



**UNIVERSITÉ DE LILLE
FACULTÉ DE MÉDECINE HENRI WAREMBOURG**

Année 2020

**THÈSE POUR LE DIPLÔME D'ÉTAT
DE DOCTEUR EN MÉDECINE**

**Évaluation de la reproductibilité inter-observateurs du score
de suspicion de cancer de prostate
PI-RADS version 2.1**

Présentée et soutenue publiquement
Le 27 avril 2020 à 18 heures
Pôle Formation, salle 3
par

Julie BRUNEREAU
Née le 18 mai 1990 à Paris (12)

JURY :

Président :

Monsieur le Professeur Arnaud VILLERS

Asseseurs :

Monsieur le Professeur Xavier LEROY

Monsieur le Docteur Benoit RENARD

Monsieur Patrick DEVOS

Directeur de thèse :

Monsieur le Professeur Philippe PUECH

*Travail effectué au sein du service d'Imagerie Génito-Urinaire
de Hôpital Claude Huriez, CHU de LILLE*

Avertissement

La faculté n'entend donner aucune approbation aux opinions émises dans les thèses. Celles-ci sont propres à leurs auteurs.

Liste des abréviations :

ADC : Apparent Diffusion Coefficient

AFU : Association Française d'Urologie

CaP : Cancer de la Prostate

CSS : Cascading Style Sheets

DICOM : Digital Imaging and Communications in Medicine

EG : Echo de Gradient

EPE : Extension Extra Prostatique

EPU : Enseignement Post Universitaire

ESUR : European Society of Urogenital Radiology

FOV : Field of View

HAS : Haute Autorité de Santé

HBP : Hypertrophie Bénigne de Prostate

HIFU : High Intensity Focal Ultrasound

HTML : HyperText Markup Language

IRM : Imagerie par Résonance Magnétique

IRM-mp : IRM prostate multiparamétrique

ISP: IntelliSpace Portal

JPEG : Joint Picture Exchange Group

PI-RADS : Prostate Imaging – Reporting and Data System

PSA : Prostate Specific Antigen

BRUNEREAU Julie

PST : Prostate Score Trainer

RTUP : Résection TransUrétéral de Prostate

SFMA : Stroma Fibro-Musculaire Antérieur

T (unité de mesure): Tesla

TE : Temps d'Echo

TNM : Tumor – Node – Metastasis

TR (dépendant du contexte) : Toucher rectal / Temps de répétition

TSE/FSE : Fast Spin Echo/Turbo Spin Echo

SIGU : Société d'Imagerie Génito-Urinaire

ZC : Zone Centrale

ZP : Zone Périphérique

ZT : Zone de Transition

SOMMAIRE

I- Résumé :	7
II- Introduction:	9
1) Généralités sur le Cancer de la Prostate :	9
A- Epidémiologie :	9
B- Anatomie:	10
C- Anatomopathologie :	11
D- Dépistage du CaP :	13
2) IRM prostatique et CaP:	15
A- Indications :	15
B- Réalisation de l'examen :	17
C- Protocole :	17
- a- Les séquences pondérées T2:	18
- b- Les séquences de diffusion:	18
- c- Les séquences dynamiques en pondération T1 avec injection de Gadolinium (DCE) :	19
D- Score de lecture PI-RADS :	20
- a- Séquence T2 :	21
- b- Séquence Diffusion (DWI) :	24
- c- Séquence dynamique pondérée T1 après injection de Gadolinium (DCE) :	26
- d- Cas particuliers de la ZC et du SFMA :	27
E) Cartographie prostatique :	29
3) Plateforme internet « Prostate Score Trainer » (PST) :	30
A- Généralités sur le PST :	30
B- Présentation des résultats aux lecteurs:	32
4) Objectifs de l'étude :	35
III- Matériels et Méthodes:	36
1) Population des lecteurs:	36
2) Données du PST :	37
3) Analyse statistiques:	40
IV- Résultats:	43
1) Population des lecteurs	43
2) Concordance lecteurs/référence pour le score PI-RADS :	43
3) Concordance lecteurs/référence pour la localisation lésionnelle:	45
4) Analyse en groupe de lecteurs:	45
5) Analyse en groupe des lésions cible :	46
6) Analyse en fonction de la topographie ZP/ZT:	47
7) Analyse de l'apprentissage au fur et à mesure des cas:	48
V- Discussion:	53
VI- Conclusion:	58
VII- Bibliographie :	62

I- Résumé :

Objectifs :

Evaluer la reproductibilité inter observateurs du score de suspicion de cancer de prostate PI-RADS v2.1 au sein d'une population de 52 radiologues grâce à la lecture de 50 IRM de prostate (IRM-mp) sur la plateforme internet « Prostate Score Trainer » (PST).

Matériel et Méthodes :

52 lecteurs ont été inclus dans cette étude entre le 1^{er} novembre 2019 et le 24 janvier 2020. Il s'agissait de radiologues séniors et juniors provenant de multiples centres d'imagerie et ayant une expérience variable dans la lecture des IRM-mp. Sur la plateforme PST, les lecteurs ont analysé 100 lésions prostatiques provenant de 50 examens consécutifs réalisés sur un aimant 1,5 T entre le 5 janvier et le 12 février 2016. Pour chacune des 100 lésions, le grade PI-RADS (5200 réponses) et la localisation lésionnelle (5200 réponses) ont été déterminés. Les 10400 réponses ont été comparées à celles d'un expert considéré comme référence. Pour étudier la reproductibilité inter observateurs, nous avons calculé le coefficient k de Cohen (weighted kappa = wk), le taux de concordance avec la référence (P_{bonREP} pour le PI-RADS et P_{topoOK} pour la localisation) et l'écart normalisé à la référence ($RMSE$). Le score PI-RADS a été étudié en 5 et 3 classes. Les lésions ont été étudiées par zone. Les lecteurs et les lésions ont été séparés en sous-catégories (clusters) en fonction de leurs niveaux de reproductibilité.

Résultats:

Pour le score PI-RADS en 5 classes / 3 classes, les valeurs des $wk - P_{bonREP}$ étaient respectivement : 0,58/0,63 - 0,51/0,70. Pour la localisation lésionnelle, la valeur P_{topoOK} était 0,77. Ces paramètres étaient significativement différents dans la zone de transition (ZT) et dans la zone périphérique (ZP). Grâce aux paramètres de reproductibilité et aux valeurs de $RMSE$, les lecteurs et les lésions ont été respectivement séparés en 2 et 4 clusters. Les lésions de la ZT étaient significativement sur-représentées dans les clusters à faible et très faible concordances. Il n'y avait pas de lien significatif entre l'expérience des lecteurs et leur répartition dans les 2 clusters.

Conclusion:

Notre étude montre une reproductibilité inter observateurs du PI-RADS v2.1 comparable à celle du PI-RADS v2. L'expérience des lecteurs ne semble pas avoir une influence majeure. Il semble persister des difficultés dans l'attribution du grade PI-RADS et la localisation lésionnelle pour les lésions de la ZT. La plateforme PST semble être un outil efficace pour évaluer les performances du score PI-RADS en proposant à un effectif important de lecteurs, une analyse d'un nombre élevé de lésions.

II- Introduction:

1) Généralités sur le Cancer de la Prostate :

A- Epidémiologie :

En France, le cancer de prostate (CaP) est de nos jours, un problème majeur de santé publique. En effet, c'est le cancer le plus fréquent chez l'homme avec une incidence annuelle de 48400 nouveaux cas (1). Il est responsable de 8200 décès par an et représente la 3^{ème} cause de décès par cancer chez l'homme après le cancer du poumon et le cancer colo-rectal.

Deux facteurs expliquent cette évolution :

- s'agissant d'un cancer du sujet âgé, l'allongement de l'espérance de vie provoque mécaniquement une augmentation de l'incidence annuelle et de la prévalence de la maladie (2),
- depuis la fin des années 80, le dosage sérique d'un antigène spécifique de la prostate (PSA) permet un diagnostic précoce du cancer, bien avant la survenue des premiers symptômes.

Les facteurs de risques connus du CaP sont :

- un âge avancé (âge moyen au moment du diagnostic = 70 ans),
- des antécédents de cancer du sein ou de la prostate au premier degré,
- certaines origines ethniques : population de souche afro-américaine et population issue des caraïbes (2).

Diagnostiqué à un stade précoce, le taux de survie d'un CaP localisé est de 80- 90% à 10 ans du traitement, quel que soit le type de traitement réalisé (3).

B- Anatomie:

La prostate est une glande située dans la partie médiane du pelvis masculin. Elle se présente sous la forme d'une pyramide inversée avec une base supérieure et un apex inférieur et tronqué. Elle se situe sous la vessie, entre le col vésical et le diaphragme uro-génital. Ses principaux rapports sont : en avant la symphyse pubienne et en arrière le rectum. La prostate englobe l'urètre prostatique et contient le carrefour urogénital. Sa fonction est exocrine. Les glandes prostatiques sécrètent un liquide plutôt acide qui représente environ un tiers de volume de l'éjaculat. **(Figure 1)**

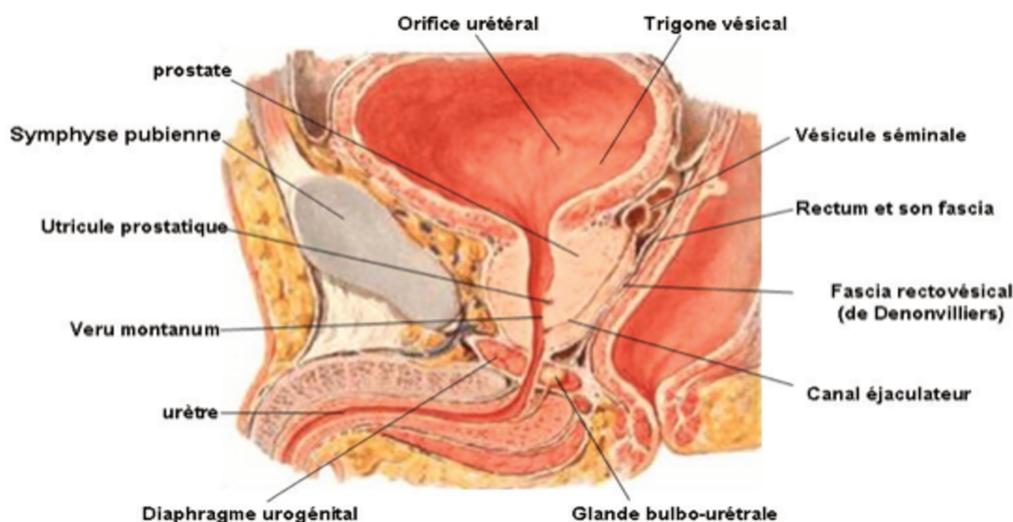


Figure 1: Coupe sagittale anatomique de la prostate (selon F.Netter, Atlas Anatomie humaine)

