A fully automatic computer aided diagnosis system for peripheral zone prostate cancer detection using multi-parametric magnetic resonance imaging

Valentina Giannini\textsuperscript{a,}\textsuperscript{*}, Simone Mazzetti\textsuperscript{a}, Anna Vignati\textsuperscript{a}, Filippo Russo\textsuperscript{a}, Enrico Bollito\textsuperscript{b}, Francesco Porpiglia\textsuperscript{c}, Michele Stasi\textsuperscript{d}, Daniele Regge\textsuperscript{a}

\textsuperscript{a} Department of Radiology at Candido Cancer Institute—FPO, IRCCS, Strada Provinciale 142 km 3.95, 10060 Candilo, Italy
\textsuperscript{b} Division of Pathology at the University of Turin, San Luigi Hospital, Regione Gonzole, 10, 10043 Orbassano, Italy
\textsuperscript{c} Division of Urology and Department of Oncology at the University of Turin, San Luigi Hospital, Regione Gonzole, 10, 10043 Orbassano, Italy
\textsuperscript{d} Department of Medical Physics at Candido Cancer Institute—FPO, IRCCS, Strada Provinciale 142 km 3.95, 10060 Candilo, Italy

\textbf{A R T I C L E   I N F O}

Article history:
Received 5 February 2015
Received in revised form 9 June 2015
Accepted 2 September 2015

Keywords:
Prostate cancer
Multiparametric MRI
Computer aided detection
Image analysis
SVM classifier

\textbf{A B S T R A C T}

Multiparametric (mp)-Magnetic Resonance Imaging (MRI) is emerging as a powerful test to diagnose and stage prostate cancer (PCa). However, its interpretation is a time consuming and complex feat requiring dedicated radiologists. Computer-aided diagnosis (CAD) tools could allow better integration of data deriving from the different MRI sequences in order to obtain accurate, reproducible, non-operator dependent information useful to identify and stage PCa. In this paper, we present a fully automatic CAD system conceived as a 2-stage process. First, a malignancy probability map for all voxels within the prostate is created. Then, a candidate segmentation step is performed to highlight suspected areas, thus evaluating both the sensitivity and the number of false positive (FP) regions detected by the system. Training and testing of the CAD scheme is performed using whole-mount histological sections as the reference standard. On a cohort of 56 patients (i.e. 65 lesions) the area under the ROC curve obtained during the voxel-wise step was 0.91, while, in the second step, a per-patient sensitivity of 97% was reached, with a median number of FP equal to 3 in the whole prostate. The system here proposed could be potentially used as first or second reader to manage patients suspected to have PCa, thus reducing both the radiologist’s reporting time and the inter-reader variability. As an innovative setup, it could also be used to help the radiologist in setting the MRI-guided biopsy target.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Multiparametric magnetic resonance imaging (mp-MRI) combines morphological and functional information and is being increasingly used to detect cancer. One of the most promising applications of mp-MRI is to detect prostate cancer (PCa) \cite{1, 2, 3}. Current indications of mp-MRI include patients with rising levels of prostatic-specific-antigen (PSA) after one or more negative transrectal ultrasound guided (TRUS) biopsies \cite{4, 5}. One good reason why mp-MRI has not yet progressed to becoming a front-line imaging modality to detect PCa is because it is a labour-intensive examination and has a steep learning curve. Indeed, interpretation requires experienced radiologists capable of analysing data extracted from the different MR sequences \cite{5, 6}.

Computer aided detection (CAD) systems have the potential to support the radiologist by indicating suspicious regions and reducing oversight and perception errors \cite{7}. In addition, some CAD applications have been shown to be time efficient \cite{8}; however, this may be accomplished only if minimal or no human interaction is required in post-processing.

