11 first order statistical features (mean, standard deviation, minimum, 10% percentile, 25% percentile, median, 75% percentile, 90% percentile, maximum, skewness, and kurtosis) of the 9 parametric maps (9 x 11 = 99), with N_{vois} the total number of VOIs. The imaging features were then used as input to the classification algorithm described in the next section.

Classification

Partial least-squares discriminant analysis (PLS-DA) is a classification method that calculates latent variables (LVs) to maximize the covariance between the input variables (in this case: the imaging features) and the response variable (the class) (*21*). In this way, large data sets with multiple variables can be visualized and interpreted using only a few dimensions (LVs). The method is well-suited for dealing with highly co-linear variables and provides valuable insight into the constructed models via scores plots, which show each sample's position in the new coordinate system defined by the LVs, and loadings plots, which show the importance of the original variables for defining the this coordinate system.

We performed PLS-DA on both the voxel- and VOI-level to evaluate the performance of a given set of imaging features for discriminating between tumor and benign tissue (i.e. BPH + prostatitis + healthy tissue). Since Gleason Grade group 3-5 tumors are associated with a significantly poorer prognosis than Gleason Grade group 1-2 tumors (*20*), we also investigated the potential discrimination between high-grade tumor and other tissue (i.e. low-grade tumor + BPH + prostatitis + healthy tissue). All imaging features were scaled to have zero mean and unit standard deviation. In order to avoid over-optimistic results, model overfitting was counteracted by nesting the PLS-DA algorithm in a wrapper for double leave-one-patient-out cross-validation. In the inner cross-validation loop of this wrapper, the number of LVs resulting in the model with the lowest average classification error in the training set (N-1 patients) was selected, with N the total number of patients.

Additionally, the variable importance in the projection (VIP) scores (*22*) were recorded as measures of relative feature importance. The trained model was then applied in the outer cross-validation loop of the wrapper to predict the class probabilities of the voxels/VOIs in the test set (1 patient), thus achieving independent classification accuracies for data that were not used to build the PLS-DA model. This whole procedure was repeated N times, giving predicted class probabilities for every voxel/VOI in the data set, which were finally compared to the true class.

Three sets of imaging features were initially evaluated and compared: MRI (nT2W + b800 + ADC + K^{trans} + v_e + v_p), PET (SUV^{5-10} + SUV^{18-23} + SUV^{33-38}), and PET/MRI (nT2W + b800 + ADC + K^{trans} + v_e + v_p + SUV^{5-10} + SUV^{18-23} + SUV^{33-38}). Based on the analysis of the VIP scores, a PET/T2W+DW MRI feature set (nT2 + b800 + ADC + SUV^{5-10} + SUV^{18-23} + SUV^{33-38}) was also evaluated and compared to the complete PET/MRI feature set. The double cross-validated performances of MRI, PET, PET/MRI and PET/T2W+DW MRI were assessed using ROC curve analysis. The AUC was used as the figure of merit and the optimal sensitivity and specificity were determined as the point on the curve closest to (1,1). Finally, scores and loading plots were constructed of the PET/MRI and PET/T2W+DW MRI models for interpretation of the results. For this purpose, the models were rebuilt using the data from all patients, i.e. without cross-validation.

Statistical Analysis

Descriptive statistics were presented as mean and standard deviation or as median and range. Statistical differences in feature intensity between classes were calculated using linear mixed effects models with the patient number as a random effect on the intercept. The resulting p-values were corrected for multiple testing by Benjamini–Hochberg correction. Permutation testing, i.e. random shuffling of the class labels (n=1000), was performed to examine whether the achieved PLS-DA predictions were significantly different from chance. Statistical differences in AUC between feature sets were evaluated using DeLong's method for comparing correlated AUCs (*23*). P-values < 0.05 were considered statistically significant for all tests. Unless indicated otherwise, MATLAB 9.0 (The MathWorks Inc., Natick, MA) was used for image processing and univariate statistics. The MATLAB toolbox PLS_toolbox 8.2.1 (Eigenvector Research, Inc., Manson, WA, USA) was used for PLS-DA classification.

RESULTS

Patients and Imaging

Twenty-eight (28) patients (median (range) age 66 (55-72) years) were included in the study. Median prostate specific antigen was 14.6 (3.7-56.9) ng/mL, median biopsy Gleason score was 8 (7-9), and clinical stage ranged from cT2b to cT3b. The median time between the PET/MRI examination and surgery was 8 (5–32) days and the median (range) administered activity was 327 (283-384) MBq.

Two hundred seventeen (217) VOIs were delineated on the T2W images. Of these, 98 were excluded based on size and/or confidence criteria, leaving a total of 119 VOIs for analysis (Table 1). The total number of voxels in these VOIs was 512717.