Les travaux publiés en 1981 par McNeal et al (4) distinguent 4 zones prostatiques délimitées par les structures canalaies intra-prostatiques que sont l'urètre prostatique et les deux canaux éjaculateurs. Ils distinguent la zone de transition (ZT), la zone centrale (ZC), la zone périphérique (ZP) et le stroma fibromusculaire antérieur (SFMA) ou zone fibreuse. **(Figure 2)**

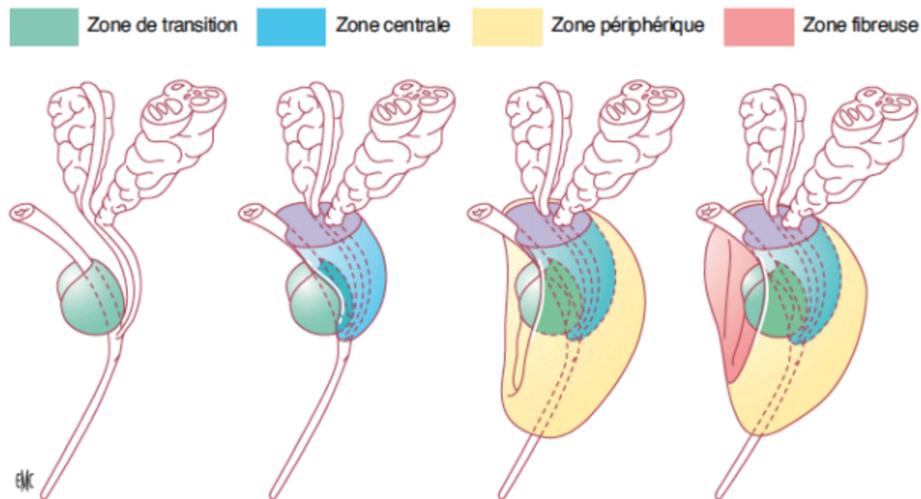


Figure 2: Anatomie zonale de la prostate selon McNeal (EM Consult).

La ZT : zone constituée de deux lobes développés de part et d'autre de l'urètre prostatique au-dessus du veru montanum et en avant des canaux éjaculateurs. Dans certains cas, cette zone comporte aussi un lobe médian situé sous la vessie faisant une empreinte postérieure sur le col vésical. Quasi inexistante chez le sujet jeune, elle s'hypertrophie avec l'âge. Environ 25% des CaP se développent dans cette zone, plutôt dans sa partie antérieure (1,5).

La ZC : zone de forme conique qui se situe en arrière de l'urètre prostatique et englobe les canaux éjaculateurs. La fréquence des CaP est de l'ordre de 4 à 7% dans cette zone (1,6).

La ZP : zone majoritaire chez le sujet jeune, elle englobe les structures précédentes et se situe en position postéro-latérale. Les cancers de cette zone représentent 70-75% des CaP (4,5).

Le SFMA : zone constituée de fibres de collagènes, de fibres musculaires lisses et parfois de quelques glandes prostatiques. Elle est située à la partie antérieure de la prostate et s'étend du col vésical jusqu'à l'apex prostatique.

C- Anatomopathologie :

L'adénocarcinome prostatique est la forme habituelle du CaP. Les formes plus rares sont des tumeurs neuroendocrines, des tumeurs urothéliales, des lymphomes, des sarcomes... (2)

Le CaP est diagnostiqué grâce à une analyse anatomo-pathologique de prélèvements tissulaires tels que des biopsies prostatiques transrectales écho-guidées (**Figure 3**) et plus rarement des biopsies prostatiques transpérinéales (7), des copeaux de résection

transurétrale de prostate (RTUP) ou une pièce de prostatectomie radicale.

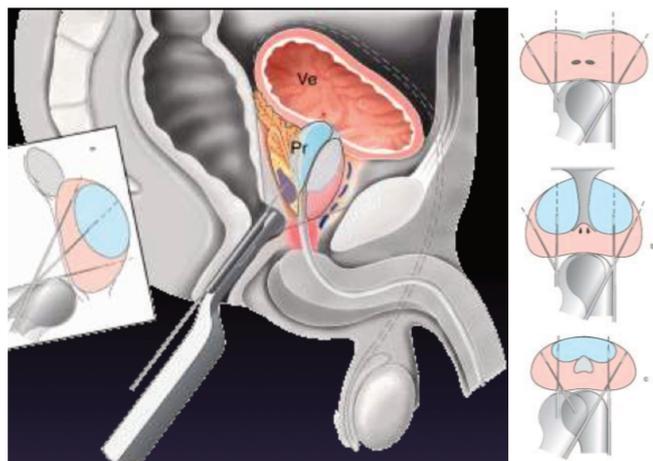


Figure 3: Schéma des 12 biopsies prostatiques transrectales systématiques écho-guidées. a) 4 prélèvements à la base, b) 4 prélèvements au milieu en regard du veru montanum (abouchement des canaux éjaculateurs), c) 4 prélèvements à l'apex (Images Pr PUECH, DINU 2019).

Le CaP est souvent multifocal et polyclonal. L'agressivité du CaP est évaluée grâce à un score anatomo-pathologique rapporté en 1966 et appelé « score de Gleason ». Différenciant 5 grades cellulaires cotés de 1 à 5 (du grade 1 = population de cellules bien différenciées au grade 5 = population de cellules dédifférenciées), ce score est basé sur l'individualisation de 2 contingents (contingent le plus représenté et contingent le plus péjoratif) et sur la sommation des grades cellulaires des 2 contingents. **(Figure 4)** Les valeurs du score de Gleason s'échelonnent en théorie de 2 à 10 mais sont exceptionnellement inférieures à 6 (8). Elles représentent un facteur pronostic majeur de la maladie.

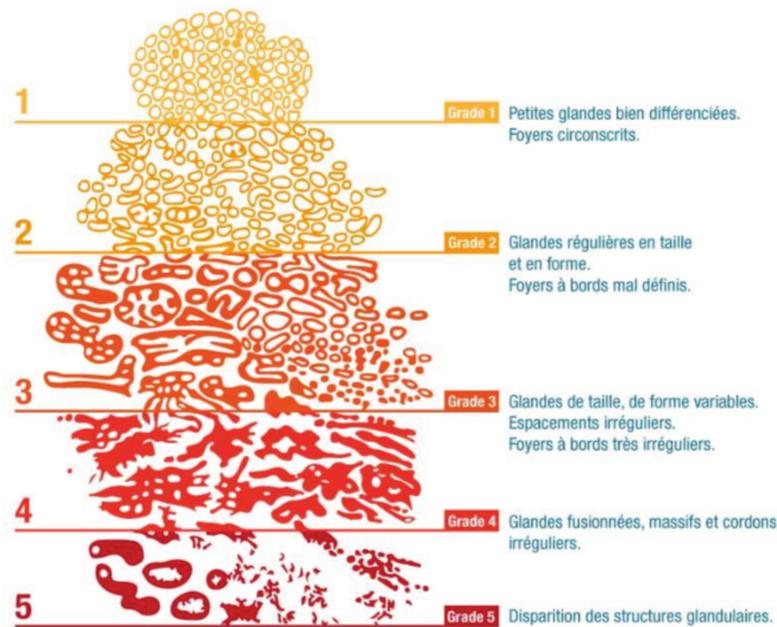


Figure 4: Représentation schématique des différents grades de Gleason.

Ce système a été revu lors des conférences de consensus de l'ISUP (International Society of Urological Pathology) en 2005 puis en 2014 (9). Désormais, on ne doit plus définir l'agressivité du CaP avec le seul score de Gleason mais avec un grade ISUP. **(Figure 5)**

The overall Gleason score for this case is based on the core with the highest Gleason score. Gleason scores can be grouped and range from Prognostic Grade Group I (most favorable) to Prognostic Grade Group V (least favorable).	
Gleason score ≤ 6 :	Prognostic Grade Group I
Gleason score $3 + 4 = 7$:	Prognostic Grade Group II
Gleason score $4 + 3 = 7$:	Prognostic Grade Group III
Gleason score 8:	Prognostic Grade Group IV
Gleason score 9-10:	Prognostic Grade Group V

Figure 5 : Grades ISUP 2014 (9)

D- Dépistage du CaP :

Le dépistage systématisé du CaP pouvant engendrer un risque de sur-traitement (10), la Haute Autorité de Santé (HAS) conseille actuellement de pratiquer un dépistage individuel après information éclairée du patient sur les risques et bénéfices de ce dépistage (11,12).

Selon la recommandation de l'Association Française d'Urologie (AFU), le dépistage du CaP s'adresse à la population des hommes de 50 et 75 ans. Il peut être pratiqué dès 45 ans en cas de facteur de risque familial ou ethnique accru (13).

Le dépistage du CaP repose sur l'association du toucher rectal (TR) et du dosage sérique du taux de PSA.

Lorsque le TR met en évidence une induration ou un nodule suspect à la palpation de la prostate, il peut justifier à lui seul la réalisation de biopsies prostatiques. Il détecte alors des cancers palpables plutôt localisés dans la zone périphérique.

Le PSA est une glycoprotéine qui favorise la liquéfaction du sperme et qui n'est sécrétée que par le parenchyme prostatique. Son taux sérique est exprimé en ng/mL. Même si une cellule cancéreuse bien différenciée possède un potentiel de sécrétion de PSA vingt fois supérieur à celui d'une cellule normale, l'élévation sérique du PSA n'est pas spécifique du CaP. Elle peut être la conséquence d'autres pathologies telles qu'une hypertrophie bénigne de prostate (HBP), une infection ou une inflammation prostatique (prostatite aigue et chronique). Le dosage du PSA est donc exposé au risque de faux positifs de CaP. Mais, il est important de noter qu'il expose aussi au risque de faux négatifs de CaP car une cellule cancéreuse dédifférenciée peut ne plus sécréter de PSA (14).

L'interprétation d'un taux de PSA chez un patient donné doit faire appel à différents paramètres d'analyse que sont la valeur initiale du PSA, la cinétique du PSA et le volume de la prostate.

En cas de premier dosage, la valeur initiale ne doit pas dépasser un seuil de 4 ng/mL (15). Entre 4 et 10ng/mL, la proportion de CaP augmente significativement et atteint 25 à 35%. Au-delà de 10 ng/mL, elle atteint 50 à 80% (16).

Pour s'aider dans l'interprétation du taux de PSA lorsque sa valeur initiale est située entre 4 et 10 ng/mL, deux informations supplémentaires sont intéressantes à connaître : la densité du PSA et le rapport PSA libre/ PSA total.

- La densité de PSA est le ratio entre le taux de PSA et le volume prostatique exprimé en mL. Si cette densité dépasse le seuil de 15% ou 20% (17) (valeur controversée), le risque de CaP augmente.

