available at www.sciencedirect.com
journal homepage: www.europeanurology.com

European Association of Urology

Platinum Priority – Prostate Cancer

Editorial by Bertrand Tombal on pp. 997–998 of this issue

Validation of the European Society of Urogenital Radiology Scoring System for Prostate Cancer Diagnosis on Multiparametric Magnetic Resonance Imaging in a Cohort of Repeat Biopsy Patients

Daniel Portalez[a,†], *Pierre Mozer*[b,†], *François Cornud*[c], *Raphaëlle Renard-Penna*[b], *Vincent Misrai*[a], *Matthieu Thoulouzan*[d], *Bernard Malavaud*[d,*]

[a] *Departments of Radiology, Clinique Pasteur, Toulouse, France;* [b] *Department of Urology, Hôpital Pitié-Salpétrière, Paris, France;* [c] *Department of Radiology, Hôpital Cochin, Paris, France;* [d] *Department of Urology, Hôpital de Rangueil, Toulouse, France*

Article info

Article history:
Accepted June 20, 2012
Published online ahead of print on June 27, 2012

Keywords:
Prostatic neoplasms
Magnetic resonance imaging
MRI functional
Biopsy

Abstract

Background: Wide variations in acquisition protocols and the lack of robust diagnostic criteria make magnetic resonance imaging (MRI) detection of prostate cancer (PCa) one of the most challenging fields in radiology and urology.
Objective: To validate the recently proposed European Society of Urogenital Radiology (ESUR) scoring system for multiparametric MRI (mpMRI) of the prostate.
Design, setting, and participants: An institutional review board–approved multicentric prospective study; 129 consecutive patients (1514 cores) referred for mpMRI after at least one set of negative biopsies.
Intervention: Transfer of mpMRI-suspicious areas on three-dimensional (3D) transrectal ultrasound images by 3D elastic surface registration; random systematic and targeted cores followed by core-by-core analysis of pathology and mpMRI characteristics of the core locations. The ESUR scores were assigned after the procedure on annotated Digital Imaging and Communications in Medicine archives.
Outcome measurements and statistical analysis: Relationships between ESUR scores and biopsy results were assessed by the Mann-Whitney *U* test. The Yates correction and Pearson χ^2 tests evaluated the association between categorical variables. A teaching set was randomly drawn to construct the receiver operating characteristic curve of the ESUR score sum (ESUR-S). The threshold to recommend biopsy was obtained from the Youden *J* statistics and tested in the remaining validation set in terms of sensitivity, specificity, positive predictive value, negative predictive value, and accuracy.
Results and limitations: Higher T2-weighted, dynamic weighted imaging and dynamic contrast-enhanced ESUR scores were observed in areas yielding cancer-positive cores. The proportion of positive cores increased with the ESUR-S aggregated in five increments (ESUR-S 3–5: 2.9%; ESUR-S 6–8: 11.1%; ESUR-S 9–10: 38.2%; ESUR-S 11–12: 63.4%; and ESUR-S 13–15: 83.3%; $p < 0.0001$). A threshold of ESUR-S $\geq$9 exhibited the following characteristics: sensitivity: 73.5%; specificity: 81.5%; positive predictive value: 38.2%; negative predictive value: 95.2%; and accuracy: 80.4%. Although the study was not designed to compare repeat biopsy strategies, more targeted cores than random systematic cores were found to be positive for cancer (36.3% compared with 4.9%, $p < 0.00001$).
Conclusions: In the challenging situation of repeat biopsies, the ESUR scoring system was shown to provide clinically relevant stratification of the risk of showing PCa in a given location.

† These authors contributed equally to this work.
* Corresponding author. Department of Urology, Hôpital de Rangueil, 1, avenue Jean Poulhès, TSA 50032–31059, Toulouse Cedex 9, France. Tel. +33 561 323 229; Fax: +33 561 323 230.
E-mail address: malavaud.b@chu-toulouse.fr, bernard.malavaud@numericable.fr (B. Malavaud).

0302-2838/$ – see back matter
http://dx.doi.org/10.1016/j.eururo.2012.06.044

1. Introduction

Variations in magnetic resonance imaging (MRI) protocols and the lack of robust diagnostic criteria make MRI detection of prostate cancer (PCa) one of the most challenging fields of radiology and urology [1–3]. Multiparametric MRI (mpMRI) provides a noninvasive approach to characterizing the anatomy (T2-weighted [T2w] MRI), angiogenesis (dynamic contrast-enhanced [DCE] MRI), and cell density (diffusion-weighted MRI) of the prostate parenchyma [4,5].

Although accumulating evidence supported the relevance of mpMRI in PCa diagnosis [4,6,7], the widespread acceptance of this method was hampered by the absence of consensus on diagnostic criteria and of systems to quantify the risk of obtaining a positive core [2,8,9].

However, European urology and radiology experts convened to research areas of consensus on the conduct, interpretation, and reporting of multiparametric MRI of the prostate [6]. The European Society of Urogenital Radiology (ESUR) recently published a unified scoring system for mpMRI [10], named the Magnetic Resonance Prostate Imaging Reporting and Data System (MR PI-RADS), to emulate the Breast Imaging Reporting and Data System (BI-RADS) score successfully implemented for breast cancer to reduce interobserver variability, increase the diagnostic value of the technique, and improve communication between clinicians and radiologists [10,11].

Although derived from a critical appraisal of the literature, the proposed ESUR scoring system has not yet been validated in a real-life study. We took advantage of the development of MRI/transrectal ultrasonography (TRUS) fusion technology [12] to evaluate the predictive values of the ESUR scoring system in a cohort of 129 consecutive patients with a total of 1514 cores.

2. Materials and methods

2.1. Patients

This prospective study was institutional review board–approved (CPP-DC2011/37), and informed consent was obtained from all patients. From May to November 2011, 129 consecutive patients with a history of negative TRUS-guided biopsies of the prostate were referred for repeat biopsies.

2.2. 1.5-T multiparametric magnetic resonance imaging

Fast spin-echo T2w-MRI images were first acquired in three planes. Diffusion-weighted imaging (DWI) comprised multiple *b*-value acquisition and apparent diffusion coefficient (ADC) mapping in the same

Table 1 – Magnetic resonance imaging units and sequences at participating centers