Feature Extraction

An example of the 9 parametric maps calculated from the combined PET/MRI exam is shown in Figure 1. A total of 9 x 512717 voxel-level features and 99 x 119 VOI-level features were extracted from the 119 VOIs. An overview of the mean feature intensities at the VOI-level for each class is provided in Table 2. ADC, $SUV^{18\text{-}23}$, and $SUV^{33\text{-}38}$ were the only features that consistently showed significant differences between malignant and benign tissue, indicating that these could be important for PLS-DA classification.

Classification

The classification performances of the different image feature sets are presented in Table 3. All PLS-DA models performed significantly better than chance as assessed by permutation testing ($p<0.001$). At the voxel-level, the imaging features from combined PET/MRI performed significantly better (i.e. higher AUCs) than those from MRI and PET alone, both for discriminating tumor vs benign tissue ($p < 0.001$ and $p < 0.001$, respectively) and for discriminating high-grade tumor vs all other voxels ($p < 0.001$ and $p < 0.001$, respectively). At the VOI-level, the highest AUCs were also observed for combined PET/MRI, but these were not always significantly different from those of MRI and PET alone (tumor vs benign tissue: $p = 0.140$ and $p= 0.049$, respectively; high-grade tumor vs all other tissue: $p = 0.831$ and $p = 0.252$, respectively).

The VIP scores shown in Figure 2 indicate that features from T2W MRI, DW MRI, and PET were consistently more important for classification than those from DCE MRI. This information encouraged us to build and evaluate an additional model based on the combination of these three modalities. As shown in Table 4 and Figure 3, the PET/T2W+DW MRI features indeed performed similar to the PET/MRI features for all classification tasks. These results suggest that, for the quantitative analysis performed in this study, DCE MRI does not add much value to PET, T2W MRI, and DW MRI. This finding is further illustrated by the scores and loadings plots of the orthogonalized PLS-DA models built on the PET/MRI features from all 28 patients, as shown in Figure 4 (VOI-level only). Tumors/high-grade tumors were

most importantly associated with lower ADC values and higher late-window SUV values than benign/other tissue. The scores and loading plots for the PET/T2W+DW MRI model, which are provided as Supplemental Figure 1, show similar patterns.

An example of how the results of this study can be translated to clinical pratice is provided in Figure 5, which shows the voxel-wise cancer probability map that corresponds to the parameter maps in Figure 1. This cancer probability map was obtained by applying the cross-validated PET/T2W+DW MRI model, which was trained on the other 27 patients, to the prostate voxels of the shown patient, followed by back-projection into image space. The highest cancer probabilities were found to accurately coincide with the location of a peripheral zone tumor as verified by whole-mount histology. These tumor probability maps could be automatically calculated and presented alongside the traditional images as a supporting tool to pinpoint the physician to regions with high likelihood of cancer when interpreting the PET/MR images.

DISCUSSION

We have previously shown that ^{18}F-Fluciclovine PET images alone may be useful for the assessment of localized prostate cancer (*14*). The objective of this study was to investigate the potential of combined ^{18}F-Fluciclovine PET/multiparametric MRI for improving the detection and characterization of primary prostate cancer, in comparison to MRI and PET alone. We built and cross-validated PLS-DA models which allowed us to evaluate the performance of combined MRI and PET-derived imaging features for discriminating between histologically verified malignant and benign tissue. We found that imaging features from PET/MRI better discriminated between tumor and benign tissue, and between high-grade tumor and other tissue (including low-grade tumor), than features from MRI or PET alone. Features from T2W MRI, DW MRI and PET were more important for classification than those from DCE MRI, indicating that the latter may be omitted for this purpose. We also showed that the PLS-DA model can be used to generate images of the cancer probability distribution, which could be interpreted by the physician alongside the traditional MRI and PET images in future clinical practice.

The results of this double cross-validated study confirm the results of our previously published analysis on the same patient cohort (*14*), which also showed that malignant prostate tissue is associated with high SUVs on late-window ^{18}F-Fluciclovine PET images. As expected (*6*), we observed significantly lower ADC and nT2W values in malignant tissue than in benign tissue, both of which were important for classification. In contrast, the quantitative DCE MRI parameters

played a very limited role in the classification tasks. This observation is in line with work from De Vischere et al (*24*) and the most recent Prostate Imaging – Reporting and Data System guidelines (*6*), where DCE MRI is only recommended as an adjuvant tool to DW MRI for potential upgrading of score-3 peripheral zone lesions.