- Le rapport PSA libre/ PSA total sous-entend de doser en plus du PSA total, la fraction du PSA non liée aux protéines dite « PSA libre ». Celle-ci diminue en présence d'un CaP. Si le rapport est supérieur à 15% et surtout à 20%, la probabilité d'une HBP

est grande. Si à l'inverse, le rapport est inférieur à 15% et surtout à 10%, la probabilité d'un CaP augmente (14).

En cas de multiples dosages successifs, l'évolution du taux de PSA doit rester progressive et concordante avec l'évolution du volume prostatique. Si la cinétique perd soudainement ce caractère linéaire pour tendre vers une croissance exponentielle, il faut suspecter l'apparition d'un CaP.

2) IRM prostatique et CaP:

A- Indications :

L'utilisation de l'IRM prostatique a nettement augmenté depuis le début des années 2000, démontrant son utilité à tous les stades de la maladie (18). Mais avant 2018, elle était surtout recommandée dans le cadre du bilan d'extension loco-régional d'un CaP prouvé par des prélèvements prostatiques (biopsies, RTUP). Son indication répondait aux critères de la classification de D'Amico publiée en 1998. (19). Cette classification basée sur les données du TR, sur le taux de PSA et sur le score de Gleason, évalue le risque évolutif d'un CaP en déterminant un taux de survie à 5 ans sans récurrence biologique après traitement (prostatectomie radicale, radiothérapie externe, curiethérapie). Elle distingue 3 groupes de patients à :

- Bas risque évolutif (90% de survie sans récurrence biologique) : TR = T1c/T2a et PSA < 10 ng/mL et Score de Gleason ≤ 6
- Risque évolutif intermédiaire (70% de survie sans récurrence biologique) : TR = T2b ou PSA entre 10 et 20 ng/mL ou Score de Gleason = 7
- Haut risque évolutif (50% de survie sans récurrence biologique) : TR = T2c ou PSA >20 ng/mL ou Score de Gleason ≥ 8 (14)

Récemment, des études (20,21) ont démontré l'intérêt de la réalisation d'une IRM avant biopsies pour augmenter la détection des CaP cliniquement significatifs. Dans la littérature, de multiples définitions ont été rapportées sur cette notion de CaP cliniquement significatif (22). Nous retiendrons les conclusions de l'étude *Precision* avec comme seul critère déterminant, un score de Gleason supérieur ou égal à 7 sur au moins une biopsie (21).

Le protocole d'examen IRM a ainsi été optimisé dans le but de détecter les cancers

avec une très bonne sensibilité. D'une IRM simple, nous sommes passés à l'ère de l'IRM « multi paramétrique » (**IRM-mp**). Dans le même temps, pour améliorer la spécificité de l'IRM-mp, un score en 5 points appelé « **score PI-RADS** » a été créé pour différencier les lésions selon leur risque cancéreux (23). Enfin, en parallèle, des techniques de fusion d'images IRM-échographie ont été développées pour guider des biopsies transrectales sur des cibles mises en évidence par l'IRM-mp.

Grâce à ces avancées récentes, l'AFU a publié de nouvelles recommandations pour améliorer la stratégie diagnostique du CaP (13). (**Figure 6**)

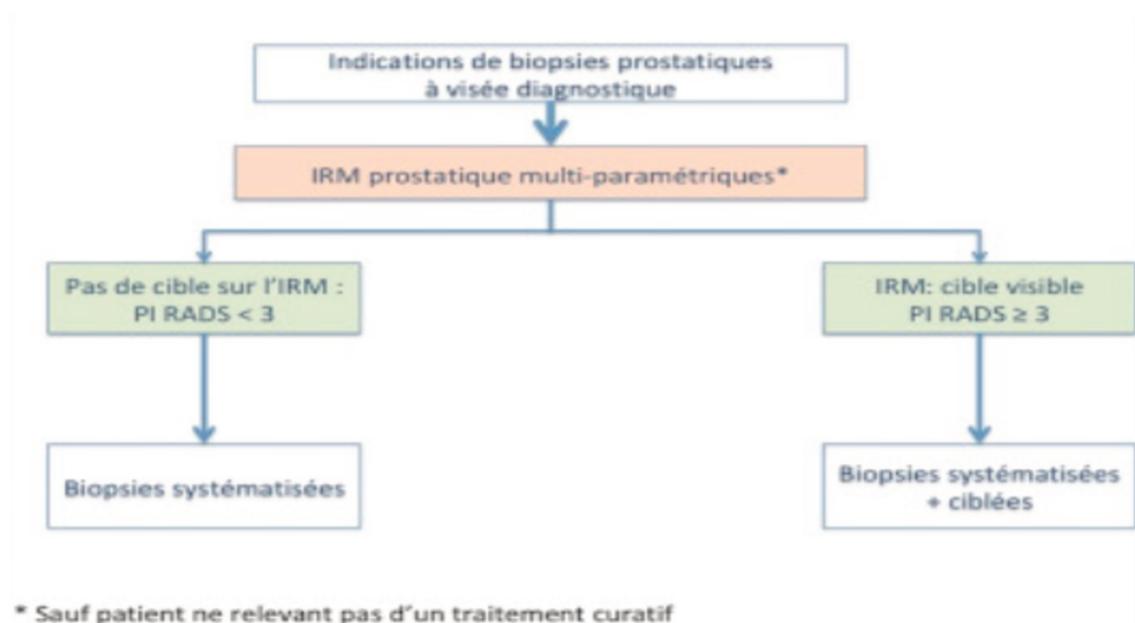


Figure 6: Recommandations de l'AFU sur l'IRM prostatique : stratégie avant biopsie pour la recherche d'un cancer cliniquement significatif.

Une nouvelle indication de l'IRM-mp est la **surveillance active** des CaP à faible potentiel évolutif. Cette surveillance est clinique (TR tous les 3 à 6 mois), biologique (PSA tous les 3 à 6 mois) et anatomopathologique (biopsies tous les 1 à 2 ans). Plusieurs études (24, 25, 26) ont comparé cette surveillance active à des traitements curatifs tels que la prostatectomie radicale et la radiothérapie, et ont rapporté dans des populations de patients à faible et très faible risques de récurrence de CaP, une absence de différence significative en termes de survie et de mortalité à 10 ans.

Les critères d'inclusion de la surveillance active sont :

Cancer à très bas risque évolutif : TR = T1c et PSA < 10 ng/mL et score de Gleason = 6 et biopsies positives < 3 et % tumoral par carotte < 50% et densité PSA <

0.15 (27).

- Cancer à bas risque : TR = T1c-T2a et PSA <10 ng/mL et Gleason = 6.

Ce choix thérapeutique impose de faire une IRM-mp initiale pour conforter la stratégie de prise en charge. En effet, 10% des patients éligibles à une surveillance active sont reclassés par cette première IRM-mp (28). L'IRM-mp peut ensuite faire partie des outils de surveillance sans qu'elle soit actuellement inscrite dans les recommandations. En cas d'évolution défavorable des critères du suivi, la surveillance active doit être interrompue et laisser place à la phase thérapeutique curative (13). Pour certains CaP évolutifs, la surveillance active ne fait donc que déplacer l'éventuel moment du traitement.

B- Réalisation de l'examen :

En l'absence de contre-indication, l'IRM-mp peut être réalisée sur des aimants 1,5 Tesla ou 3 Tesla. Les antennes utilisées sont surtout des antennes externes en réseau phasé de type antenne « corps ou cardiaque ». En raison de leur inconfort et de leur coût, l'usage des antennes endorectales est de nos jours très limité. Il faut aussi s'assurer de l'absence d'intolérance à l'injection d'un produit de contraste gadoliné (allergies, insuffisance rénale sévère...) car le recours à une séquence dynamique injectée est souvent nécessaire.

Il n'y a pas de réel consensus au sujet de la préparation des patients avant la réalisation d'une IRM prostatique. Il existe cependant certaines recommandations (29) :

- L'usage d'antispasmodique pour diminuer les artéfacts de mouvements dus au tube digestif et au rectum,
- La vidange vésicale pour le confort du patient pendant toute la durée de l'examen,
- Une préparation rectale pour s'affranchir des artéfacts de distorsion liés à la présence d'air,
- L'absence d'éjaculation dans les 3 jours précédant l'examen, pour obtenir un meilleur remplissage des vésicules séminales.

C- Protocole :

L'IRM-mp doit comporter des séquences morphologiques pondérées T2 et des séquences fonctionnelles de diffusion et dynamiques après injection IV de Gadolinium (DCE = Dynamic Contrast Enhanced) (29). Une séquence pondérée T2 à champ élargi couvrant l'abdomen et le pelvis, peut être ajoutée pour la recherche d'adénomégalies.

- a- Les séquences pondérées T2:

Ce sont les séquences morphologiques qui permettent l'étude de l'anatomie zonale prostatique. On doit au minimum acquérir 2 plans orthogonaux dans cette pondération. Il s'agit de séquences réalisées en écho de spin rapide (FSE ou TSE). Il est commun de faire appel à des séquences 2D (axiales, sagittales et/ou coronales). **(Figure 7)** Cependant, l'utilisation d'une séquence 3D est envisageable mais avec un risque accru d'artéfacts de mouvement dégradant la qualité des images.



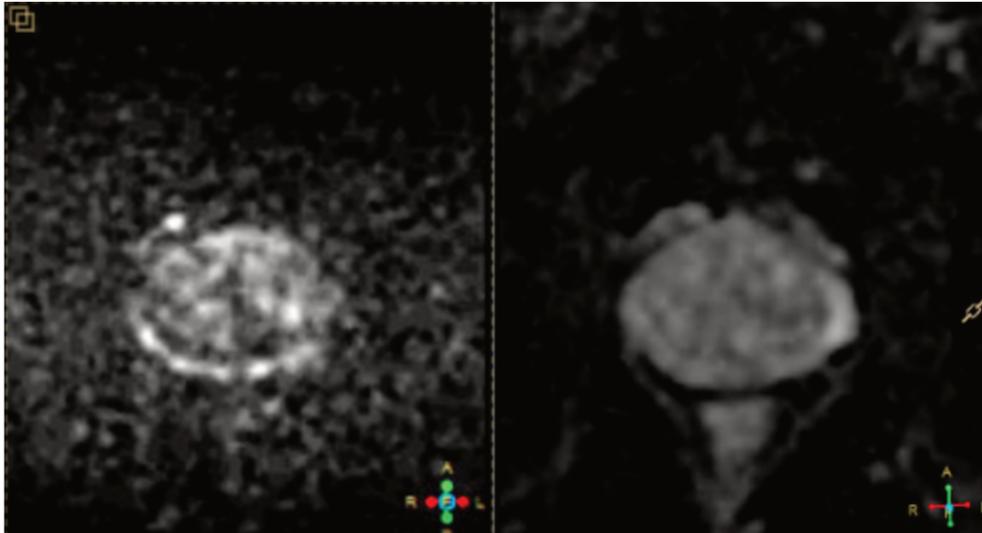
Figure 7: Séquences en pondération T2 axiale à gauche et sagittale à droite
(<https://www.pcih.fr/portal>, Prostate Score Trainer).