The implementation of a fully automatic CAD system is not a trivial problem. Chan et al. first implemented in 2003 a CAD system for the diagnosis of peripheral zone (PZ) PCa, using a support vector machine (SVM) classifier \cite{9}. Vos et al. \cite{6} developed in 2012 a CAD scheme using multiple sequential steps, including initial blob detection on apparent diffusion coefficient (ADC) maps followed by a local feature analysis by a supervised classifier. Similarly, Niaf et al. \cite{10} used a SVM classifier combined with a \(t\)-test feature selection method, and achieved an area under the ROC curve (AUROC) equal to 0.82. Finally, Litjens et al. \cite{5} included different stages in their CAD system, investigating a novel voxel classification step in combination with a candidate classification stage.
None of the described studies has adopted a fully automatic registration step to align both the Dynamic Contrast Enhanced (DCE-MR) and the Diffusion Weighted (DW) to the T2-weighted (T2w) image. In this study, we present a fully automatic CAD system for PCA detection, which seeks to overcome the limitation of previous related works. Training and testing of the CAD scheme is performed using whole-mount histological sections as the reference standard.

2. Materials and methods

2.1. Patients

This was a single institution study. We enrolled all individuals that complied with the following inclusion criteria: (a) biopsy-proven prostate adenocarcinoma, (b) mp-MRI examination between April 2010 and November 2012, including axial T2w, DW, and DCE–MR sequences, (c) radical prostatectomy (RP) within 3 months of MRI, and (d) a clinically significant PZ lesion (tumour volume ≥0.5 ml) [11] at the whole-mount histopathologic analysis. The local Ethics Committee approved the study and participants in the study signed informed consent forms. This study was in accordance with the Helsinki Declaration.

2.2. MR image acquisition

Images were acquired with a 1.5 T scanner (Signa Excite HD, GE Healthcare, Milwaukee, Wisconsin, USA) using a four-channel phased-array coil combined with an endorectal coil (Medrad, Indianola, Pa). Axial T2w images were obtained using the following protocol: slice thickness, 3 mm; field of view (FOV), 16 x 16 cm; NEX, 2; acquisition matrix, 384 x 288; reconstruction matrix, 512 x 512; TR/TE, 3020/85 ms. DW imaging was obtained using axial Echo-Planar Imaging sequences as follows: slice thickness, 3 mm; FOV, 16 x 16 cm; acquisition matrix, 128 x 128; reconstruction matrix, 256 x 256; NEX, 6; TR/TE, 7000/101 ms; b-values, 0 and 1000 s/mm². Finally, a 13 s time resolution DCE study was performed, with an axial 3D Spoiled Gradient echo (SPGR) sequence using the following parameters: TR/TE/FA, 3.6 ms/1.3 ms/20°; FOV, 20 x 20 cm; slice thickness, 3 mm; acquisition matrix, 224 x 192; reconstruction matrix, 512 x 512. Scanning started simultaneously with the intravenous injection of 0.1 mmol/kg gadobutrol (Gadovist, Bayer Schering, Berlin, Germany) through a peripheral line at 0.7 ml/s, using a power injector (Medrad Spectris, Maastricht, The Netherlands), followed by infusion of 20 cm³ normal saline at same rate. Twenty-six contrast-enhanced frames were obtained. The average time to complete the whole MR exam, including two additional T2w scans in the sagittal and coronal plane, was 40 min. Overall imaging parameters satisfied the minimal scanning requirements [4].