MRI unit and workstation	Type	T2w MRI	DCE MRI	DWI MRI
Toulouse and Paris I MRI unit 1.5T Achieva™ (Philips Medical Systems, Best, The Netherlands)	Endorectal pelvic phased-array coil and SENSE cardiac coil (Philips Medical Systems)*	Fat spin echo in three planes, TR/TE 3300 ms/130 ms	Fat-saturated T1w FFE, 3D scan mode, TR/TE 11 ms/4.6 ms	TR/TE shortest as possible/72 ms
Extended WorkSpace™ (Philips Healthcare, Best, The Netherlands)	Slice parameters	Thickness 2.5 mm, gap 0.25 mm	Thickness 3 mm, gap 0 mm	Thickness 5 mm, SENSE sequence with P reduction 2.00
* Paris 1: no endorectal coil	FOV and matrix	FOV 170 × 153 mm, TSE factor 20, voxel 0.72 × 0.96 mm	FOV 130 × 116 mm; flip angle 25°, voxel 1.16 × 1.16 mm	FOV 370 × 286 mm, flip angle 90°, voxel 2.26 × 2.89 mm
	Functional imaging		14 consecutive postgadolinium dynamic sequences, reconstruction voxel 0.51 × 0.51 × 5.0	*b* values 0, 600, 800 s/mm^2 in the same plane as T2; voxel 1.93 × 1.95 × 5.0
	Duration	Duration 1 min 45 s	Gradient echo sequence 13 s per sequence	First acquisition 1 min 47 s, second for diffusion-weighted ADC 1 min 20 s
Paris 2 1.5T Avanto (Siemens Healthcare, Erlangen, Germany) ICAD Vividlook™ (ICAD, Nashua, NH, USA)	Eight-channel pelvic phased-array coil (Signa Excite, GE)	Fat spin echo in three planes, TR/TE 1300 ms/120 ms	Fat-saturated T1w FFE, 3D scan mode, TR/TE 5.11 ms/1.85 ms	TR/TE 3700 ms/104 ms
	Slice parameters	Thickness 3.5 mm, gap 0 mm	Thickness 3.5 mm, gap 0 mm	Thickness 3.5 mm, SENSE sequence with P reduction 2.00
	FOV and matrix	FOV 18 × 18 cm, matrix 192 × 256 mm, flip angle 140°, voxel 0.9 × 0.7 mm	FOV 20 × 20 cm, flip angle 10°, voxel 1.4 × 1.4 mm	FOV 18 cm, matrix 192 × 256 mm; flip angle 140°, voxel 1.27 × 1.27 mm
	Functional imaging		Consecutive postgadolinium dynamic sequences	*b* values 0, 100, 500,1000, 2000 s/mm^2 in the same plane as T2; voxel 1.93 × 1.95 × 3.5
	Acquisition time	5 min	Gradient echo sequence 8.5 s/sequence over 5 min	20 slices, 4 min

MRI = magnetic resonance imaging; T2w = T2-weighted; DCE = dynamic contrast-enhanced; DWI = diffusion-weighted imaging; SENSE = sensitivity encoding; TR/TE = time of repetition/time of echo; T1w = T1-weighted; FFE = fast field echo; 3D = three-dimensional; FOV = field of view; TSE = turbo spin echo; ADC = apparent diffusion coefficient.

planes as T2w sequences. DCE MRI was obtained by fat-saturated T1-weighted fast-field echo sequence and a temporal resolution of 8.5–15 s, following an intravenous bolus injection of gadoterate meglumine (Guerbet, France). Details of MRI sequences have been described elsewhere [2,13], MRI units and sequences are presented in Table 1.

2.3. Magnetic resonance imaging/three-dimensional real-time ultrasonography fusion–guided biopsy

Biopsies comprised sextant random systematic cores and targeted cores of suspicious areas on mpMRI. A typical suspicious lesion was well circumscribed and of low signal intensity on T2w imaging, showing restricted diffusion on ADC maps and early and intense enhancement with rapid washout on DCE imaging [2]. The decision to target a specific location and the technique for biopsy were left to the discretion of each investigator.

Cores were taken using a commercially available MRI/three-dimensional (3D) TRUS fusion–guided system (Koelis, LaTronche, France) that provided three separate functions: (1) precise 3D TRUS targeting, in the millimeter range, of a predefined location; (2) recording of the core spatial location; and (3) the ability to fuse MRI–Digital Imaging and Communications in Medicine archives and real-time TRUS images [12,14].

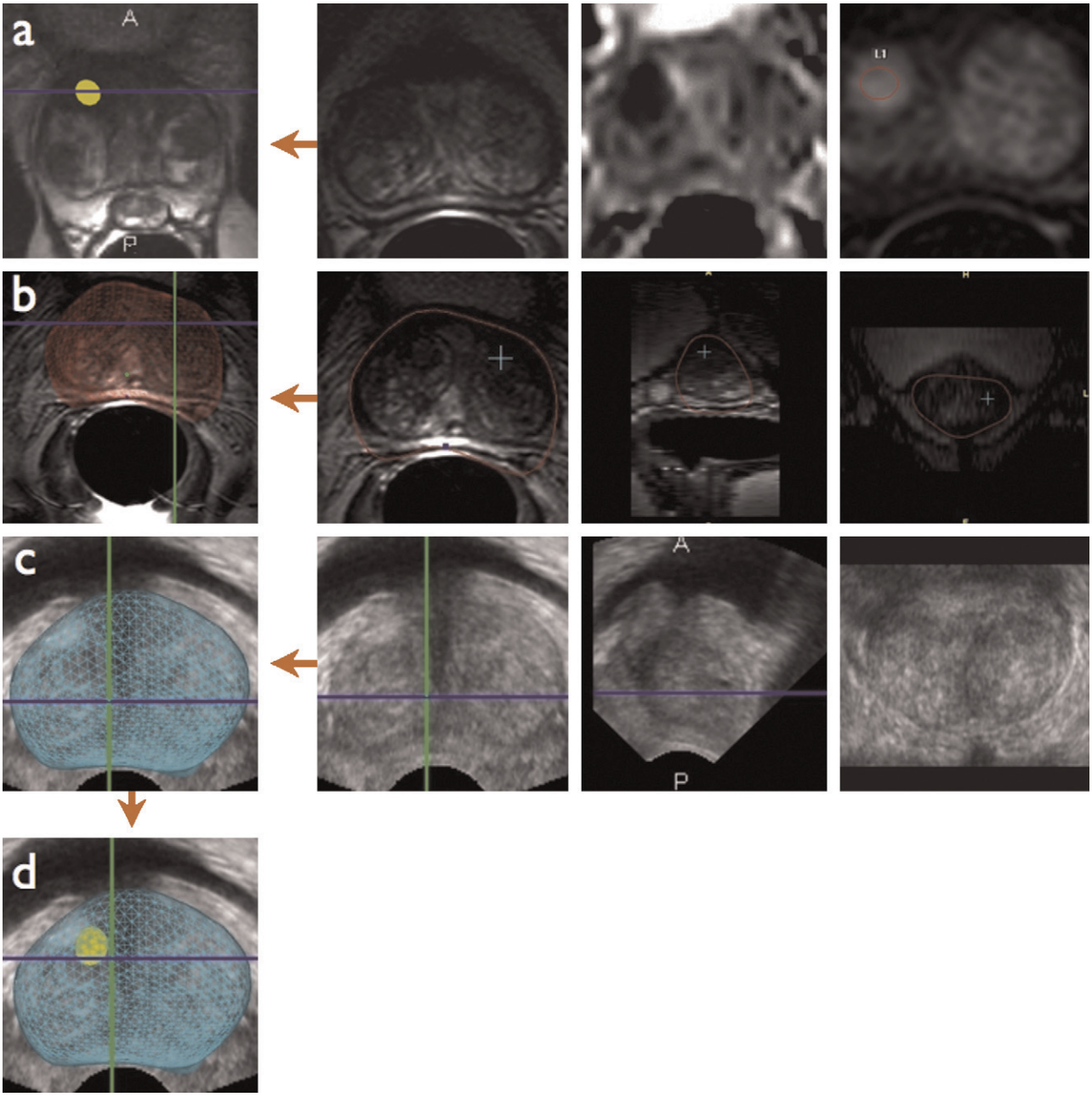

Fig. 1 – Target definition and elastic surface registration, axial representations of the prostate, are presented in the first column. (a) T2-weighted magnetic resonance imaging (MRI), diffusion-weighted imaging MRI, and dynamic contrast-enhanced MRI demonstrate a European Society of Urogenital Radiology score sum (ESUR-S) of 13 lesion in the anterior aspect of the right transition zone that is highlighted by a yellow tag in the first column. (b) Surface reconstruction from MRI; the prostate volume is contoured in the axial, sagittal, and coronal T2-weighted MRI acquisitions. (c) Surface reconstruction of three-dimensional transrectal ultrasonography (3D TRUS). (d) Elastic surface registration by means of an algorithm that controls the prostate constraints driven by changes in the patient's positioning, deformations in rectal wall, or displacement of the prostate. After registration, the MRI-suspicious areas highlighted by colored tags could alternatively be displayed within MRI- or 3D TRUS–registered volumes.