- b- Les séquences de diffusion:

Ce sont des séquences fonctionnelles qui explorent l'architecture tissulaire en étudiant les mouvements spontanés et aléatoires des molécules d'eau au sein des tissus. Une perte de signal tissulaire est attendue lorsque l'architecture tissulaire est conservée. Cette décroissance du signal est liée à l'intensité du coefficient de diffusion b exprimé en seconde/mm². La pente de décroissance linéaire du signal est appelée coefficient apparent de diffusion (ADC). Cette donnée est systématiquement fournie sous la forme d'une image associée aux images de diffusion à valeurs de b croissantes. Pour obtenir un coefficient ADC fiable, il est conseillé d'utiliser 3 valeurs de b : une entre 0 et 100 s/mm², une aux alentours de 400 s/mm² et une ≤ 1000 s/mm². Une séquence acquise ou reconstruite avec une valeur de b très élevée ($b = 2000$ s/mm²) est aussi recommandée. **(Figure 8)**

Un tissu normal ayant une architecture respectée aura un signal faible sur les images à b élevé et un coefficient ADC élevé. A l'inverse, un tissu néoplasique ayant perdu son architecture tissulaire normale aura un signal conservé sur les images à b élevé et un

coefficient ADC faible.



*Figure 8: Séquences de Diffusion à $b = 2000 \text{ s/mm}^2$ à gauche et ADC à droite
(<https://www.pcih.fr/portal>, Prostate Score Trainer).*

- c- Les séquences dynamiques en pondération T1 avec injection de Gadolinium (DCE) :

Il s'agit de séquences dynamiques répétées plusieurs fois après une injection intraveineuse de gadolinium. Leur intérêt réside dans l'analyse de la vascularisation tissulaire. Pour obtenir une résolution temporelle faible, il faut disposer de séquences courtes d'une durée $\leq 15 \text{ s}$. Il est donc nécessaire de faire appel à des séquences en écho de gradient 2D ou 3D (EG). Une séquence test avant injection permet de repérer les hypersignaux T1 spontanés liés à des reliquats hémorragiques et sert de masque pour reconstruire les images soustraites. Des soustractions sont en effet très utiles pour bien analyser les prises de contraste au sein du parenchyme prostatique. Après injection, les séquences sont acquises sur une durée totale de 2 minutes au moins.

(Figure 9)

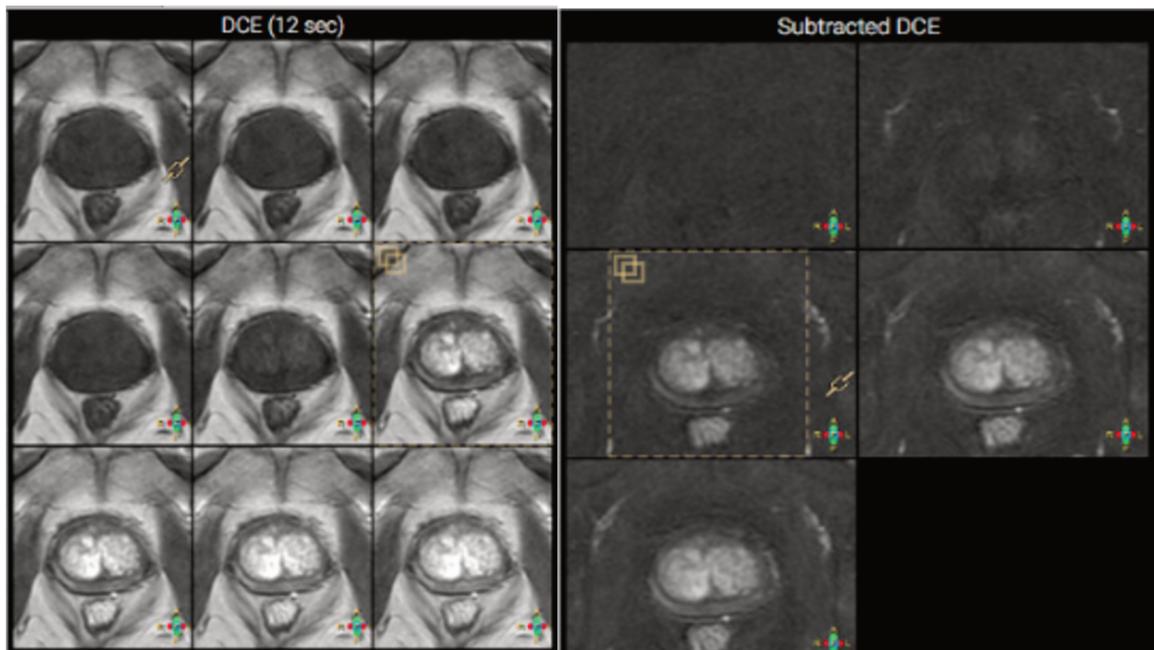


Figure 9: Séquences dynamiques pondérées T1 après injection de Gadolinium (DCE : Dynamic Contrast Enhancement,) : images natives à gauche et soustraites à droite (<https://www.pcih.fr/portal>, Prostate Score Trainer).

D- Score de lecture PI-RADS :

Avant 2012, les radiologues utilisaient pour l'interprétation d'IRM-mp le score de Likert (score subjectif issu des travaux du psychologue Rensis Likert) qui comprend 5 classes (1-2 lésions d'allure bénigne, 3 équivoques et 4-5 suspectes), sans atlas d'image associé. (30)

Pour standardiser et homogénéiser l'acquisition et l'interprétation des IRM-mp, l'American College of Radiology (ACR), l'European Society of Urogenital Radiology (ESUR) et l'AdMeTech Foundation ont créé grâce à un panel d'experts, le score de lecture PI-RADS (Prostate Imaging- Reporting and Data System). Il s'agit d'une classification permettant d'évaluer à partir des données d'une IRM-mp, la probabilité de CaP cliniquement significatif. Cette classification comporte 5 grades côtés de 1 à 5 : du PI-RADS 1 pour les très faibles probabilités de cancer, au PI-RADS 5 pour les très fortes probabilités (29). **(Figure 10)**

PIRADS 1 – Very low (clinically significant cancer is highly unlikely to be present)

PIRADS 2 – Low (clinically significant cancer is unlikely to be present)

PIRADS 3 – Intermediate (the presence of clinically significant cancer is equivocal)

PIRADS 4 – High (clinically significant cancer is likely to be present)

PIRADS 5 – Very high (clinically significant cancer is highly likely to be present)

Figure 10 : Score PI-RADS de 1 à 5

L'application du PI-RADS ne doit pas tenir compte des facteurs cliniques et biologiques comme le toucher rectal et le taux de PSA. Elle doit uniquement se baser sur les données fournies par les séquences d'IRM-mp.

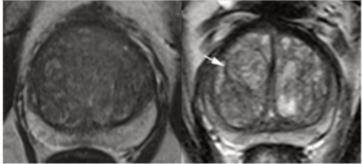
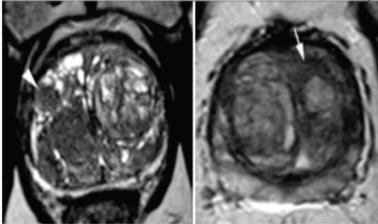
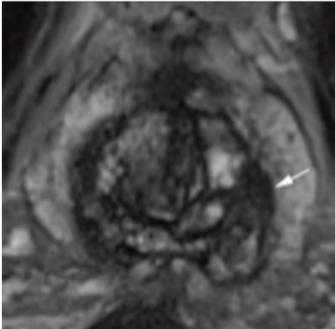
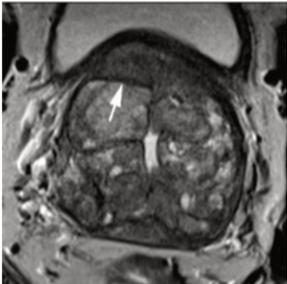
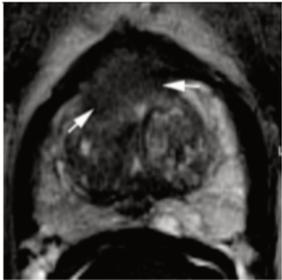
L'intérêt du score PI-RADS est d'apporter une aide à la décision dans l'indication de biopsies ciblées sur une ou plusieurs lésions. Les scores PI-RADS 4 et 5 nécessitent une vérification anatomopathologique ce qui n'est pas le cas pour les scores PI-RADS 1 et 2. Le score PI-RADS 3 étant intermédiaire, la réalisation de biopsies doit s'appuyer sur des données clinico-biologiques.

Trois versions du score PI-RADS ont été publiées jusqu'à ce jour : PI-RADS V1 en 2012 (23), PI-RADS V2 en 2015 (31) et PI-RADS V2.1 en 2019 (29).

Les séquences d'IRM qu'il faut prioritairement analyser pour appliquer le score PI-RADS, sont les séquences T2, diffusion (DWI) et dynamiques en pondération T1 après injection de gadolinium (DCE). Selon la zone de la prostate, il existe une séquence dominante qui est soit le T2, soit le DWI. Les recommandations de la dernière version PI-RADS V2.1, sont :

- a- Séquence T2 :

La séquence T2 est la séquence dominante pour l'analyse de la ZT.