2.3. Reference standard and MR correlation

The prostate specimen was step-sectioned at 3 mm intervals perpendicular to the long axis (apical-basal) of the gland [12]. This confidently reproduces the inclination of axial T2w images, which were acquired perpendicular to the rear gland surface. The bases and the axes were cut parasagittally. Five μm sections were then obtained and coloured with hematoxylin eosin. The pathologist (E.B., with 24 years of experience in pathology, 20 attending uropathology) outlined each clinically significant peripheral tumour on microscopic slices and assigned a pathological Gleason Score (pGS). The radiologist (F.R., with an experience of more than 500 prostate mp-MRI studies interpreted per year for 6 years) in consensus with the pathologist, established the reference standard for PCa on T2w images drawing freehand regions of interest (ROIs) on cancer foci, following the outlines drawn by the pathologist on digital images of the pathologic slices. When pathological microslices and axial T2w images were not perfectly overlapped, usually due to modified prostate shape soaked by formaldehyde, the radiologist and the pathologist established the locations of tumours with respect to identifiable anatomic landmarks (e.g., adenoma nodule, urethra, ejaculatory ducts, and benign prostatic hyperplasia). If a lesion extended into more than one histopathologic slice, a ROI was drawn on each corresponding MR slice. For each patient a ROI, with extension similar to the tumoural region, was also drawn on normal gland located in the contralateral PZ.

2.4. CAD pipeline

The pipeline of the CAD system is shown in Fig. 1; it is conceived as a 2-stage process. First, a parametric colour-coded map of the prostate gland is created; colours are assigned to the map based on the probability of each voxel to be cancerous (Fig. 1A). Then, a candidate segmentation step is performed to highlight suspected areas (Fig. 1B). Different fully automatic steps, thoroughly described in the following subsections, compose each of these stages. All methods are implemented using C++ and the ITK libraries [13].

2.4.1. Image registration

Image registration has been described in detail elsewhere [14]. Registration is an important step as it allows to correctly align different types of images so that features, derived from all the MR sequences and referring to the same pixel or group of pixels, may be compared and studied. Before applying the registration methods, both the DCE volumes and the DW images have been upsampled to the T2w image resolution, and the DCE volumes were automatically cropped to match the same FOV of the T2w image. The algorithm first aligns DW to the T2w images, by applying a non-rigid registration step. In particular, the deformation field is modelled as a linear decay field along the vertical direction (1), assuming that the pixel shifts caused by magnetic field inhomogeneities occur particularly in the phase encode direction and decrease linearly with distance from the coil (1).

\[
T(y) = \begin{cases} 
  d_1 - k \times y & 0 < y < \frac{d_1}{k} \\
  0 & y > \frac{d_1}{k} 
\end{cases}
\]

Moreover, the DCE images are aligned to the T2w sequence using a multi-resolution rigid registration algorithm. In this case, the registration is solved as an optimization problem with the goal of finding the optimal transformation \( T(x,y,z) \rightarrow (x',y',z') \) which maps any point in the moving DCE dynamic image sequence \( l(x,y,z,t) \) at time \( t \) into its corresponding point in the reference image \( f(x,y,z,0) \), i.e., the T2w. The mutual information, which is a measure of statistical dependency between two datasets, has been used as similarity metric and the regular step gradient descent algorithm has been used as optimizer [14].

2.4.2. Prostate segmentation

The segmentation of the prostate is of key importance to reduce the computational burden of the CAD system. In our method we first automatically identify on each slice a rectangular region of fixed size (i.e., width = 7 cm, height = 6 cm). The rectangle is automatically generated in such a way that its posterior border is in contact with the anterior profile of the coil, which is segmented using the Hough transform on the T2w image. It may confidently enclose both a normal (size = 4 x 2 x 3 cm) and an enlarged prostate [15]. Then, we extract the ADC map on this rectangular region and we apply, on the selected region, the multi-level Otsu threshold.
[16], i.e., 3 levels. This method is able to select different threshold values by maximizing the between-class variance in a gray level image, thus separating the prostate from the darker background and the brighter coil. Once the thresholding has been applied, a 3-values map is provided, representing the three different classes.

Finally, we select all voxels belonging to the second class, which represents the prostate, and we apply some morphological operations to fill holes, i.e. the darker voxels within the prostate, without enlarging the segmented prostate.

2.4.3. Features extraction

For all the voxels belonging to the segmented prostate, we extract anatomical and pharmacokinetic features, as detailed below.