Briefly, MRI data were loaded in the workstation to obtain segmentation of the prostate volume (Fig. 1). Regions of interest were superimposed on mpMRI-suspicious areas. Real-time prostate 3D reconstruction was obtained from one axial and two oblique 60° acquisitions using a 3D end-fire endorectal probe (HD9, Philips Medical Systems, Best, The Netherlands). 3D TRUS and MRI reconstructions were finally registered by means of an algorithm that controls the prostate constraints [12].

Targeted cores were first taken, followed by random systematic cores. For both modalities, 3D TRUS acquisition was repeated with the needle in situ to register its precise location within the prostate volume. At the end of the procedure, all cores, numbered consecutively, were referenced to MRI and 3D TRUS archives (Fig. 2).

2.4. Postprocedure analysis

Each biopsy was annotated in terms of (1) location in 1 of the 16 sectors of the diagram provided by the European Consensus Meeting [6] and the ESUR [10], (2) random systematic compared with targeted cores, and (3) presence and length of cancer, Gleason grade. The length of the biopsy was not available. The mpMRI features of the area of the biopsy were characterized according to the ESUR system [10], which graded the level of suspicion for each MRI sequence from 1 to 5 (Table 2), the sum of the ESUR scores (ESUR-S) (range: 3–15), and the Likert scale [15].

2.5. Statistical analysis

The elementary unit for analysis was the core annotated with pathology and mpMRI results. We present mean and standard deviations or 95% confidence intervals (CIs) for continuous variables and percentages for categorical variables.

The relationship between mpMRI scores and biopsy results was assessed by the nonparametric Mann-Whitney *U* test. The Yates correction and Pearson χ^2 tests were used to evaluate the association between categorical variables. All reported p values are two-sided. Statistical significance was set at $p < 0.05$.

A teaching set of two-thirds of the study population was randomly drawn to construct the receiver operating characteristic curves of the mpMRI scoring systems. The thresholds to recommend a biopsy were

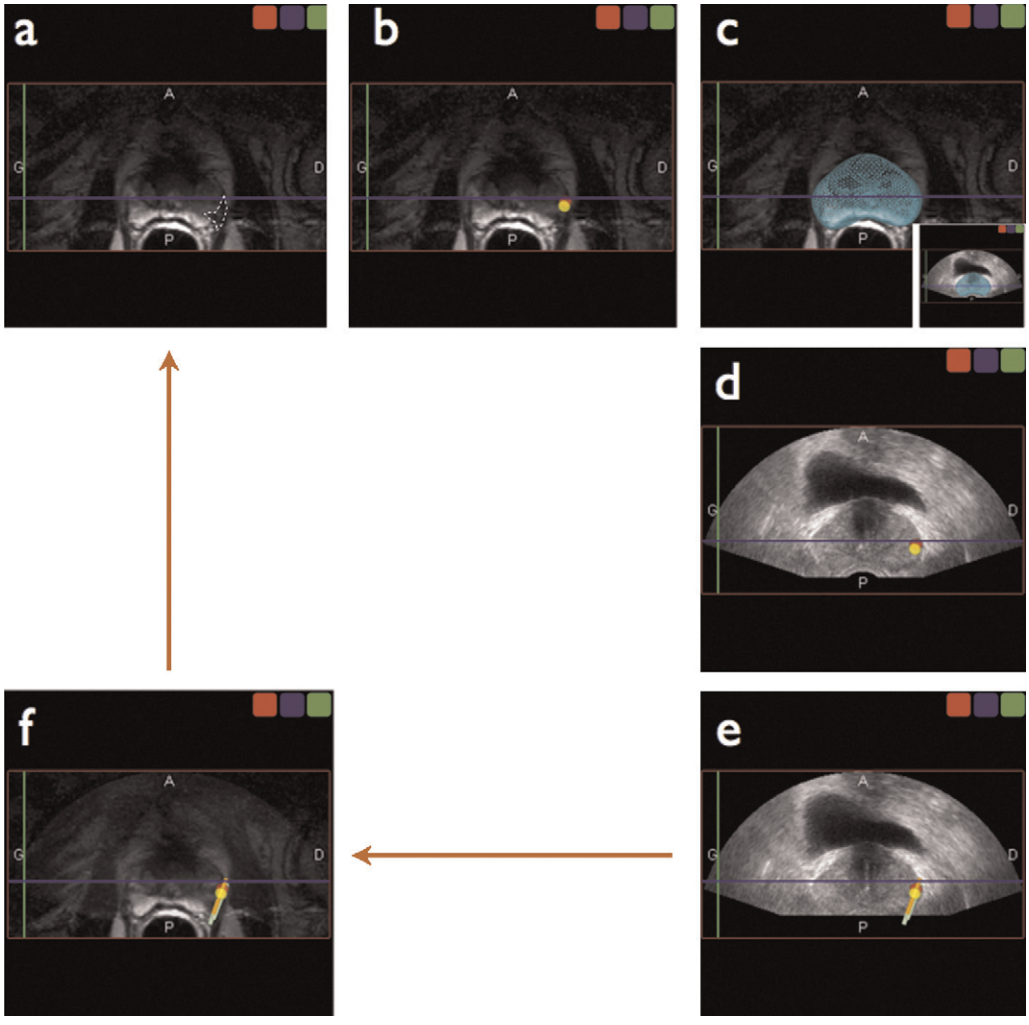

Fig. 2 – Postprocedure analysis of the European Society of Urogenital Radiology (ESUR) characteristics of multiparametric magnetic resonance imaging (mpMRI) Digital Imaging and Communications in Medicine archives; for clarity's sake, only the pictures pertaining to the target are presented. (a,b) Tagging of a suspicious focus in the right peripheral zone. (c,d) Transfer to the three-dimensional transrectal ultrasonography (3D TRUS) model by elastic surface registration ([c] T2-weighted magnetic resonance imaging; [inset] 3D TRUS). (e) Recording of the spatial location of the cores by repeating 3D TRUS with the needle left in place. (f) Reverse transfer to the mpMRI model. The actual ESUR characteristics of the site sampled by the cores was determined after the procedure by comparison with the initial mpMRI Digital Imaging and Communications in Medicine archives (a).