The observed absolute differences in SUV between tumors and benign lesions were relatively small (e.g. 2.5 vs 2.0 for BPH, $p<0.05$), which indicates that PET images alone may not be sufficient for evaluation of primary prostate cancer in clinical practice. Nevertheless, the discriminative power of PET was strong enough to improve MRI-based classification by approximately 5 percentage points when combined. The benefit of such a multivariate approach to cancer probability mapping could be higher in patient cohorts where MRI alone has a less impressive performance. This effect was more apparent at the voxel-level than at the VOI-level, which may be caused by the difference in statistical power. In a study combining sequential multiparametric MRI and ^{18}F-Fluciclovine PET/CT, Turkbey et al found comparable results: the combination of PET and MRI features showed a higher positive predictive value for distinguishing tumor-containing from non-tumor-containing prostate sectors than PET or MRI alone (*12*). Similarly, Eiber et al found that the combination of ^{68}Ga-PSMA PET and multiparametric MRI was more accurate for localization of primary prostate cancer than MRI alone (*25*). They reported a high tumor-to-benign tissue SUV ratio of 4.48, which could be a potential benefit of ^{68}Ga-PSMA in comparison to ^{18}F-Fluciclovine.

This study has some limitations. The patient cohort was relatively small, but homogenous, as it consisted of high-risk prostate cancer patients only. We avoided model overfitting by using a double cross-validation scheme, but our results may not be extended to low and intermediate-risk patients without proper validation. Furthermore, we realize that the results of this quantitative analysis, in which we retrospectively identified the VOIs on the T2W MR images while using histology as a reference, may not be directly translatable to radiological reading as performed in clinical practice. Another implication of the retrospective approach was the possible introduction of a positive bias towards T2W MRI, as the VOIs were delineated on these images. Nevertheless, ^{18}F-Fluciclovine PET was still shown to add value to multiparametric MRI. Validation of the clinical diagnostic value of combined ^{18}F-Fluciclovine PET/MRI in a broader patient population will be part of future prospective studies.

Including ^{18}F-Fluciclovine PET to multiparamteric MRI for diagnosis of primary prostate cancer is clinically feasible on an integrated PET/MRI system and does not increase the total scan time of the examination (*14*). However, the additional costs associated with ^{18}F-Fluciclovine PET may hinder its use for routine diagnostic imaging in the overall prostate cancer population. Nevertheless, the high diagnostic potential of combined ^{18}F-Fluciclovine PET/MRI could be fully exploited in selected cases, such as for diagnosis of patients at high risk for lymph node metastases and for planning of targeted prostate biopsies in highly-suspected patients with previous negative biopsies.

CONCLUSION

The combination of imaging features from T2-weighted MRI, diffusion-weighted MRI, and ^{18}F-Fluciclovine PET shows excellent discriminative performance between malignant and benign tissue and may improve the detection and characterization of high-risk prostate cancer in selected cases.

DISCLOSURE

This study was funded by The Norwegian Cancer Society (grant number 100792).

ACKNOWLEDGEMENTS

N/A

REFERENCES

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin.* 2015;65:87-108.
2. Mottet N, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG Guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent. *Eur Urol.* 2017;71:618-629.
3. Hansen NL, Kesch C, Barrett T, et al. Multicentre evaluation of targeted and systematic biopsies using magnetic resonance and ultrasound image-fusion guided transperineal prostate biopsy in patients with a previous negative biopsy. *BJU Int.* November 10, 2016 [Epub ahead of print].
4. Vigneault E, Mbodji K, Racine LG, et al. Image-guided high-dose-rate brachytherapy boost to the dominant intraprostatic lesion using multiparametric magnetic resonance imaging including spectroscopy: results of a prospective study. *Brachytherapy.* 2016;15:746-751.
5. Ahmed HU, El-Shater Bosaily A, Brown LC, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet.* 2017;389:815-822.
6. Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS prostate imaging - reporting and data system: 2015, version 2. *Eur Urol.* 2016;69:16-40.
7. Futterer JJ, Briganti A, De Visschere P, et al. Can clinically significant prostate cancer be detected with multiparametric magnetic resonance imaging? A systematic review of the literature. *Eur Urol.* 2015;68:1045-1053.
8. Kobus T, Vos PC, Hambrock T, et al. Prostate cancer aggressiveness: in vivo assessment of MR spectroscopy and diffusion-weighted imaging at 3 T. *Radiology.* 2012;265:457-467.
9. Odewole OA, Tade FI, Nieh PT, et al. Recurrent prostate cancer detection with anti-3-[(18)F]FACBC PET/CT: comparison with CT. *Eur J Nucl Med Mol Imaging.* 2016;43:1773-1783.
10. Schuster DM, Nieh PT, Jani AB, et al. Anti-3-[(18)F]FACBC positron emission tomography-computerized tomography and (111)In-capromab pendetide single photon emission computerized tomography-computerized tomography for recurrent prostate carcinoma: results of a prospective clinical trial. *J Urol.* 2014;191:1446-1453.
11. Nanni C, Zanoni L, Pultrone C, et al. (18)F-FACBC (anti1-amino-3-(18)F-fluorocyclobutane-1-carboxylic acid) versus (11)C-choline PET/CT in prostate cancer relapse: results of a prospective trial. *Eur J Nucl Med Mol Imaging.* 2016;43:1601-10.
12. Turkbey B, Mena E, Shih J, et al. Localized prostate cancer detection with 18F FACBC PET/CT: comparison with MR imaging and histopathologic analysis. *Radiology.* 2014;270:849-856.
13. Schuster DM, Taleghani PA, Nieh PT, et al. Characterization of primary prostate carcinoma by anti-1-amino-2-[(18)F] -fluorocyclobutane-1-