2.4.3.1. Image intensity from ADC maps and T2w images. One of the major issues complicating MR image analysis is the absence of standardized signal intensity values, like the Hounsfield units in computed tomography. This usually means that an algorithm will give different results as scanners, sequences or even sequence parameters change [5]. To reduce this potential limitation, in this study we adopt two standardized measurements.

First, the pixel-wise ADC values are calculated with in-house C++ algorithms, developed using ITK open source libraries [13], by using a monoexponential model [17].

Second, we normalize the T2w signal intensity by using the signal intensity of the obturator muscle, automatically segmented by our system as previously described [18]. However, before applying image normalization, we correct the image inhomogeneities by using a variant of the popular non-parametric non-uniform intensity normalization (N3) algorithm introduced by Tustison et al. [19], with the following parameters: FWHM = 0.5, Wiener = 0.1, convergence threshold = 0.001.

2.4.3.2. Pharmacokinetic features. Before extracting any parameters, the curve $S(t)$ was normalized to the first DCE volumes, as following detailed: $(S(i) - S(0))/S(0)$, where $S(i)$ is the signal intensity at time $i$th and $S(0)$ is the signal intensity at time 0, i.e., the first DCE frame. Then some model-free features are derived from the normalized time-intensity curve $S(t)$: (a) maximum uptake (MU), (b) time to peak (TTP), defined as the frame index at which the maximum enhancement occurs, (c) wash-in rate (WI) defined as MU/TTP, (d) washout rate (WO) defined as

$$W0 = \frac{MU - S(t_{END})}{t_{END} - TTP}$$

where $t_{END}$ is the time at the last contrast-enhanced volume, (e) area under the curve $S(t)$ (AC), and (f) and area under the curve $S(t)$ computed within the first 60 s after the contrast agent injection (iAC).

These parameters are related to the curve type method approach commonly used by the radiologists [4], however they have some limitations, i.e. they are scanner and patient dependent, and they do not use all information present in the curve $S(t)$ [5].

To overcome these limitations, we decided to extract the maximum number of information from DCE curves, exploiting different approaches in quantifying tissue perfusion [10,20].

Therefore, we implemented both the well-known Tofts pharmacokinetic model to account for physiological parameters, and two empirical functions used to fit the [20] normalized $S(t)$ without making assumption on tumour physiology [21,22].

Using the bicompartimental Tofts model [23] we extracted: (a) $K_{trans}$, which is the transendothelial transport of contrast medium from vascular compartment to the tumor interstitium, (b) $k_{ep}$, which represents the transport parameter of contrast medium back into the vascular space, and (c) $v_p$, which is the plasma volume fraction. To model the arterial input function we have used a literature-based function [24], while the conversion of MR signal to contrast concentration curve is performed according to Medved et al. [25], by using the signal intensity of the obturator muscle as reference tissue ($T1 = 1026$ ms).

Tofts pharmacokinetic model provides a better understanding of the interaction between contrast agent and tissues, but it suffers from complexity, because of the required conversion of MR signal intensities into contrast agent concentration. Therefore, the following two empirical functions have been also used to fit the normalized $S(t)$: the Weibull function (3), and the Phenomenological Universalities (PUN) approach [22,26] (4):

$$y_{weib}(t)_{END} = At \exp (-t^B)$$

$$y_{PUN}(t) = \exp \left[ rt + \frac{1}{B} (a_0 - r) \exp (Br) - 1 \right]$$

The nonlinear curve fitting has been solved in the least-squares sense with the Levenberg–Marquardt algorithm, starting from an initial guess estimated with a grid parameters search. Upper and lower bounds are also chosen to reflect physiological behaviour normally found in healthy and tumoural tissues [27]. Advantages of non-parametric techniques are: (i) no need to convert MR signals to contrast agent concentration, and (ii) the ability to analyse the whole enhancement curve cycle from wash-in to wash-out phases.