Table 2 – The Likert scale and the European Society of Urogenital Radiology Prostate Imaging Reporting and Data System scoring system for T2-weighted, diffusion-weighted, and dynamic contrast-enhanced imaging

Likert scale

Score 1 Clinically significant disease highly unlikely to be present
Score 2 Clinically significant cancer unlikely to be present
Score 3 The presence of clinically significant cancer is equivocal
Score 4 Clinically significant cancer likely to be present
Score 5 Clinically significant disease highly likely to be present

European Society of Urogenital Radiology scoring system

T2-weighted imaging for the peripheral zone

1 Uniform high signal intensity
2 Linear, wedge-shaped, or geographical areas of lower signal intensity, usually not well demarcated
3 Intermediate appearances not in categories 1/2 or 4/5
4 Discrete, homogeneous low-signal focus/mass confined to the prostate
5 Discrete, homogeneous low-signal-intensity focus with extracapsular extension/invasive behavior or mass effect on the capsule (bulging), or broad (>1.5-cm) contact with the surface

T2-weighted imaging for the transition zone

1 Heterogeneous transition zone adenoma with well-defined margins: "organized chaos"
2 Areas of more homogeneous low signal intensity; however, well marginated, originating from the transition zone/benign prostatic hyperplasia
3 Intermediate appearances not in categories 1/2 or 4/5
4 Areas of more homogeneous low signal intensity, ill defined: "erased charcoal sign"
5 Same as 4, but involving the anterior fibromuscular stroma or the anterior horn of the peripheral zone, usually lenticular or water-drop shaped

Diffusion-weighted imaging

1 No reduction in ADC compared with normal glandular tissue; no increase in signal intensity on any high–*b* value image (≥*b*800)
2 Diffuse, hyper signal intensity on ≥*b*800 image with low ADC; no focal features; however, linear, triangular, or geographical features are allowed
3 Intermediate appearances not in categories 1/2 or 4/5
4 Focal area(s) of reduced ADC but isointense signal intensity on high–*b* value images (≥*b*800)
5 Focal area/mass of hyper signal intensity on the high–*b* value images (≥*b*800) with reduced ADC

Dynamic contrast-enhanced imaging

1 Type 1 enhancement curve
2 Type 2 enhancement curve
3 Type 3 enhancement curve

+1 For focal enhancing lesion with curve type 2–3
+1 For asymmetric lesion or lesion at an unusual place with curve type 2–3

ADC = apparent diffusion coefficient.

estimated from the Youden *J* statistics (sensitivity + specificity − 1) and tested in the remaining one-third of the population in terms of diagnostic performance.

The characteristics of positive cores were compared according to the thresholds of the Likert scale and ESUR-S. For statistical analysis, the pathologic results were split into two groups adapted from Harnden et al. [16] (cancer length <3 mm and no Gleason pattern 4/5 vs cancer length ≥3 mm or Gleason pattern 4/5).

3. Results

3.1. Population

A total of 129 consecutive patients (mean age: 64.7 ± 6.9 yr [range: 47–79]) were enrolled for repeat biopsies (mean number of prior biopsy sessions: 1.3 ± 0.7 [range: 1–4]). Digital rectal examination (DRE) was unremarkable in 107 patients and suspicious in 22 patients (nodule smaller than half a lobe). Mean prostate-specific antigen (PSA) at biopsy was 9.6 ± 5.9 ng/ml (range: 2.7–40.0), and prostate volume (MRI estimate) was 51.1 ± 28.1 cm^3 (range: 12–192). Locations in the 16-sector map of random systematic cores (n = 1125) and targeted cores (n = 399) and proportions of positive cores are presented in Figure 3. PCa was demonstrated in 200 of 1524 cores (13.1%) (Table 3) or, alternatively, in 62 of 129 patients (48.1%).

Table 3 – Multiparametric magnetic resonance imaging and pathologic characteristics in 1125 random systematic and 399 targeted cores†

	Random systematic cores, n = 1125	Targeted cores, n = 399	p value
ESUR T2 score (1–5)	1.5 ± 0.6	3.3 ± 0.9	<0.00001*
ESUR DWI score (1–5)	1.4 ± 0.6	3.6 ± 1.2	<0.00001*
ESUR DCE score (1–5)	1.2 ± 0.6	2.5 ± 1.3	<0.00001*
ESUR sum of scores (3–15)	4.2 ± 1.4	9.4 ± 2.6	<0.00001*
Likert score (1–5)	1.6 ± 0.7	3.5 ± 1.0	<0.00001*
Peripheral zone, no.	1120	277	<0.00001°
Transition zone, no.	5	122	
Positive for cancer, no.	55	145	<0.00001°
Negative, no.	1070	254	*OR: 11.4, 95% CI, 8.1–16.1*
Length of cancer in positive cores, mm	2.0 ± 1.9	5.2 ± 3.1	<0.00001*
Gleason score			
≤3 + 3	21	57	
3 + 4	25	75	(NS) p = 0.3
≥4 + 3	9	13	

ESUR = European Society of Urogenital Radiology; T2 = T2-weighted imaging; DWI = diffusion-weighted imaging; DCE = dynamic contrast-enhanced; OR = odds ratio; CI = confidence interval; NS = not significant.
† Mean and standard deviation are provided for quantitative variables.
* Mann-Whitney *U* test.
° Yates corrected χ^2.