carboxylic acid (anti-3-[(18)F] FACBC) uptake. *Am J Nucl Med Mol Imaging.* 2013;3:85-96.
14. Elschot M, Selnaes KM, Sandsmark E, et al. A PET/MRI study towards finding the optimal [18F]Fluciclovine PET protocol for detection and characterisation of primary prostate cancer. *Eur J Nucl Med Mol Imaging.* 2017;44:695-703.
15. Holland D, Kuperman JM, Dale AM. Efficient correction of inhomogeneous static magnetic field-induced distortion in Echo Planar Imaging. *Neuroimage.* 2010;50:175-183.
16. Tofts PS, Brix G, Buckley DL, et al. Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusable tracer: standardized quantities and symbols. *J Magn Reson Imaging.* 1999;10:223-232.
17. Parker GJ, Roberts C, Macdonald A, et al. Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI. *Magn Reson Med.* 2006;56:993-1000.
18. Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging.* 2010;29:196-205.
19. Epstein JI. An update of the Gleason grading system. *J Urol.* 2010;183:433-440.
20. Epstein JI, Zelefsky MJ, Sjoberg DD, et al. A contemporary prostate cancer grading system: a validated alternative to the Gleason score. *Eur Urol.* 2016;69:428-435.
21. Ståhle L, Wold S. Partial least squares analysis with cross-validation for the two-class problem: A Monte Carlo study. *J Chemom.* 1987;1:185-196.
22. Chong I-G, Jun C-H. Performance of some variable selection methods when multicollinearity is present. *Chemometr Intell Lab Syst.* 2005;78:103-112.
23. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44:837-845.
24. De Visschere P, Lumen N, Ost P, Decaestecker K, Pattyn E, Villeirs G. Dynamic contrast-enhanced imaging has limited added value over T2-weighted imaging and diffusion-weighted imaging when using PI-RADSv2 for diagnosis of clinically significant prostate cancer in patients with elevated PSA. *Clin Radiol.* 2017;72:23-32.
25. Eiber M, Weirich G, Holzapfel K, et al. Simultaneous Ga-PSMA HBED-CC PET/MRI improves the localization of primary prostate cancer. *Eur Urol.* 2016;70:829-836.