2.4.4. Feature selection and voxel classification

Feature selection is a process of choosing a subset of original features to reduce dimensionality, remove irrelevant data, increase learning accuracy, and improve result comprehensibility [28]. To reduce the number of parameters, we use the feature selection
Table 1

Correlation matrix between all the features. Numbers in parentheses in the second column are the areas under the ROC curves of each parameter. In bold are all the selected parameters, and the correlation higher than 0.8.

<table>
<thead>
<tr>
<th>Semi-quantitative features</th>
<th>MU (0.677)</th>
<th>TTP</th>
<th>WI</th>
<th>VO</th>
<th>AC</th>
<th>IA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.00</td>
<td>−0.39</td>
<td>0.73</td>
<td>0.44</td>
<td><strong>0.92</strong></td>
<td><strong>0.84</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Weibull</th>
<th>PUN</th>
<th>Toft</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>A</strong></td>
<td><strong>B</strong></td>
</tr>
<tr>
<td></td>
<td>0.84</td>
<td>0.35</td>
</tr>
</tbody>
</table>

method based on the correlation matrix. For each feature, we evaluate its ability in discriminating between normal and tumour voxels, by measuring the AUROC curve (Table 1). Then, we compute the correlation matrix between the features, to detect which pairs are highly correlated. When a couple of features show a correlation ≥0.8, we discard the feature with the lower AUROC (Table 1). At the end of this step, for each voxel a 9-dimensional vector of features has been obtained, including ADC and T2w signal intensity, \(K_{trans}\) of Tofts model, \(A\) and \(B\) of Weibull function, \(β\) of PUN model, AC, TTP, and VO. Finally, a voxel classification is performed, which results in a likelihood between 0 and 1 per voxel, \(O\) indicating no suspicion of PCa and 1 indicating very high suspicion of PCa. Some pilot experiments have been conducted to evaluate the best classifier, and its best setting. The SVM classifier [29], which uses the radial basis kernel, yielded the best results and has been therefore chosen over both the Bayesian and the Linear Discriminant Analysis classifier. Parameter \(C\) and \(γ\) have been set to 2 and 0.002, respectively, after performing, on a preliminary dataset, a “grid search” using cross validation. To speed up training, thereby shortening the time for model selection, we have implemented a sampling strategy, that selects a small portion of the training data, taking into account the differences due to the heterogeneity of the tumour between different slices. Therefore, a reduced training dataset has been created by sampling the dataset with a step size of 50 voxels.

2.4.5. Candidate segmentation

On the likelihood map obtained during the previous step, we apply the second stage of the CAD pipeline to extract 3D candidates highly suspected to be cancers and reduce the number of false positives (FP) voxels.

First, voxels having probability higher than 60% to be malignant are extracted from the voxel-wise malignancy probability map, and connected regions with a size <100 mm\(^2\) are discarded (see Fig. 2). This size represents 60% of the volume of the smallest clinically significant PCa, i.e. 0.5 cm\(^3\) [11], therefore it has been chosen in order not to discard tumours that might have been only partially segmented.

Then, for each of the remaining voxels we compute the time-intensity curves from the DCE-MR dataset, previously smoothed with a median filter having kernel size of 11 × 11 × 0, and normalized to the first contrast-enhanced frame. Some heuristic criteria, inferred from a preliminary dataset, have been used to discard some voxels, thus reducing the number of FP: (a) MU between the 5th and 6th enhanced frame <1, (b) average signal intensity from the 12th and the last contrast-enhanced frame <1, and (c) increase or decrease of the signal intensity \(ε\) >60%, computed as

\[
ε = \frac{\bar{S}(t_{end}, t_{end-1}) - \bar{S}(t_4, t_5)}{\bar{S}(t_4, t_5)}
\]

where \(\bar{S}(t_{end}, t_{end-1})\) is the average of the signal intensity of the last and the second to last contrast-enhanced frame, and \(\bar{S}(t_4, t_5)\) is the average of the signal intensity of the 4th and the 5th contrast-enhanced frame.