SCORE	Critères T2 dans la ZT
<p>PI-RADS 1</p> 	<ul style="list-style-type: none"> - apparence normale - nodule d'HBP rond complètement encapsulé - nodule dans le nodule - voir Figure 11
<p>PI-RADS 2</p> 	<ul style="list-style-type: none"> - nodule circonscrit incomplètement encapsulé - nodule circonscrit, homogène, non encapsulé - ZT homogène discrètement hypo intense entre les nodules - voir Figure 11
<p>PI-RADS 3</p> 	<ul style="list-style-type: none"> - hétérogénéité de signal mal limitée - ce qui n'est pas inclus dans les critères 1, 2 et 4
<p>PI-RADS 4</p> 	<ul style="list-style-type: none"> - lésion lenticulaire, non circonscrite en hyposignal modéré homogène <15 mm
<p>PI-RADS 5</p> 	<ul style="list-style-type: none"> - similaire mais ≥ 15 mm ou signe d'extension extra prostatique

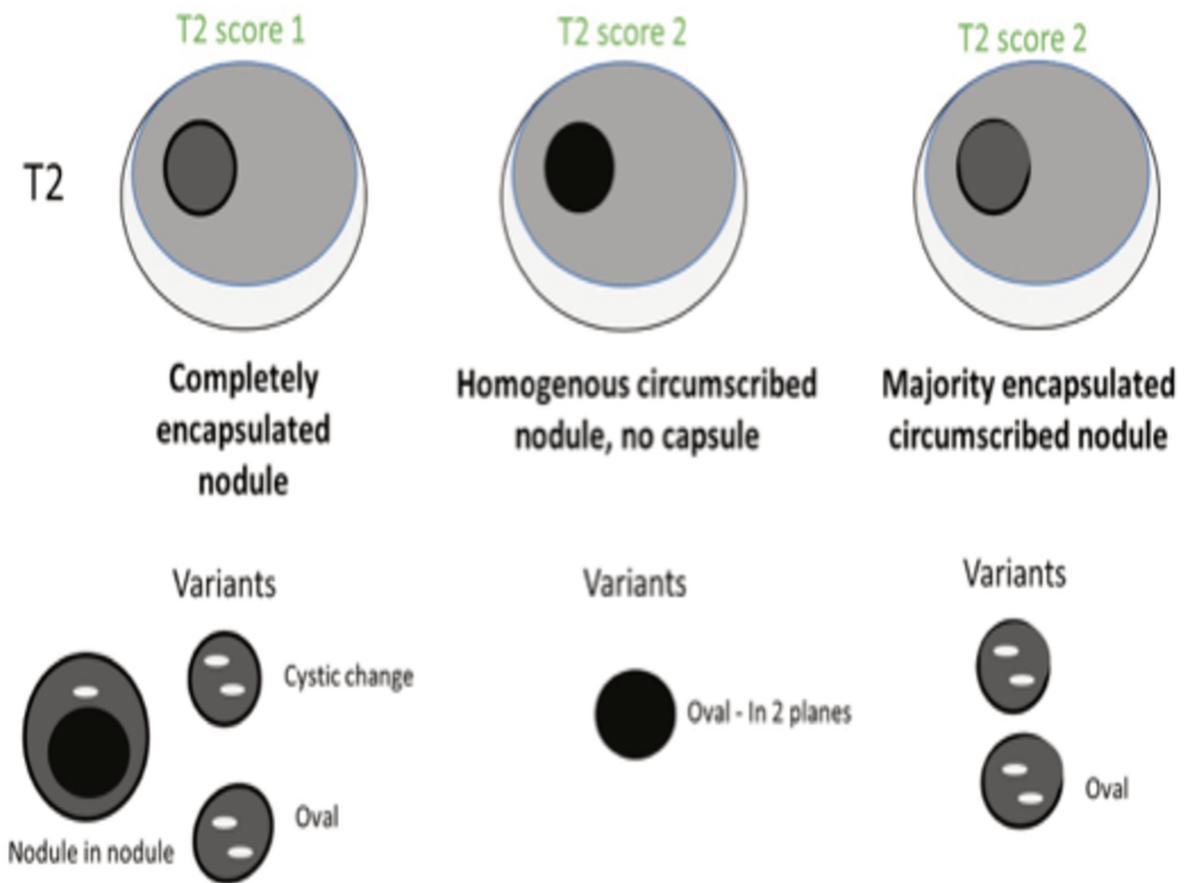
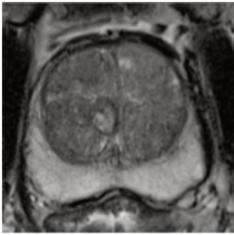
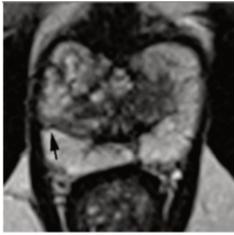
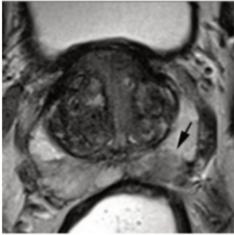
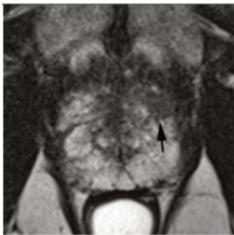
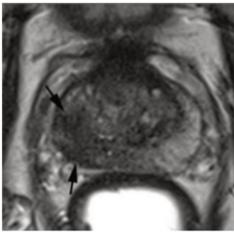


Figure 11: particularités de certains nodules classés PI-RADS 1 et PI-RADS 2 dans la ZT

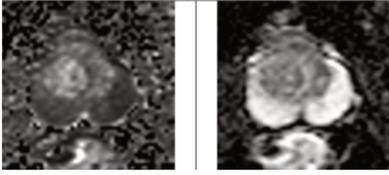
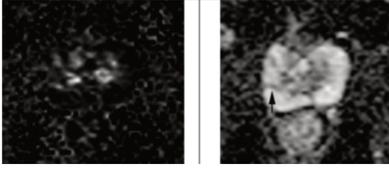
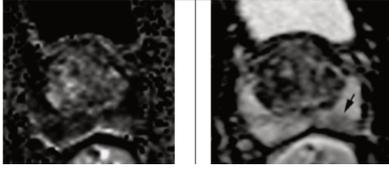
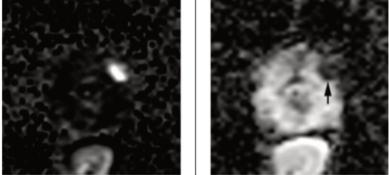
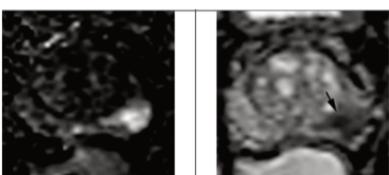
La séquence T2 n'est pas la séquence dominante pour l'interprétation d'une lésion de la ZP. Cependant, en cas d'échec des séquences de diffusion (artefacts liés à une prothèse de hanche par exemple), il est possible d'utiliser la séquence T2 dans la ZP.

SCORE	Critères T2 dans la ZP
PI-RADS 1 	<ul style="list-style-type: none"> - signal uniforme hyperintense
PI-RADS 2 	<ul style="list-style-type: none"> - hyposignal linéaire ou en forme de coin - signaux modérément hypointenses diffus aux marges indistinctes
PI-RADS 3 	<ul style="list-style-type: none"> - hétérogénéité de signal non circonscrite, arrondie, modérément hypointense - ce qui n'est pas inclus dans les critères 1, 2 et 4
PI-RADS 4 	<ul style="list-style-type: none"> - lésion focale circonscrite, homogène, limitée à la prostate, <15 mm
PI-RADS 5 	<ul style="list-style-type: none"> - similaire mais ≥ 15 mm ou extension extra prostatique

- *b- Séquence Diffusion (DWI) :*

La séquence de diffusion est la séquence dominante pour analyser une lésion de la ZP.

Les critères détaillés ci-dessous sont aussi utilisés pour certaines lésions de la ZT.

SCORE	Critères ZP et ZT
PI-RADS 1 	<ul style="list-style-type: none"> - absence d'anomalie de signal en diffusion à b élevé ou sur la reconstruction ADC
PI-RADS 2 	<ul style="list-style-type: none"> - hyposignal linéaire ou en forme de coin sur l'ADC ou hyposignal linéaire ou en forme de coin sur la diffusion.
PI-RADS 3 	<ul style="list-style-type: none"> - hyposignal focal discret sur l'ADC et/ou hypersignal discret sur la diffusion. - hyposignal focal franc en ADC ou hypersignal franc en diffusion mais pas les deux.
PI-RADS 4 	<ul style="list-style-type: none"> - hyposignal focal marqué sur l'ADC associé à un hypersignal focal marqué sur la diffusion à haute valeur de b, < 15 mm dans ses plus grandes dimensions.
PI-RADS 5 	<ul style="list-style-type: none"> - similaire mais ≥ 15 mm ou signe d'extension extra prostatique

Pour les nodules de la ZT classés PI-RADS 2 et 3, la séquence DWI permet d'upgrader le score en 3 et 4 dans un second temps comme schématisé ci-dessous (**Figure 12**) :

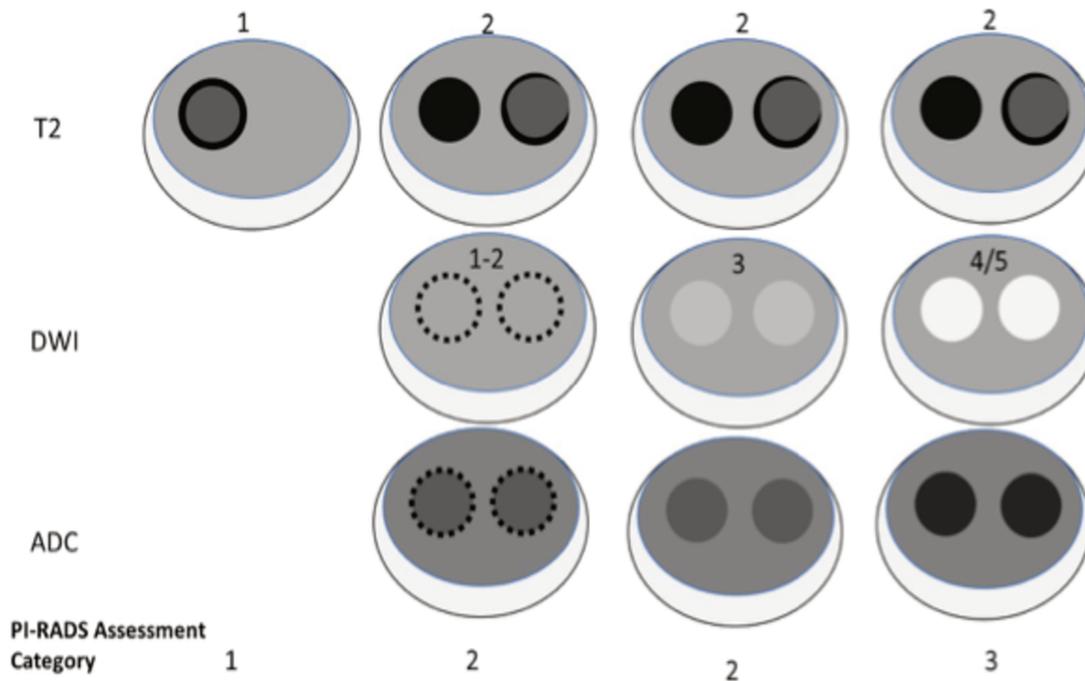


Figure 12 : Règles d'upgrading de nodules classés PI-RADS 2 et PI-RADS 3 dans la ZT

Le tableau suivant détaille l'influence respective des séquences T2, DWI et DCE dans l'analyse de la ZT. La séquence dominante T2 peut être upgradée par la séquence DWI dans certaines lésions PI-RADS 2 et 3. Pour les lésions PI-RADS 1, 4 et 5, la séquence T2 se suffit à elle-même.