FIGURES

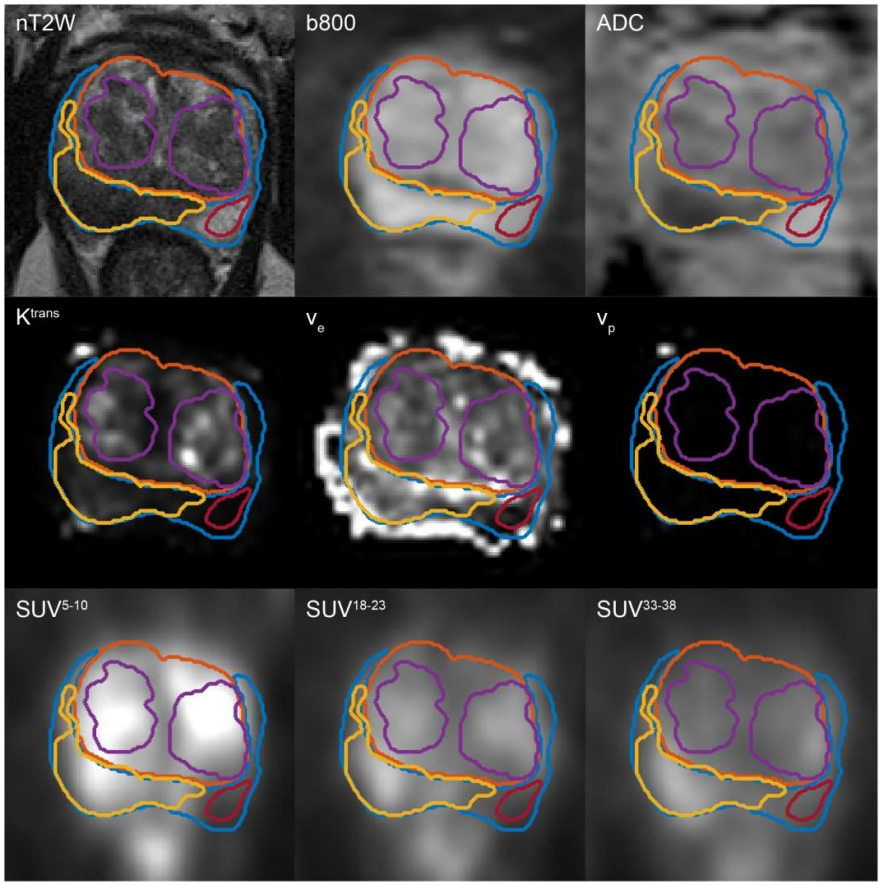


Figure 1 Collage of the central slices through the nine parametric maps of a 64 year-old patient. The contour of a tumor VOI is indicated in yellow, two benign prostatic hyperplasia nodules in purple, and a healthy tissue VOI in red. The contours of the central gland and peripheral zone are visualized in orange and dark blue, respectively.

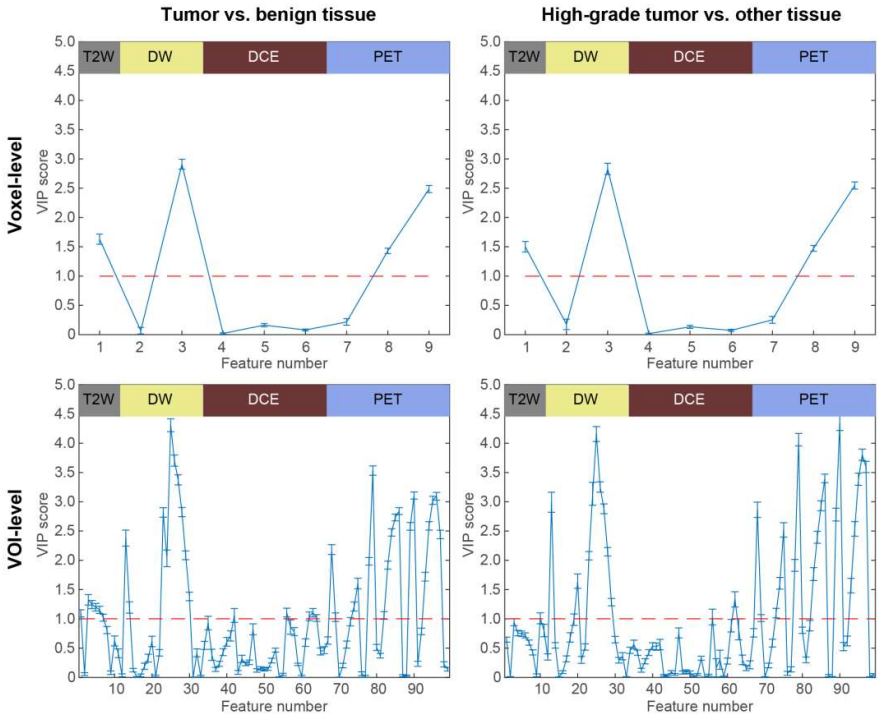


Figure 2 Overview of the VIP scores for all classification tasks, showing which variables are important for discriminating between the groups. The features from DCE MRI were consistently least important for classification. Voxel-level features 1-9: nT2W, b800, ADC, K^{trans}, v_e, v_p, SUV^{5-10}, SUV^{18-23}; and SUV^{33-38}. VOI-level features 1-11: mean, standard deviation, minimum, 10% percentile, 25% percentile, median, 75% percentile, 90% percentile, maximum, skewness, and kurtosis, respectively, for nT2W; 12-22: for b800; 23-33: for ADC; 34-44: for K^{trans}; 45-55: for V_e; 56-66: for v_p; 67-77: for SUV^{5-10}; 78-88: for SUV^{18-23}; 89-99: for SUV^{33-38}