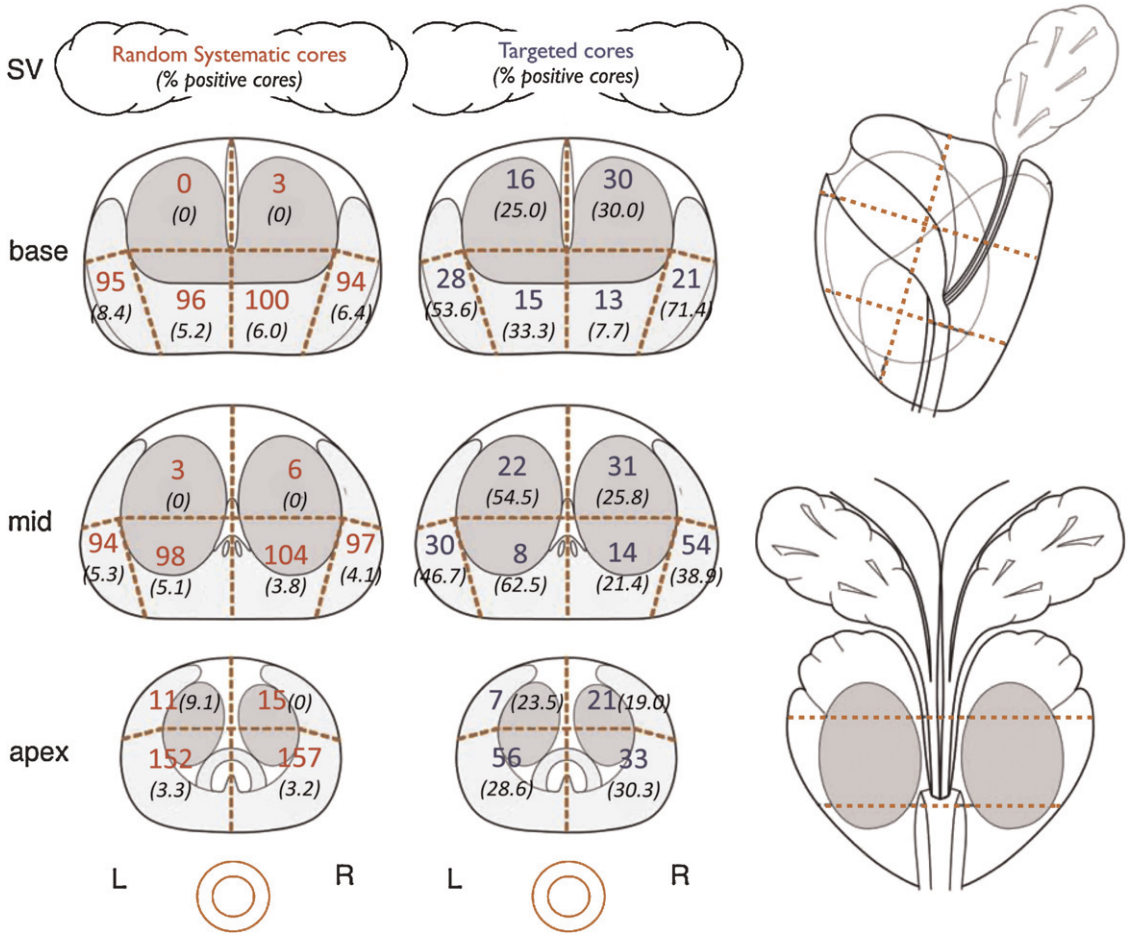

Fig. 3 – Absolute number of cores taken and proportion positive for cancer (in parentheses) per sector. Random systematic cores are in red (n = 1125), and targeted cores are in blue (n = 399). L = left; R = right.

3.2. Biopsy results

Targeted cores were more often found positive than were random cores (36.3% compared with 4.9%, $p < 0.00001$). The odds of demonstrating the presence of cancer were 11.4 higher in a targeted core than in a random core (Table 3) and varied within the prostate (odds ratios [95% CI]: 9.6 [5.6–16.5], 14.0 [7.9–24.7], and 12.1 [5.9–24.8] for the base, midprostate, and apex, respectively; all $p < 0.00001$) (Fig. 3). Targeted cores yielded longer segments of cancer (5.2 ± 3.1 mm and 2.0 ± 1.9 mm in targeted and random cores, respectively; $p < 0.00001$), although no differences were found in terms of Gleason patterns. Cancer was shown in 62 patients: in 21 patients by random and targeted cores, 35 patients by targeted cores only, and 6 patients by random cores only.

3.3. Predictive values of multiparametric magnetic resonance imaging characteristics

Higher ESUR scores were observed in positive cores (Table 4). The percentage of positive cores increased proportionally to T2w and DWI ESUR scores, whereas an indentation was observed for DCE MRI (Fig. 4). Consequently, the ESUR-S was also strongly related to the percentage of positive cores (p for trend < 0.0001).

A continuous increase in the proportion of positive cores was observed when ESUR-S was aggregated in five increments (mean [95% CI]: 2.9% [1.8–3.9], 11.1% [7.7–14.6], 38.2% [28.6–47.8], 63.4% [52.8–74.1], and 83.3% [73.1–93.6] for increasing increments of ESUR scores; p for trend < 0.0001) (Fig. 5). The yield also increased in

Table 4 – Multiparametric magnetic resonance imaging characteristics of 1524 cores, as a function of the presence or absence of prostate cancer

	Negative cores, n = 1324	Positive cores, n = 200	p value*
ESUR T2 score (1–5)	1.8 ± 0.9	3.3 ± 1.1	<0.00001
ESUR DWI score (1–5)	1.7 ± 0.9	3.6 ± 1.5	<0.00001
ESUR DCE score (1–5)	1.4 ± 0.8	2.7 ± 1.4	<0.00001
Sum of scores (3–15)	4.9 ± 2.2	9.7 ± 3.5	<0.00001
Likert score (1–5)	1.8 ± 0.9	3.6 ± 1.3	<0.00001

ESUR = European Society of Urogenital Radiology; T2 = T2-weighted imaging; DWI = diffusion-weighted imaging; DCE = dynamic contrast-enhanced.
* Mann-Whitney U test.

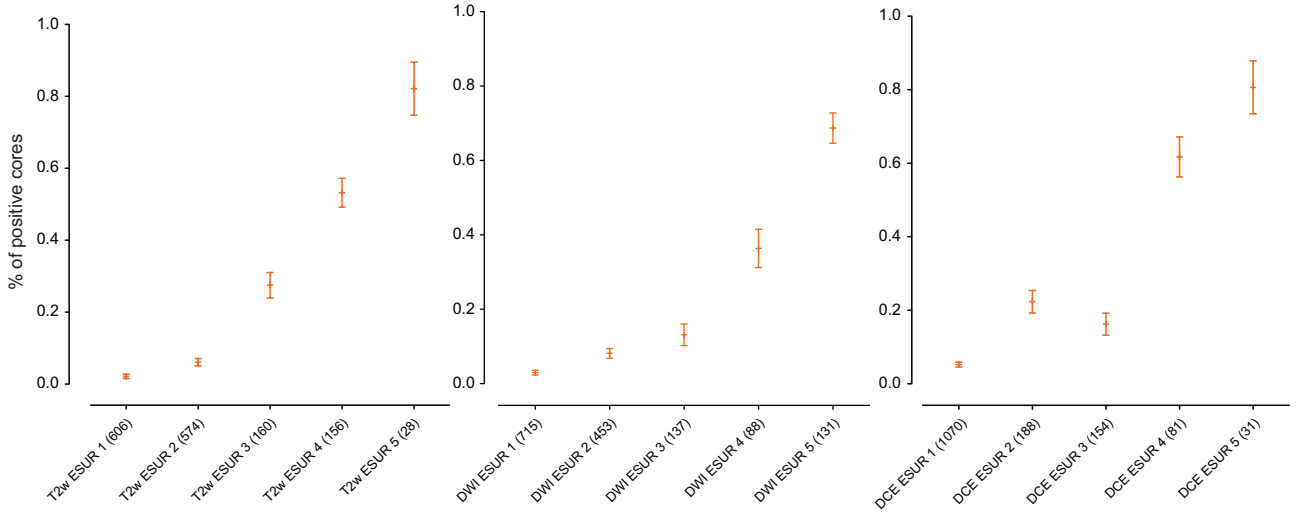

Fig. 4 – Proportion of cores positive for cancer according to T2-weighted (T2w) magnetic resonance imaging (MRI), diffusion-weighted imaging (DWI) MRI, and dynamic contrast-enhanced (DCE) MRI European Society of Urogenital Radiology (ESUR) score sum. The number of cores taken for each condition is presented in parentheses.