2.4.6. Statistical analysis

To evaluate the performance of the system we used the leave-one-out method (LOO), which involves training on all but one case, estimating the likelihood of malignancy of the pixels belonging to the left-out patient, and repeating the procedure until each case has been tested individually. Both the voxel classification and the candidate selection steps have been evaluated.

The ROC analysis was used to evaluate the ability of the system in discriminating between normal and malignant tissue. Since the LOO strategy has been used, the median value of the AUROC, the median sensitivity and the median specificity have been provided.

The ROC analysis, however, misses information about the number of FP candidates. Therefore, the method may have a high discriminative performance though it presents many FP candidates. Therefore, for the candidate selection step, we provided the per-region and per-patient sensitivity, and the number of FP in both the peripheral zone and the whole prostate for each patient. A 3D connected region was considered a true positive (TP) if the segmentation lays within the ROI drawn by the radiologist, vice versa it was considered a FP. We evaluated the system both for the detection of all tumours and the detection of high-grade tumours (pGS ≥ 7).

3. Results

The study flowchart is presented in Fig. 3. From the initial cohort of 88 patients, 32 were excluded because of the following: the whole-mount section pathologic tumour maps were not available as patients received radiotherapy \(n = 22\); received hormonal therapy at the time of MR examination \(n = 2\); did not have cancer
on the excised prostate (n = 1); the whole-mount histopathologic analysis did not find any clinically significant lesion in the PZ of the prostate gland (n = 7).

The study population (56 patients) included 47 patients (84%) with one clinically significant PZ tumour focus, and 9 patients (16%) with two clinically significant PZ tumour foci, for a total of 65 clinically significant peripheral PCAs. The patient and lesion characteristics are summarized in Table 2. Among the features, $K_{\text{T2max}}$ and ADC showed the best individual classification performance, having an AUROC equal to 0.83 and 0.76, respectively, while the signal intensity of the T2w image was one of the worst parameters in discriminating between normal and tumoural voxels. However, the correlation matrix shows that T2w is not correlated with any other parameters, i.e. the highest correlation is equal to 0.41 with the ADC, demonstrating that it provides some additional independent information to PCA detection. Moreover, our results confirm the importance of using mp-MRI over single-parameter MR imaging, since it leads to an increased AUROC. Our CAD system reached a median AUROC of 0.91 (1st–3rd quartile: 0.83–0.98), a median sensitivity of 0.84 (1st–3rd quartile: 0.77–0.93), and a median specificity of 0.86 (1st–3rd quartile: 0.76–0.95), in the voxel classification stage. In the candidate segmentation step, it detected 62 out of 65 lesions, leading to a per-lesion sensitivity of 96%, with a median number of FP per patient equal to 1 (1st–3rd quartile: 0–2), and 3 (1st–3rd quartile: 1–4) in the PZ and in the whole prostate, respectively. The per-lesion sensitivity increased to 98%, if we consider only the high grade tumours (pGS ≥ 7). One of the 3 false negative (FN) lesions was a secondary tumour in a patient in which the primary lesion was correctly detected by the system, while the other 2 FNs belonged to two patients with only one clinically

### Table 2

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients included in study</td>
<td>56</td>
</tr>
<tr>
<td>Patients median age [y] (1st–3rd quartile)</td>
<td>64 (60–70)</td>
</tr>
<tr>
<td>Median PSA at diagnosis [ng/ml] (1st–3rd quartile)</td>
<td>5.9 (4.9–8.6)</td>
</tr>
<tr>
<td>Median no. of days between biopsy and MR examination [d] (1st–3rd quartile)</td>
<td>92 (61–112)</td>
</tr>
<tr>
<td>Median time between MR imaging and prostatectomy [d] (1st–3rd quartile)</td>
<td>26 (13–47)</td>
</tr>
<tr>
<td>Median prostate volume [ml] (1st–3rd quartile)</td>
<td>44.8 (37.3–59.5)</td>
</tr>
<tr>
<td>No. of lesions with tumour volume ≥0.5ml</td>
<td>65</td>
</tr>
<tr>
<td>Median volume (ml) (1st–3rd quartile)</td>
<td>1.6 (0.9–2.5)</td>
</tr>
</tbody>
</table>