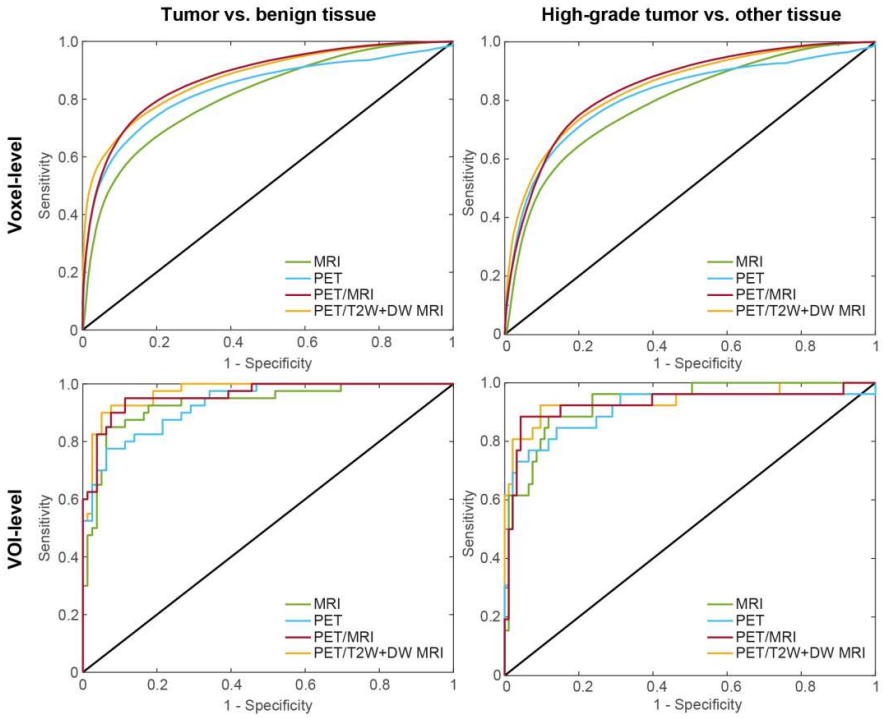


Figure 3 Overview of the ROC curves for all classification tasks. PET/MRI performed consistently better than MRI and PET, whereas PET/T2W+DW MRI and PET/MRI performed similarly well.

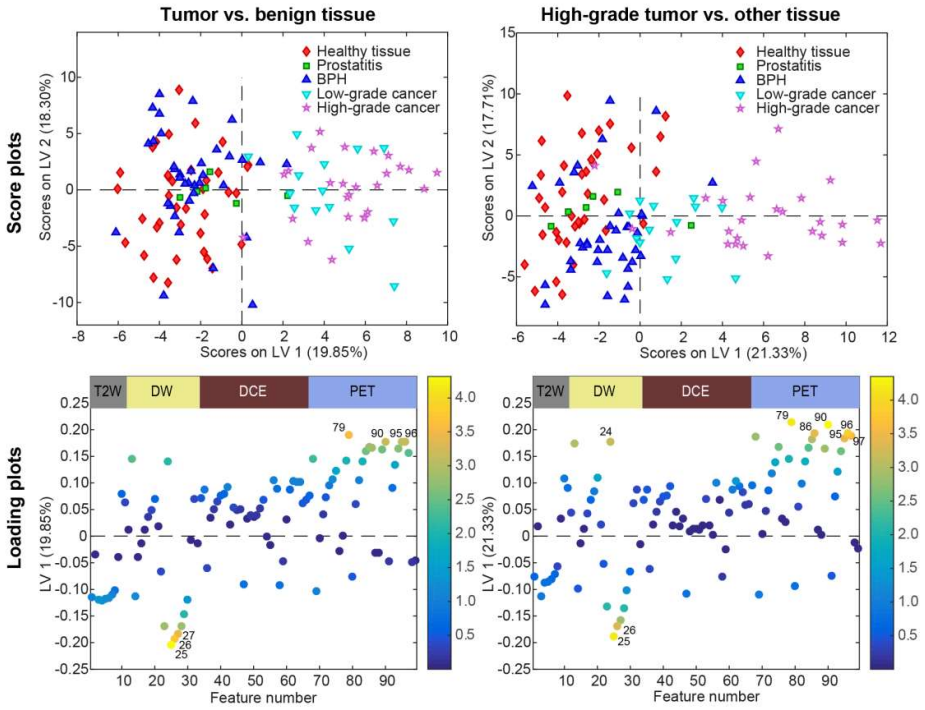


Figure 4 Overview of the scores and loadings plots of the PET/MRI models for the VOI-level classification tasks. The scores plots (upper row) show the excellent discrimination between tumor and benign tissue (left) and between high-grade tumor and other tissue (right) on LV1. The loadings (bottom row) are colored by their VIP score, showing that low ADC values (labels 25, 26, 27), high SUV values (labels 86, 95, 96, 97), and high standard deviations (labels 24, 79, 90) were the most important features characterizing the malignant classes. Features 1-11: mean, standard deviation, minimum, 10% percentile, 25% percentile, median, 75% percentile, 90% percentile, maximum, skewness, and kurtosis, respectively, for nT2W; 12-22: for b800; 23-33: for ADC; 34-44: for K^{trans}; 45-55: for V_e; 56-66: for v_p; 67-77: for SUV^{5-10}; 78-88: for SUV^{18-23}; 89-99: for SUV^{33-38}