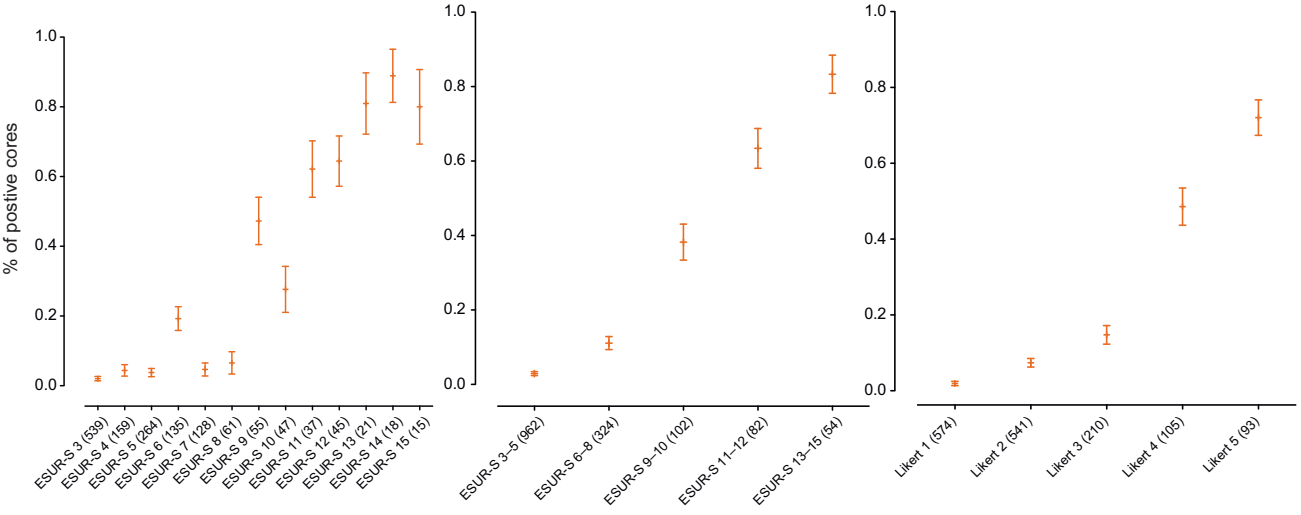

Fig. 5 – Proportion of cores positive for cancer according to the European Society of Urogenital Radiology score sum (ESUR-S), the ESUR-S in five increments, and the five classes of the Likert scale.

proportion to the five increments of the Likert scale (mean [95% CI]: 1.9% [0.7–3.0], 7.3% [5.2–9.6], 14.8% [9.9–19.6], 48.6% [38.9–58.3], and 72.0% [62.8–81.3]; p for trend < 0.0001)) (Fig. 5).

The balance between sensitivity and specificity for different thresholds was analyzed by the receiver operating characteristic curves of ESUR-S and the Likert scoring systems (Fig. 6). Both showed excellent areas under the curve (0.86 [95% CI, 0.83–0.89] and 0.84 [95% CI, 0.82–0.88], respectively), suggesting clinically relevant predictive characteristics.

The Youden J statistics indicated threshold values of 3 for the Likert scale and 9 for the ESUR-S in the random training set of 1005 cores. The predictive characteristics for these thresholds were then assessed in the validation set (Table 5). While both systems exhibited excellent negative predictive values in the 95% range, the ESUR-S showed a higher positive predictive value (58.0% compared with 38.2% for ESUR-S and the Likert systems in the validation set, respectively) and accuracy (89.1% compared with 80.4%).

We then used the threshold value of 9 to compare the cancer features of positive cores. Of the 1524 cores, 1286 cores (84.4%) were ESUR-S <9, of which only 64 cores (4.9%) were found positive for cancer (45 cores with adverse pathologic features [16], ie, cancer length ≥ 3 mm or Gleason pattern 4/5). Of the 238 cores collected in ESUR-S ≥ 9 locations, 136 cores (57.1%) were positive for cancer, of which 118 cores exhibited adverse pathologic features.

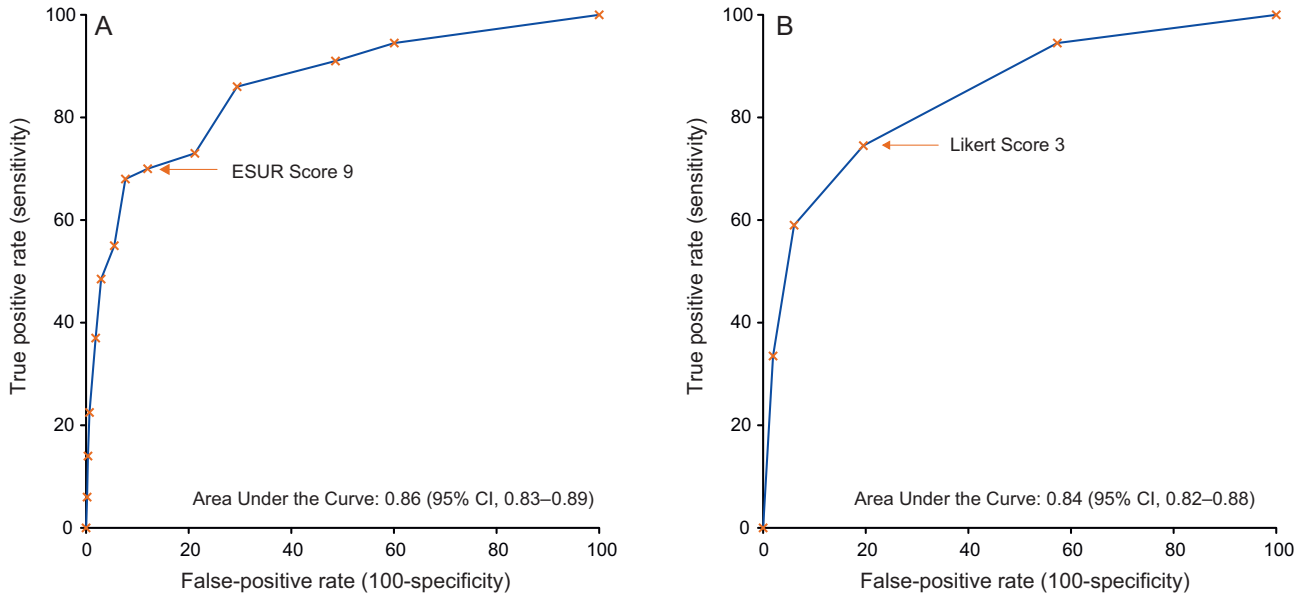

Fig. 6 – Receiver operating characteristic curves for (A) the European Society of Urogenital Radiology (ESUR) score sum and (B) the Likert scale. CI = confidence interval.

Table 5 – Receiver operating characteristics of the European Society of Urogenital Radiology sum of scores and the Likert scale in a training set and a validation set randomly drawn from the total cohort of 1524 cores*

	Training set, $n = 1018$		Validation set, $n = 506$	
Positive cores, no. (%)	132 (12.9)		68 (13.4)	
Random systematic cores, no. (%)	750 (73.7)		375 (74.1)	
Targeted cores, no. (%)	268 (26.3)		131 (25.9)	
	ESUR score	Likert Scale	ESUR score	Likert scale
AUC of the ROC curve	0.855 ± 0.019	0.845 ± 0.019	0.873 ± 0.022	0.848 ± 0.024
Youden-selected threshold	≥9	≥3	–	–
Sensitivity, % (95% CI)	67.4 (58.7–75.3)	75.0 (66.7–82.1)	69.1 (56.7–79.8)	73.5 (61.4–83.5)
Specificity, % (95% CI)	92.3 (90.3–94.0)	79.9 (77.1–82.5)	92.2 (89.2–94.5)	81.5 (77.5–85.0)
Positive predictive value, % (95% CI)	56.7 (48.6–64.6)	35.7 (30.1–41.7)	58.0 (46.5–68.9)	38.2 (29.8–47.1)
Negative predictive value, % (95% CI)	95.0 (93.3–96.3)	95.5 (93.7–96.9)	95.1 (92.4–96.8)	95.2 (92.4–97.0)
Overall accuracy, % (95% CI)	89.1 (87.0–90.9)	79.3 (76.6–81.7)	89.1 (86.0–91.6)	80.4 (76.6–83.7)

ESUR = European Society of Urogenital Radiology; AUC = area under the curve; ROC = receiver operating characteristics; CI = confidence interval.
* The predictive characteristics of multiparametric magnetic resonance imaging are presented for ESUR sum of scores ≥9 and Likert scale ≥3 thresholds, as determined by the Youden *J* statistics.