Distribution of pathologic Gleason scores [no. of patients]

<table>
<thead>
<tr>
<th>Number</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 + 3</td>
<td>7 (11%)</td>
</tr>
<tr>
<td>3 + 4</td>
<td>33 (51%)</td>
</tr>
<tr>
<td>4 + 3</td>
<td>15 (23%)</td>
</tr>
<tr>
<td>4 + 4</td>
<td>7 (11%)</td>
</tr>
<tr>
<td>4 + 5</td>
<td>3 (4%)</td>
</tr>
</tbody>
</table>

![Flowchart of study population](image)

**Fig. 3.** Flowchart of study population.
significant tumour. Therefore, the system was able to detect at least one PCa in 54 over 56 patients, yielding a per-patient sensitivity of 97%. Fig. 4 shows two examples of how lesions are detected and visualized on CAD interface. Two out of 3 FN were PCas with pGS = 3 + 3 and size of 0.75 cm³ and 0.59 cm³; one FN was a PCa with pGS = 4 + 3 and size 1.28 cm³. The first one was not segmented because it did not show a suspicious behaviour in any of the MR images, while a low probability (i.e., around 20%) to be malignant was assigned to the other two FNs, probably because they showed a very high ADC (>1600 mm²/s), which is atypical for PCs [30]. Fig. 5 shows an example of a FN.

4. Discussion

In this work we present a two-stage CAD system providing both a voxel-wise malignancy probability map of the entire prostate gland and the segmentation of PCa candidates. In the voxel-wise
step, the AUROC for our CAD system was 0.91, while in the candidate segmentation step per-patient sensitivity was 97%—i.e., at least one PCA lesion was detected in 54 of the 56 patients—and the median number of FP per exam was 3 (1st–3rd quartile: 1–4). Other CAD systems relying on mp-MRI for PCa detection showed slightly lower AUROC values (0.89 vs 0.91) [56,10]. Besides, CAD scheme accuracy may be influenced by the method used for drawing ROIs. Previously, only Langer et al. [31] computed the AUROC on ROIs drawn by an experienced radiologist using histopathology as the ground truth, obtaining an AUROC of 0.71, which is lower than our value. Conversely, authors that obtained AUROC similar to ours either adopted the overall lesion detection method [10,32], or considered only voxels belonging to an area, with predefined size, around the radiologist annotations [5]. The latter approaches tend to improve the apparent performance since they do not take into consideration tumour heterogeneity and the overlaying image noise; therefore, the ability to determine a strict threshold between tumour and normal values is increased [31]. The method implemented by our group, which relies on the analysis of every voxel belonging to the ROIs, may potentially extend the role of the CAD system to the evaluation of tumour aggressiveness, which might not be uniform within the lesion.

The CAD system we implemented detected 62 of the 65 PZ PCAs (per-lesion sensitivity of 96%). Two out of the three FNs lesions were tumours with a pGS of 3+3; one of the 2 patients had a second dominant tumour that was detected by the CAD system. In both cases, ADC values were above the voxel inclusion threshold of ADC which was set at $1.6 \times 10^{-3}$ mm$^2$/s and one lesion had no contrast uptake. Recently, it has been argued that low GS lesions—i.e. ≤6—should not be defined as cancers due to their nihil propensity to metastatization [33] and should therefore not be treated with surgery. Indeed, missing such indolent lesions at imaging could be beneficial, as it may reduce overtreatment which, with the conventional diagnostic pipeline, is estimated to occur in approximately 20 to 60% of cases [34]. The third FN was a lesion with pGS of 4+3, which was discarded in the candidate selection step due to its high diffusion values (ADC > $1.6 \times 10^{-3}$ mm$^2$/s).