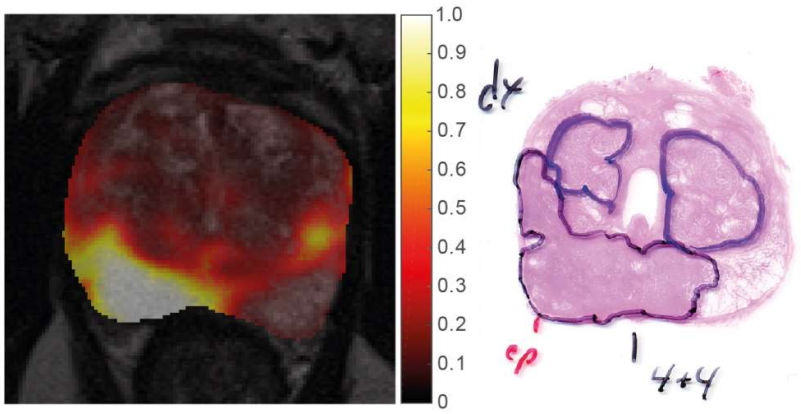


Figure 5. The central slice through the tumor probability map of a 64 year-old patient, fused with the T2-weighted image for anatomical reference (left). The region with the highest tumor probability values corresponds to a peripheral zone tumor with Gleason Score 4+4, as verified by histopathology (right).

TABLES

Table 1. VOI characteristics presented as median (range) unless indicated otherwise.

	Whole prostate	Peripheral zone	Central gland
Tumor n	40	34	6
size in mL	2.3 (0.5 – 31.9)	3.1 (0.5 – 31.9)	1.7 (1.1 – 19.9)
Gleason Score	7 (6 – 9)	7 (6 – 9)	6 (6 – 7)
High-grade n (%)	26 (65%)	24 (71%)	2 (33%)
Low-grade n (%)	14 (35%)	10 (29%)	4 (67%)
BPH n	36	4	32
size in mL	2.1 (0.5 – 11.7)	3.5 (1.1 – 7.3)	1.6 (0.5 – 11.7)
Prostatitis n	6	4	2
size in mL	1.2 (0.6 – 5.4)	1.9 (0.6 – 5.4)	0.6 (0.6 – 0.6)
Healthy n	37	18	19
size in mL	0.8 (0.5 – 1.9)	0.8 (0.5 – 1.9)	0.7 (0.5 – 1.6)

Table 2. Feature intensities (mean of VOI) presented as mean (standard deviation) over all VOIs belonging to the indicated classes.

	Tumor	BPH	Prostatitis	Healthy
T2W				
nT2W	3537 (666)	4369 (1311)*	3675 (533)	5139 (1872)*
DWI (x 10^{-3})				
b800	75 (26)	70 (24)*	74 (12)	72 (29)*
ADC	1122 (129)	1346 (163)*	1340 (124)*	1520 (257)*
DCE (x 10^{-3})				
K^{trans}	176 (75)	163 (82)	143 (53)	107 (54)*
v_e	313 (82)	269 (103)	299 (85)	284 (106)
v_p	3 (3)	2 (2)*	2 (2)	1 (1)*
PET (x 10^{-3})				
SUV^{5-10}	3093 (827)	2960 (830)	2227 (510)*	2315 (658)*
SUV^{18-23}	2762 (653)	2313 (556)*	2005 (250)*	1892 (403)*
SUV^{33-38}	2530 (554)	2036 (440)*	1673 (218)*	1765 (309)*

* significantly different from tumor tissue after Benjamini Hochberg correction ($p < 0.05$)

Table 3. Performance measures of the MRI, PET and PET/MRI feature sets for discriminating between tumor and benign tissue and between high-grade tumor and other tissue