4. Discussion

In this paper, we showed the ability of the ESUR score to stratify mpMRI findings by cancer suspicion, as befits a clinically relevant scoring system. The objective of the present study was not to compare two modalities—random systematic and targeted—of sampling the prostate volume.

Although the ESUR system was based on literature evidence and consensus, it still lacked validation in a real-life setting. We took advantage of the archives produced by the 3D elastic registration system of Koelis image fusion technology to study the relationship between MRI scores and pathology findings in 129 patients referred for repeat biopsies.

As recently emphasized, cancer detection was traditionally assessed on a patient-by-patient basis; however, with the development of targeted biopsy techniques, cancer detection may be better assessed by lesion or biopsy core [17]. Indeed, one major limitation in studies pertaining to the diagnosis of PCa is that no proof of the absence of cancer can be obtained in patients found negative on TRUS-guided biopsies, while minute correlations (eg, whole-mount specimens) between pathology and mpMRI are possible in the others, amounting to a classic verification bias. To avoid this pitfall, the present study used the core, not the sextant or the general geometry of the cancer in specimens, as the elementary unit for analysis. The large number of cores targeted on suspicious foci ($n = 399$) and of random cores ($n = 1125$) encompassed all possible combinations of

mpMRI and pathology findings, thereby allowing an unbiased evaluation.

The proportional increase between percentage of positive cores and T2w and DWI scores fully validated the ESUR system for these two modalities (Fig. 4). For DCE, a drop was observed for score 3 that is assigned to several different conditions—that is, a typical type 3 enhancement curve or a less suspicious type 2 curve but with "focal enhancement" or "asymmetric lesion" or "at an unusual place" [10]. It was therefore a complex indicator that showed limitations in the present experiment. In addition, DCE is subject to variations in acquisition and spatial resolution [2]. We suggest that this indicator might be adapted or clarified by a compendium of reference cases, similar to that which ensured the rapid promotion of the BI-RADS system [18].

However, the recommended use of MRI in PCa is multiparametric [1,10], bringing into question the best way to sum up multidimensional results into a single score. The ESUR recommended a five-point scale to standardize the level of suspicion of the referent radiologist [10].

The Likert five-point scale showed continuous progression of the percentage of positive cores with increasing scores. Using score 3 and higher as threshold, the scale provided adequate sensitivity and specificity (73.5% and 81.5%, respectively) but a suboptimal positive predictive value (38.2%) (Table 5). We hypothesized that the structured analytic process of determining the ESUR scores for each of the three modes would perform better than the less structured Likert scale. The sum of the three ESUR scores provided a simple indicator (ESUR-S) that was split into classes to meet the ESUR recommendation of a five-point scale.

ESUR-S showed an almost linear progression in the percentage of positive biopsies, ranging from 2.9% for ESUR-S 3–5 to 83.3% for ESUR-S 13–15 (Fig. 5), suggesting that the ESUR-S would constitute a clinically relevant indicator to recommend a biopsy in a specific location.

The threshold of 9, as suggested by the Youden *J* statistics in the teaching set, exhibited in the validation set clinically relevant sensitivity (69.1%), specificity (92.2%), and positive and negative predictive values (58.0% and 95.1%, respectively) with adequate accuracy (89.1%). In the context of repeat biopsies, the observed excellent specificity would prove invaluable to drive attention to suspicious locations, while the ability to predict with high confidence the absence of cancer (negative predictive value) would control the number of unnecessary cores.

Still, because of suboptimal sensitivity, some cancer foci would be missed. In the present cohort, 1286 cores were taken in ESUR-S <9 areas, of which 64 cores (4.9%) were found positive for cancer. Conversely, restricting the biopsies to ESUR-S ≥ 9 would have detected 136 of 200 positive cores (68%). Although it is acknowledged that the significance of a core cannot be asserted without taking into account the results of the others, these figures strongly support the concept of targeting cores to locations found suspicious on mpMRI, as proposed by Emberton's group and others [17,19,20]. In addition, further studies will be needed to assess the respective values of clinical predictors (eg, DRE and PSA) and innovative predictors (eg, mpMRI and new biomarkers).

Precise registration of the core location was a crucial prerequisite. Indeed, to achieve homogeneous sampling of the prostate [21], one has to mentally integrate the real-time two-dimensional TRUS images into a 3D representation of the anatomic environment. Although this cognitive process was recently shown to achieve relevant accuracy in targeting mpMRI-suspicious foci [20], the process remained prone to imprecision, as evidenced by the differences in detection rates among operators of similar experience [22] and the discordances between consecutive sets or biopsies [23,24] or between biopsies and prostatectomy specimens [25–28].

Different expedients have been proposed. One is the use of a fixed grid as in brachytherapy for the transperineal approach; however, because of needle deflection during insertion [29], the method was reported to be liable to inaccuracy. Others reported a millimeter range of precision in transperineal targeting with a custom-made system of biplane ultrasonography–MRI and ad hoc software. Although ill adapted to the diagnostic biopsies that are routinely done transrectally, this method would provide a valuable asset in future MRI-guided focal therapy [30]. Other systems using tracking sensors on the needle guide [31] or the probe [17,32] managed to monitor their positions within the pelvis, but they could not account for the motions of the prostate induced by the probe or the needle.

Elastic surface registration was developed to overcome these limitations. This technology was shown to control the deformations induced by the biopsy [12] and was proved in a phantom model to target, with 100% accuracy and millimetric precision, randomly located 0.5-cm^3 hypoechoic lesions [14]. As shown in Table 3, image fusion targeted biopsies were 11-fold more likely to show PCa than random cores, providing strong simultaneous evidence for the relevance of mpMRI in PCa diagnosis and of fusion technology in guiding the cores. The relevance of targeted biopsies was also shown with a prototype electromagnetic tracking device [17] and with in-gantry MRI-compatible tracking devices [33]. However, to our knowledge, the present report is the first to use freely available commercial technology [17] in that setting.

Finally, the sole objective of this study was to report the relationship between mpMRI characteristics and pathology. For this purpose, data on 1524 cores was prospectively accumulated and analyzed on a core-by-core basis. Conversely, as the evaluation of strategies for biopsy was not an objective, no recommendations were given to the investigators in that respect. Therefore, the differences in the yields of targeted and random systematic biopsies should be considered suggestive but not conclusive, pending further studies.

5. Conclusions

The ESUR scoring system for multiparametric MRI of the prostate was shown to provide clinically relevant stratification of the risk of showing PCa in a given location. Only a few cancers were detected solely by random cores, as

opposed to the larger yield of cores targeted at mpMRI-suspicious locations.

As a whole, the validation of the ESUR scoring system and precise targeting of TRUS-guided biopsies now provide convincing leverage in the challenging field of PCa diagnosis.

Author contributions: Bernard Malavaud had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Portalez, Malavaud.
Acquisition of data: Portalez, Mozer, Cornud, Renard-Penna, Misrai, Thoulouzan, Malavaud.
Analysis and interpretation of data: Portalez, Malavaud.
Drafting of the manuscript: Portalez, Malavaud.
Critical revision of the manuscript for important intellectual content: Portalez, Mozer, Cornud, Malavaud.
Statistical analysis: Malavaud.
Obtaining funding: None.
Administrative, technical, or material support: None.
Supervision: Portalez, Malavaud.
Other (specify): None.