T2	DWI	DCE	Overall
1	Any	Any	1
2	≤3	Any	2
	≥4	Any	3
3	≤4	Any	3
	5	Any	4
4	Any	Any	4
5	Any	Any	5

- c- Séquence dynamique pondérée T1 après injection de Gadolinium (DCE) :

Elle est uniquement utilisée pour analyser les lésions de la ZP.

Les critères positifs sont la présence d'une prise de contraste focale, précoce et contemporaine au tissu prostatique adjacent normal, et correspondant à une lésion suspecte sur les séquences T2 et diffusion.

Les critères négatifs sont une absence de prise de contraste précoce ou contemporaine au tissu prostatique adjacent normal, ou un rehaussement multifocal, diffus du tissu prostatique ne correspondant pas à une image suspecte sur les séquences T2 et diffusion.

Le tableau suivant détaille l'influence respective des séquences T2, DWI et DCE dans l'analyse de la ZP. La séquence dominante DWI peut être upgradée par la séquence DCE dans certaines lésions PI-RADS 3. Pour les lésions PI-RADS 1, 2, 4 et 5, la séquence DWI se suffit à elle-même.

Peripheral Zone (PZ)			
DWI	T2W	DCE	PIRADS
1	Any*	Any	1
2	Any	Any	2
3	Any	- +	3 4
4	Any	Any	4
5	Any	Any	5

- *d- Cas particuliers de la ZC et du SFMA :*

Deux zones sortent du cadre standard décrit ci-dessus pour la ZT et la ZP. Il s'agit de la zone centrale (ZC) et du stroma fibromusculaire antérieur (SFMA).

La ZC est en hyposignal T2 et hypersignal diffusion discret à b élevé. Elle encercle les canaux éjaculateurs de façon symétrique à la base prostatique (**Figure 13**). Toute asymétrie de la ZC constatée sur les séquences T2, DWI /ADC et /ou DCE doit faire suspecter un cancer.

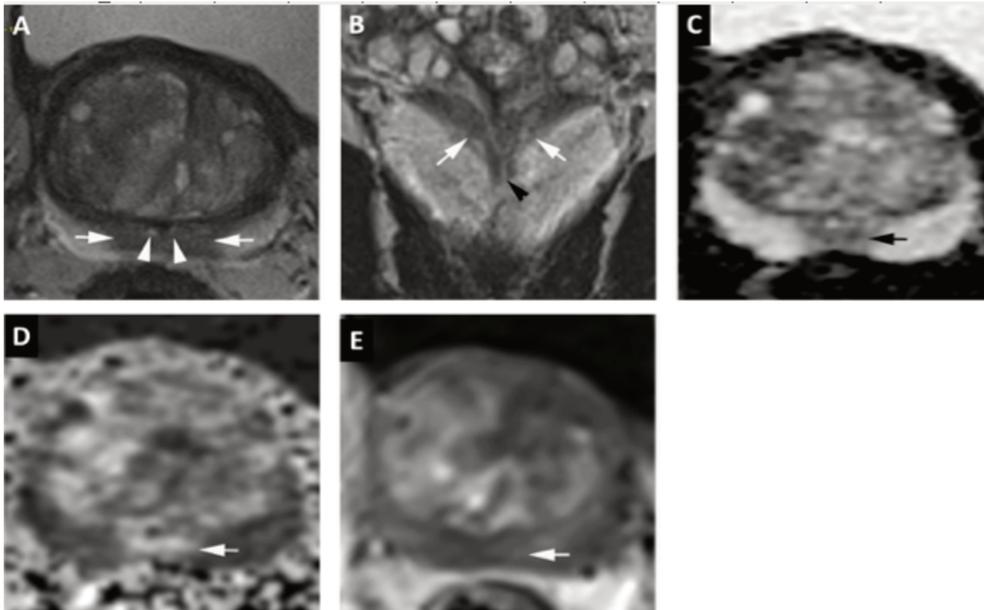


Figure 13 : ZC normale A. T2 axiale, B. T2 coronale, C. ADC, D. Diffusion, E. DCE.
(PI-RADS v2.1 2019)

Le SFMA est naturellement en hyposignal T2 franc, similaire aux muscles obturateurs, en hyposignal DWI sans restriction de l'ADC et sans prise de contraste après injection de Gadolinium sur la séquence DCE (**Figure 14**). De la même manière que la ZC, toute anomalie de signal du SFMA doit être considérée comme suspecte surtout s'il existe un CaP au sein de la ZP ou de la ZC à son contact.

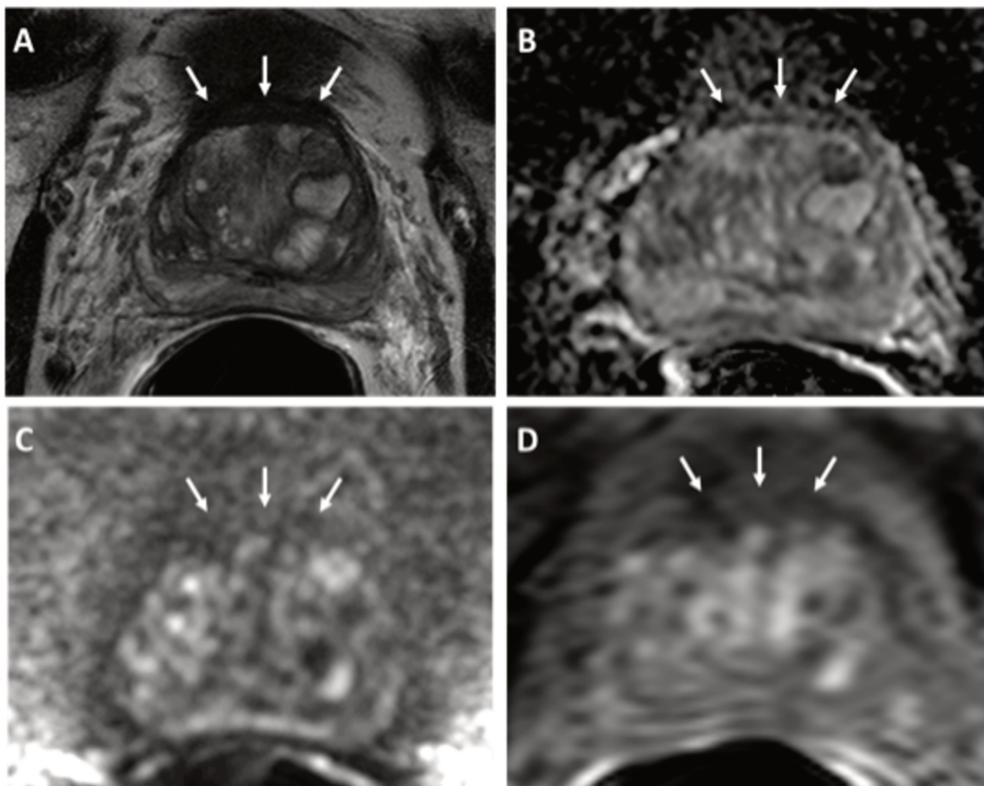


Figure 14 : SFMA d'aspect normal. A. T2 axiale, B. ADC, C. Diffusion, D. DCE.
(PI-RADS v2.1 2019)

E) Cartographie prostatique :

Selon la version 2 du score PI-RADS, la prostate a été divisée en 41 secteurs (**Figure 15**) :

- 38 secteurs intra prostatique (soit 19 à droite et 19 à gauche),
- 2 secteurs pour les vésicules séminales,
- 1 secteur pour l'urètre membraneux.

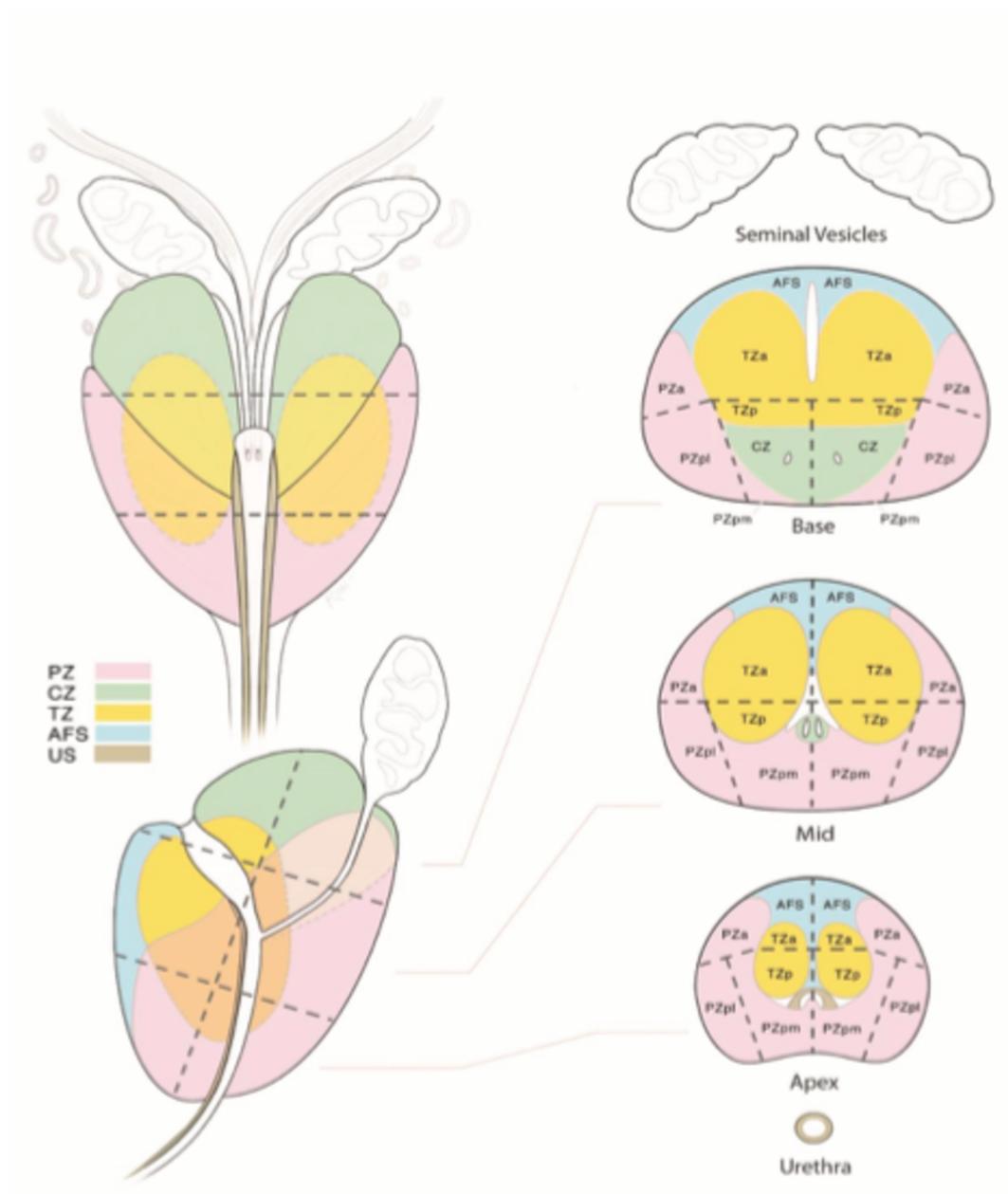


Figure 15 : Anatomie zonale de la prostate en 41 secteurs.