Our CAD system has important strengths. First, the reference standard was established using whole-mount pathological slices obtained from the prostate specimen, rather than relying on the radiologist’s annotations even if confirmed by biopsy [5,6]. The latter approach has an important limitation as no information is available on regions annotated as normal by the radiologist that could actually harbour PCAs. Indeed, in a recent paper Le et al. [35] showed that radiologists can miss up to 40% of clinically significant PCAs. Conversely, using the whole-mount sections as histologic reference standard provides more accurate label information for training a CAD system [2], as it assesses the ground truth for both PCAs extent, grade, and disease infiltration [36]. Previously, other authors used the whole-mount histology as gold standard [31,37], however our method includes an automatic registration method.

Second, we introduced a registration step to align both the DCE and the DW images to the reference T2w image [14]. This step is of key importance to correct misalignment due to both patient movements and image distortions that occur especially during DW images acquisition [2,38]. Most of the recent works describing the development of a CAD system did not apply a registration algorithm before feature extraction [5,10]. Previously, only Vos et al. [32] and Viswanath et al. [39] embedded a fully automatic registration algorithm into a CAD system for PCa on mp-MRI. Vos et al. [32] aligned T2w and Proton Density (PD)-T1-w images, from which they extracted pharmacokinetics maps, while Viswanath et al. [39] performed an affine registration between the T2w image and the 5th contrast-enhanced frame of the DCE acquisition. The latter approach does not consider that registration errors could occur due to patients movements [5]. Conversely, our method is able to align each contrast-enhanced frame to the T2w image. Moreover, neither Vos et al. [32] nor Viswanath [39] assess a specific framework to correct for image distortion and patients movements during EPI-DW imaging acquisitions.

Indeed, although the mp-MRI is implicitly registered (all sequences are acquired in one go, without the patient leaving the scanner), image deformation on DW images and registration errors between the different DCE volumes could occur due to patient movement [5]. The latter may strongly influence parameter estimation and cause apparently extreme changes in enhancement [40]. The registration step, together with the features selection method, may account for the lower number of FPs obtained by our CAD system. This study has some limitations. First, it does not include the assessment of tumours of the transitional and central zone. The latter have different imaging features as they grow in or anteriorly to the central lobe which is inhomogeneous and usually includes patchy areas of highly vascularized tissue. Therefore, to train our SVM classifier, we will need a higher number of lesions than those available in our database ($n = 7$). Second, this database does not include patients without PCAs. However, this does not represent a limitation in the development and testing of the CAD system as normal tissue is outlined by using the pathological images as ground truth. We are now evaluating the CAD system prospectively in a screening cohort to assess its performance both in patients with and without PCAs. Third, our dataset includes patients acquired with the same scanner and using the endorectal coil; further analysis is ongoing to assess CAD performances on other MR scanners, at different MR field strength, with and without the use of endorectal coils. Finally, in a few cases the CAD system segmented a volume larger than the ground truth. While this might not affect per-patient sensitivity, it could be difficult to accurately localize the lesion within the gland. To overcome this limitation we are planning to refine the segmentation step by evaluating the local maxima in the probability map for each candidate selected by our system.

In conclusion, we developed a fully automated CAD system using mp-MRI that has demonstrated a high sensitivity in the detection of PZ PCa and a low number of FP findings per patient. Future developments will include the improvement of the classifier to assess centrally and anteriorly located lesions and a prospective evaluation of the system in a screening cohort.

Acknowledgments

This work was funded by Fondazione Piemontese per la Ricerca sul Cancro FPRC-onlus, grant Pro-Cure, 5 per Mille 2009 Ministero della Salute.

References