Financial disclosures: Bernard Malavaud certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/ affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: Pierre Mozer holds patents on the elastic registration software.

Funding/Support and role of the sponsor: None.

References

[1] Hoeks CM, Barentsz JO, Hambrock T, et al. Prostate cancer: multiparametric MR imaging for detection, localization, and staging. Radiology 2011;261:46–66.
[2] Portalez D, Rollin G, Leandri P, et al. Prospective comparison of T2w-MRI and dynamic-contrast-enhanced MRI, 3D-MR spectroscopic imaging or diffusion-weighted MRI in repeat TRUS-guided biopsies. Eur Radiol 2010;20:2781–90.
[3] Heidenreich A, Bellmunt J, Bolla M, et al. EAU guidelines on prostate cancer. Part 1: screening, diagnosis, and treatment of clinically localised disease. Eur Urol 2011;59:61–71.
[4] Heijmink SW, Futterer JJ, Strum SS, et al. State-of-the-art uroradiologic imaging in the diagnosis of prostate cancer. Acta Oncol 2011;50(Suppl 1):25–38.
[5] Giannarini G, Petralia G, Thoeny HC. Potential and limitations of diffusion-weighted magnetic resonance imaging in kidney, prostate, and bladder cancer including pelvic lymph node staging: a critical analysis of the literature. Eur Urol 2012;61:326–40.
[6] Dickinson L, Ahmed HU, Allen C, et al. Magnetic resonance imaging for the detection, localisation, and characterisation of prostate cancer: recommendations from a European consensus meeting. Eur Urol 2011;59:477–94.
[7] Kirkham APS, Emberton M, Allen C. How good is MRI at detecting and characterising cancer within the prostate? Eur Urol 2006;50:1163–75, discussion 1175.
[8] Weinreb JC, Blume JD, Coakley FV, et al. Prostate cancer: sextant localization at MR imaging and MR spectroscopic imaging before prostatectomy—results of ACRIN prospective multi-institutional clinicopathologic study. Radiology 2009;251:122–33.
[9] Villers A, Lemaitre L, Haffner J, Puech P. Current status of MRI for the diagnosis, staging and prognosis of prostate cancer: implications for focal therapy and active surveillance. Curr Opin Urol 2009;19:274–82.
[10] Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. Eur Radiol 2012;22:746–57.
[11] Molleran V, Mahoney MC. The BI-RADS breast magnetic resonance imaging lexicon. Magn Reson Imaging Clin N Am 2010;18:171–85, vii.
[12] Baumann M, Mozer P, Daanen V, Troccaz J. Prostate biopsy tracking with deformation estimation. Med Image Analysis 2012;16:562–76.
[13] Cornud F, Rouanne M, Beuvon F, et al. Endorectal 3D T2-weighted 1mm-slice thickness MRI for prostate cancer staging at 1.5Tesla: should we reconsider the indirects signs of extracapsular extension according to the D'Amico tumor risk criteria? Eur J Radiol 2012;81:e591–7.
[14] Ukimura O, Desai MM, Palmer S, et al. 3-dimensional elastic registration system of prostate biopsy location by real-time 3-dimensional transrectal ultrasound guidance with magnetic resonance/transrectal ultrasound image fusion. J Urol 2012;187:1080–6.
[15] Likert R. A technique for the measurement of attitudes. Arch Psychol 1932;140:1–55.
[16] Harnden P, Naylor B, Shelley MD, et al. The clinical management of patients with a small volume of prostatic cancer on biopsy: what are the risks of progression?. a systematic review and meta-analysis. Cancer 2008;112:971–81.
[17] Pinto PA, Chung PH, Rastinehad AR, et al. Magnetic resonance imaging/ultrasound fusion guided prostate biopsy improves cancer detection following transrectal ultrasound biopsy and correlates with multiparametric magnetic resonance imaging. J Urol 2011;186:1281–5.
[18] Agrawal G, Su MY, Nalcioglu O, Feig SA, Chen JH. Significance of breast lesion descriptors in the ACR BI-RADS MRI lexicon. Cancer 2009;115:1363–80.
[19] Rouse P, Shaw G, Ahmed HU, et al. Multi-parametric magnetic resonance imaging to rule-in and rule-out clinically important prostate cancer in men at risk: a cohort study. Urol Int 2011;87:49–53.
[20] Haffner J, Lemaitre L, Puech P, et al. Role of magnetic resonance imaging before initial biopsy: comparison of magnetic resonance imaging-targeted and systematic biopsy for significant prostate cancer detection. BJU Int 2011;108:E171–8.
[21] Sclaverano S, Chevreau G, Vadcard L, Mozer P, Troccaz J. BiopSym: a simulator for enhanced learning of ultrasound-guided prostate biopsy. Stud Health Technol Inform 2009;142:301–6.
[22] Lawrentschuk N, Toi A, Lockwood GA, et al. Operator is an independent predictor of detecting prostate cancer at transrectal ultrasound guided prostate biopsy. J Urol 2009;182:2659–63.
[23] Levine MA, Ittman M, Melamed J, Lepor H. Two consecutive sets of transrectal ultrasound guided sextant biopsies of the prostate for the detection of prostate cancer. J Urol 1998;159:471–5, discussion 475–6.
[24] Djavan B, Waldert M, Zlotta A, et al. Safety and morbidity of first and repeat transrectal ultrasound guided prostate needle biopsies: results of a prospective European prostate cancer detection study. J Urol 2001;166:856–60.
[25] Washington SL, Bonham M, Whitson JM, Cowan JE, Carroll PR. Transrectal ultrasonography-guided biopsy does not reliably identify dominant cancer location in men with low-risk prostate cancer. BJU Int 2012;110:50–5.
[26] Sinnott M, Falzarano SM, Hernandez AV, et al. Discrepancy in prostate cancer localization between biopsy and prostatectomy

specimens in patients with unilateral positive biopsy: implications for focal therapy. Prostate 2012;72:1179–86.
[27] Lattouf JB, Saad F. Gleason score on biopsy: is it reliable for predicting the final grade on pathology? BJU Int 2002;90:694–8, discussion 698–9.
[28] De Laet K, de la Taille A, Ploussard G, et al. Predicting tumour location in radical prostatectomy specimens: same-patient comparisons of 21-sample versus sextant biopsy. BJU Int 2009;104:616–20.
[29] McGill CS, Schwartz JA, Moore JZ, McLaughlin PW, Shih AJ. Precision grid and hand motion for accurate needle insertion in brachytherapy. Med Phys 2011;38:4749–59.
[30] Hadaschik BA, Kuru TH, Tulea C, et al. A novel stereotactic prostate biopsy system integrating pre-interventional magnetic resonance imaging and live ultrasound fusion. J Urol 2011;186:2214–20.
[31] Turkbey B, Xu S, Kruecker J, et al. Documenting the location of prostate biopsies with image fusion. BJU Int 2011;107:53–7.
[32] Natarajan S, Marks LS, Margolis DJ, et al. Clinical application of a 3D ultrasound-guided prostate biopsy system. Urol Oncol 2011;29:334–42.
[33] Hambrock T, Somford DM, Hoeks C, et al. Magnetic resonance imaging guided prostate biopsy in men with repeat negative biopsies and increased prostate specific antigen. J Urol 2010;183:520–7.