PI-RADS v2.1, 2019

3) Plateforme internet « Prostate Score Trainer » (PST) :

A- Généralités sur le PST :

Le PST est un outil de standardisation de la lecture des IRM-mp et d'apprentissage du score PI-RADS v2.1. Il s'agit d'une plateforme informatique élaborée en France (OVH, Roubaix) et dont le cahier des charges (32) imposait la mise à disposition des lecteurs, d'images provenant d'IRM-mp anonymisées sans autres données cliniques et biologiques (âge des patients, TR, taux de PSA...). Cet outil est disponible en ligne en entrant le lien <https://www.pcih.fr/portal/> dans la barre de recherche d'un navigateur Internet (**Figure 16**).



Figure 16 : Portail d'accueil sur <https://www.pcih.fr/portal/>

Lors de la première connexion, le lecteur doit s'inscrire sur le site et suivre les instructions de la page d'accueil en cliquant sur "*pas encore inscrit*". Il reçoit ensuite un mail confirmant l'inscription au site "Lille Imaging Portal", et l'informant de la validation de ses log-in (adresse mail) et mot de passe (celui choisi au moment de l'inscription).

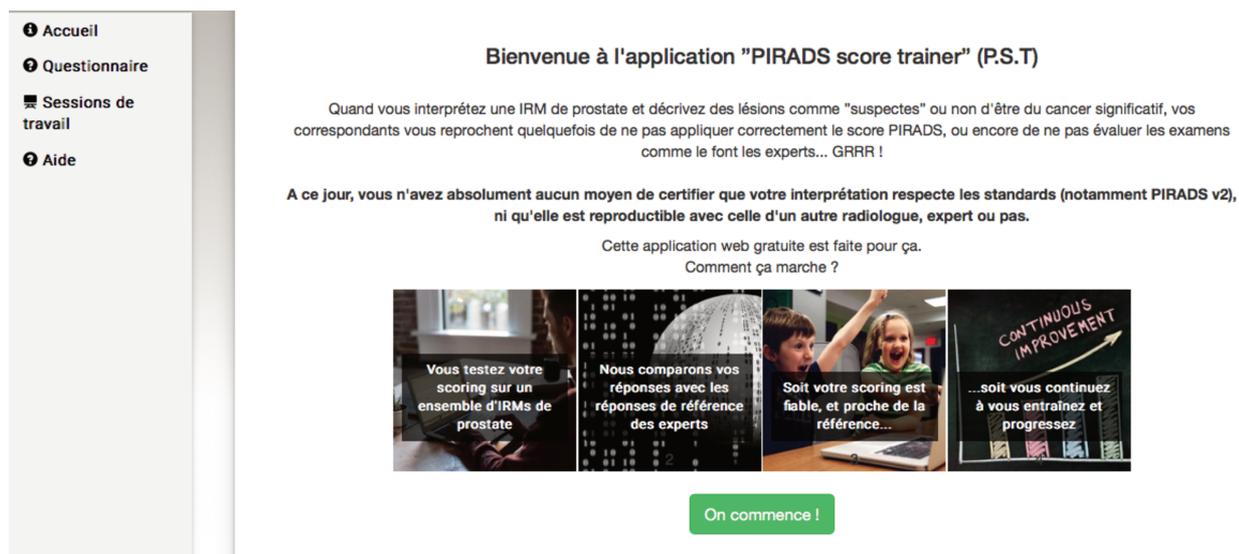


Figure 17 : Page d'accueil du PST après connexion avec son log-in et son mot de passe.

Une fois connecté, l'utilisateur arrive sur la page d'accueil du PST et clique sur l'icône « *On commence !* » pour ouvrir une session de travail (Figure 17). Lors de la première ouverture, l'utilisateur répond à un questionnaire (annexe 1) évaluant sa pratique de l'imagerie prostatique et sa connaissance du score PI-RADS v2.1. Ce questionnaire composé de 20 questions, permet de préciser le profil du lecteur.

Le **consentement du lecteur** à la participation à cette étude est recueilli conformément au Règlement Général de Protection des Données (RGPD) du 25 mai 2018.

Une session de travail se présente sous la forme de 50 dossiers provenant de 50 patients consécutifs. Le lecteur doit évaluer 2 lésions prostatiques par patient, chaque lésion étant identifiée par un cercle rouge. Ces lésions ont été préalablement choisies par un expert qui en a déterminé le score PI-RADS v2.1 et la localisation.

Chaque cas comporte plusieurs images JPEG : 2 plans T2, 1 plan axial DWI, 1 plan axial ADC, plusieurs images axiales DCE sans et avec soustractions.

Pour chaque patient et pour chacune des deux lésions, le lecteur doit se prononcer sur le score PI-RADS v2.1 et sur la localisation lésionnelle en prenant pour référence la segmentation prostatique en 41 secteurs (Figure 18).

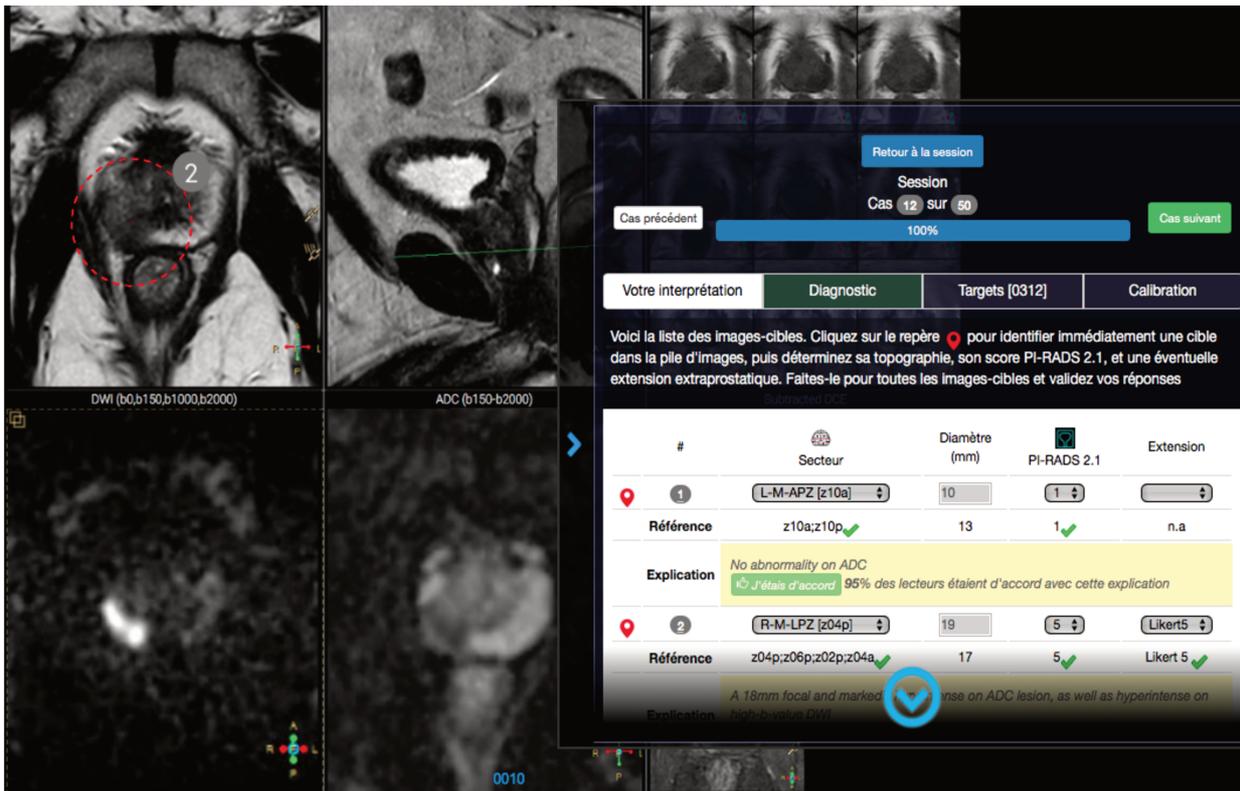


Figure 18 : Session de travail avec grille de remplissage à droite de l'écran.

L'analyse de chaque cas dure environ 5 minutes. La totalité des 50 dossiers peut se faire en une ou plusieurs fois, le parcours du lecteur étant mémorisé à chaque déconnexion. L'analyse complète des 50 cas dure entre 3 et 5 heures.

B- Présentation des résultats aux lecteurs:

Une fois ses réponses validées, le lecteur est informé de la comparaison de ses réponses avec celles de l'expert (✔️ si concordance, ❌ si discordance). Le lecteur a alors la possibilité d'indiquer s'il est d'accord ou non avec l'expert en cliquant sur des icônes "Je suis d'accord" ou "Je ne suis pas d'accord" (Figure 19).

#	Secteur	Diamètre (mm)	PI-RADS 2.1	Extension
1	L-M-LPZ [z10p]	10	3	
Référence	z10p;z08p;z09p;z07p ✓	15	3 ✓	n.a
Explication	<p>Moderately hypointense on ADC, and mildly hyperintense on high-b-value DWI. Enhancement is not focal : remains PI-RADS 3</p> <p>👍 J'étais d'accord 88% des lecteurs étaient d'accord avec cette explication</p>			
2	L-M-ATZ [z09a]	10	1	
Référence	z09a ✓	11	2 ✗	n.a
Explication	<p>This is a homogenous mildly hypointense area between 2 (benign) nodules : definition of PI-RADS 2</p> <p>👍 J'étais d'accord 87% des lecteurs étaient d'accord avec cette explication</p>			

Figure 19 : comparaison des réponses du lecteur avec celles de l'expert :

✓ si les réponses sont similaires, ✗ si les réponses diffèrent.

Opinion du lecteur "Je suis d'accord" ou "Je ne suis pas d'accord".

Les résultats du lecteur s'affichent sous la forme de 2 tableaux. Pour chacun des deux tableaux, l'ordonnée représente les réponses du lecteur et l'abscisse celles de l'expert. Le tableau 1 (**Figure 20**) détaille l'ensemble des 5 valeurs du PI-RADS v2.1. En regroupant les valeurs 1 et 2 et les valeurs 4 et 5, le tableau 2 (**Figure 21**) illustre 3 types de lésions : les lésions probablement bénignes (PI-RADS 1 et 2), les lésions indéterminées (PI-RADS 3) et les lésions probablement malignes (PI-RADS 4 et 5). Cette représentation avec seulement 3 catégories de lésions, a pour but de simplifier les prises de décisions à partir des données de l'IRM-mp.