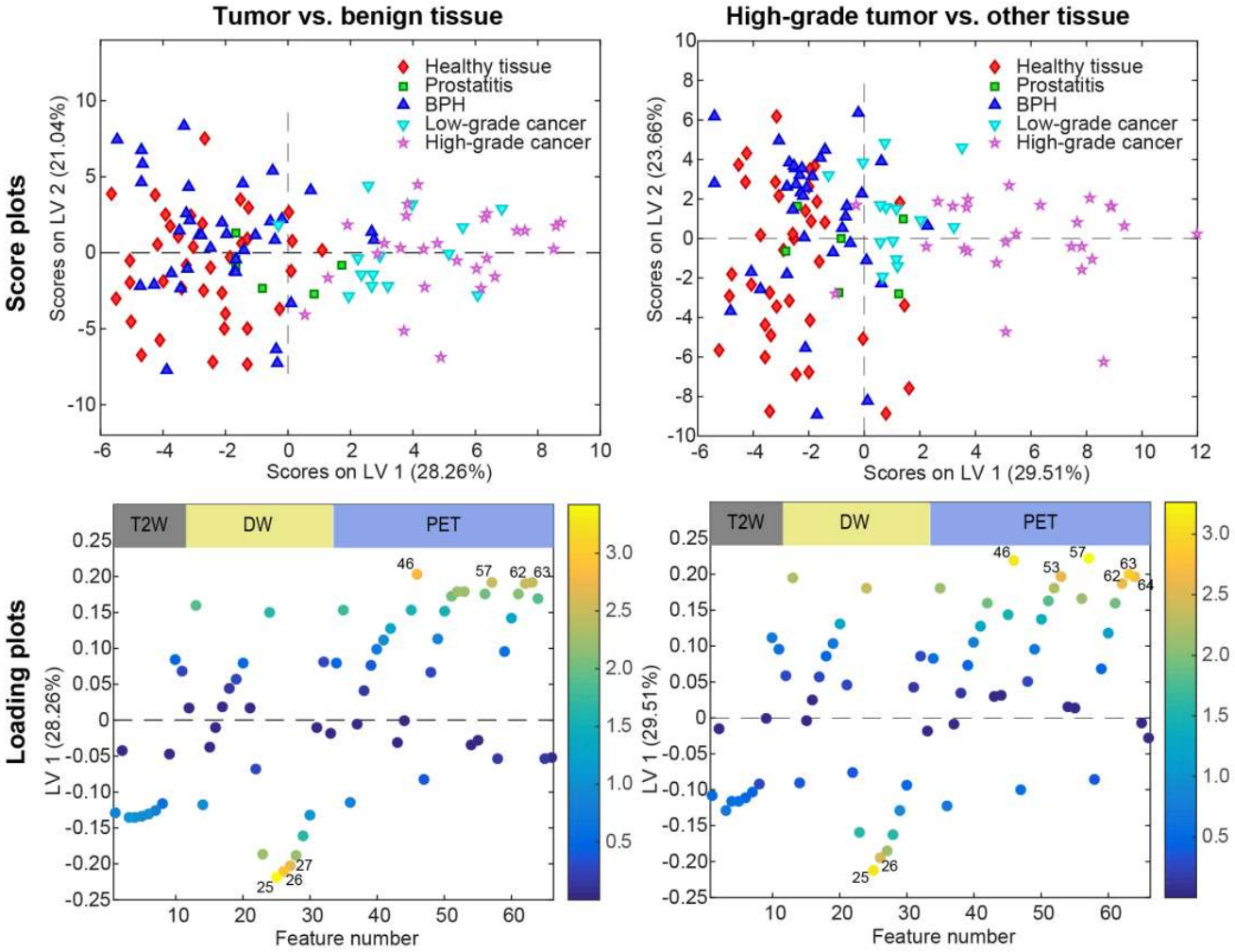
	Tumor vs benign tissue			High-grade tumor vs other tissue		
	MRI	PET	PET/MRI	MRI	PET	PET/MRI
Voxel-level						
AUC	81%	83%	87%*†	79%	81%	85%*†
Sensitivity	70%	76%	78%	69%	74%	76%
Specificity	77%	78%	81%	75%	77%	79%
VOI-level						
AUC	93%	93%	96%†	93%	91%	93%
Sensitivity	85%	83%	95%	88%	85%	88%
Specificity	94%	86%	89%	88%	86%	96%

* significantly different from MRI ($p < 0.05$)
† significantly different from PET ($p < 0.05$)

Table 4. Performance measures of the PET/T2W+DWI MRI feature set for discriminating between tumor and benign tissue and between high-grade tumor and all other tissue. The values of the complete PET/MRI feature set are given in parentheses for comparison

	Tumor vs benign tissue	High-grade tumor vs other tissue
Voxel-level		
AUC	87% (87%)	84% (85%)
Sensitivity	76% (78%)	75% (76%)
Specificity	82% (81%)	78% (79%)
VOI-level		
AUC	97% (96%)	94% (93%)
Sensitivity	93% (95%)	92% (88%)
Specificity	92% (89%)	90% (96%)

SUPPLEMENTAL MATERIAL



Supplemental Figure 1. Overview of the scores and loadings plots of the PET/T2W+DW MRI models for the VOI-level classification tasks. The scores plots (upper row) show the excellent discrimination between tumor and benign tissue (left) and between high-grade tumor and other tissue (right) on LV1. The loadings (bottom row) are colored by their VIP score, showing that low ADC values (labels 25, 26, 27), high SUV values (labels 53, 62, 63, 64), and high standard deviations (labels 46, 57) were the most important features characterizing the malignant classes. Features 1-11: mean, standard deviation, minimum, 10% percentile, 25% percentile, median, 75% percentile, 90% percentile, maximum, skewness, and kurtosis, respectively, for nT2W; 12-22: for b800; 23-33: for ADC; 34-44: for SUV^{5-10}; 45-55: for SUV^{18-23}; 56-66: for SUV^{33-38}