Pour chaque tableau, une diagonale de cases vertes correspond aux lésions classées de façon identique par le lecteur et l'expert. Au-dessus de cette diagonale, les lésions sous-classées par le lecteur, sont représentées en bleu et en-dessous de cette diagonale, les lésions sur-classées par le lecteur sont représentées en rouge. Le lecteur est renseigné sur le nombre de ses réponses en accord et en désaccord avec l'expert. Son taux d'accord est illustré par la valeur PP appelée probabilité de bonne réponse ou PbonREP, et par le taux de concordance « Weighted Kappa » (WK) avec un intervalle de confiance. Chaque lecteur peut lire ses statistiques tous les 10 ou 20 cas et voir son résultat final à la fin des 50 cas.

		Référence					Total
		PI-RADS 1	PI-RADS 2	PI-RADS 3	PI-RADS 4	PI-RADS 5	
Vous	PI-RADS 1	4	1	1	0	0	6
	PI-RADS 2	3	11	2	0	0	16
	PI-RADS 3	0	7	10	1	0	18
	PI-RADS 4	0	0	1	6	0	7
	PI-RADS 5	0	0	1	1	11	13
Total		7	19	15	8	11	60
Exactitude selon score		57.1%	57.9%	66.7%	75.0%	100.0%	

PP (pourcentage de lésions **Parfaitement en accord avec le score de référence**) = 42/60 = **70%**
 PO (pourcentage de lésions **Sur-classant le score de référence**) = 13/60 = 22%
 PU (pourcentage de lésions **Sous-classant le score de référence**) = 5/60 = 8%
 Weighted Kappa **wk=0.88**: Accord presque parfait (wk]0.8-1])
 wk confidence interval (95%) : [0.69-1]

Figure 20 : Tableau 1 en 5 catégories : résultats du lecteur (Vous) par rapport à l'expert (Référence) avec valeurs du PbonREP et du WK.

		Référence			Total
		PI-RADS 1-2	PI-RADS 3	PI-RADS 4-5	
Vous	PI-RADS 1-2	35	4	1	40
	PI-RADS 3	9	13	2	24
	PI-RADS 4-5	1	3	32	36
Total		45	20	35	100
Exactitude selon score		77.8%	65.0%	91.4%	

PP (pourcentage de lésions **Parfaitement en accord avec le score de référence**) = 80/100 = **80%**
 PO (pourcentage de lésions **Sur-classant le score de référence**) = 13/100 = 13%
 PU (pourcentage de lésions **Sous-classant le score de référence**) = 7/100 = 7%
 Weighted Kappa **wk=0.83**: Accord presque parfait (wk]0.8-1])
 wk confidence interval (95%) : [0.71-0.96]

Figure 21: Tableau 2 en 3 catégories : résultats du lecteur (Vous) par rapport à l'expert (Référence) avec valeurs du PbonREP et du WK.

Concernant la localisation des lésions, le taux de concordance entre le lecteur et l'expert est illustré par un pourcentage appelé **PTopoOK (Figure 22)**.

**Précision pour la localisation
des images-cibles**

72% (dont 50% parfaitement localisées)
Précision requise idéalement : $\geq 70\%$

Figure 22 : taux de concordance entre le lecteur et l'expert : PTopoOK

4) Objectifs de l'étude :

Devant l'utilisation croissante de l'IRM-mp dans le dépistage du CaP, le score PI-RADS a été créé dans le but d'homogénéiser la réalisation et l'interprétation des examens, et de diminuer le taux de variabilité entre les sites et les lecteurs. L'évolution de ce score s'est progressivement effectuée dans une logique d'amélioration des performances. De nombreuses études ont été publiées sur les versions 1 et 2 du PI-RADS et se sont concentrées sur la notion de reproductibilité.

Actuellement la version 2.1 est la 3^{ème} version du PI-RADS. Sa publication récente doit nous conduire à tester sa reproductibilité pour savoir si ce score atteint son objectif de standardisation de lecture des IRM-mp.

Le logiciel « Prostate Score Trainer (PST) » a été élaboré sous la forme d'une plateforme internet. Il a pour but d'apprendre aux lecteurs, les critères de la version 2.1. Il s'agit avant tout, d'un outil d'auto-évaluation mais il permet dans le même temps, d'analyser la reproductibilité du score PI-RADS dans une large population d'utilisateurs.

Notre étude avait pour objectif principal, d'analyser la reproductibilité inter-observateurs du PI-RADS v2.1 en utilisant la plateforme PST comme outil d'évaluation. Cette étude a fait appel à un panel de 52 lecteurs plus ou moins expérimentés dans la lecture des IRM-mp et issus de multiples centres d'imagerie français.

Ses objectifs secondaires étaient :

- d'évaluer la progression des performances des lecteurs grâce à la plateforme PST,
- d'évaluer le nombre de cas nécessaires pour atteindre le meilleur niveau possible,
- de proposer à terme un outil d'auto-évaluation.

III- Matériels et Méthodes:

1) Population des lecteurs:

Des invitations à participer à notre étude ont été envoyées à des radiologues seniors et juniors français ayant des connaissances en imagerie génito-urinaire. Ont été ciblés les membres de la SIGU (Société française d'Imagerie Génito-Urinaire) et de l'ESUR (European Society of Urogenital Radiology), des radiologues ayant récemment participé à un EPU (Enseignement Post Universitaire) sur l'imagerie de la prostate à Lille et des internes ayant validé un stage d'imagerie génito-urinaire à Lille ou dans d'autres CHU de France. Les invitations ont été envoyées par mail par le Pr PUECH à partir du 1^{er} novembre 2019. Après acceptation, les lecteurs recevaient un lien de connexion à la plateforme PST et des recommandations écrites. Ils étaient informés qu'ils avaient au plus 3 mois pour finir les 50 cas et que l'étude serait clôturée à la fin du mois de janvier 2020.

Sur les 235 invitations envoyées, 71 réponses positives ont été reçues.

Sur ces 71 radiologues, 2 lecteurs se sont finalement récusés, 12 n'ont pas ouvert leur session de travail et 5 ne l'ont pas achevé dans le temps imparti.

52 lecteurs ont donc terminé les 50 cas (**Figure 23** : flow-chart des lecteurs).

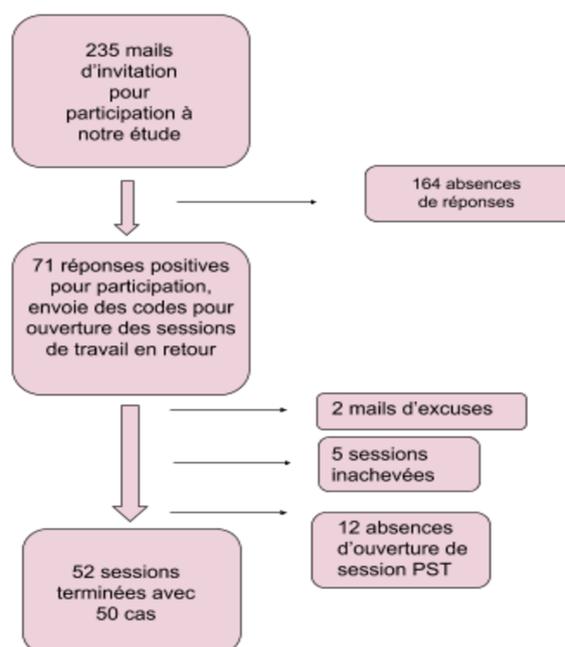


Figure 23: Flow-chart des lecteurs

Les données des 52 lecteurs ont été traitées dans le respect du Règlement Général sur la Protection des Données RGPD (25 Mai 2018) c'est-à-dire de façon anonyme, avec possibilité de retrait à tout moment. Nous avons évalué pour chaque lecteur :

- le sexe,
- le niveau de formation (interne, CCA, radiologue sénior PH/PUPH/libéral),
- le type d'activité professionnelle : publique, libérale ou mixte,
- le niveau de compétence en imagerie génito-urinaire : spécialisé en imagerie génito-urinaire, partiellement spécialisé en imagerie génito-urinaire, activité généraliste,
- l'ancienneté de la 1^{ère} IRM-mp : < 1 an, > 1 an, > 3 ans,
- le nombre d'IRM-mp réalisé par an : <50, entre 50 et 100, entre 100 et 200,
- le délai depuis la dernière formation en IRM-mp : < 1 an, > 1 an, > 3 ans,
- la participation aux RCP de cancérologie génito-urinaire.

2) Données du PST :

Patients et IRM :

Ont été incluses les données IRM de 50 patients consécutifs adressés au Centre Hospitalo-Universitaire Régional de Lille (Hôpital Claude Huriez) pour suspicion de CaP et ayant bénéficié d'une IRM-mp entre le 5 Janvier et le 12 Février 2016. Tous les patients avaient été informés par une note explicative, de l'éventualité de l'utilisation des données de leur imagerie dans un but scientifique (attestation CNIL n°DEC18-366).

Les IRM-mp ont été réalisées sur une IRM 1.5 Tesla Ingenia (Philips Healthcare, Koninklijke, the Netherlands) avec un protocole d'acquisition en accord avec les exigences du score PI-RADS v2.1 :

- une antenne abdominale à 32 canaux en réseau phasé.
- Séquences en pondération 2D T2 TSE dans les 3 plans, axial, sagittal et coronal : TR = 3000ms, TE = 110ms, épaisseur de coupes = 3 mm, FOV = 180x180 mm, matrice = 224x220.
- Séquence de diffusion EPI dans le plan axial avec multiples valeurs de b (0,150, 1000 s/mm²) donnant une cartographie ADC, et une séquence avec une seule valeur à 2000 s/mm² : TR = 4347 ms, TE = 95 ms, épaisseur de coupes = 4 mm, FOV = 200 x178mm, matrice = 68x59.

- Séquence dynamique EG en pondération T1 (DCE) dans le plan axial avec injection IV de gadolinium (Dotarem, Guerbet, Villepinte, France) et un débit d'injection à 2 mL/s : TR = 8,9ms, TE = 4,2 ms, angle de bascule = 17°, épaisseur de coupe = 3 mm, FOV = 160x150mm, matrice = 128 x114, résolution temporelle = 15 secondes, durée totale = 5min
- Deux séquences en pondération T2 EG et Diffusion sur le pelvis pour la recherche d'adénopathies.

Choix et répartition des lésions « cible » :

Tous les examens ont été relus en utilisant le score PI-RADS v2.1 sans tenir compte des données du compte rendu initial qui utilisait le score PI-RADS v2.

Deux lésions par patient ont été sélectionnées et 100 lésions en tout ont été ciblées dans le PST. Ces lésions étaient réparties dans les 4 zones prostatiques (ZP, ZT, SFMA et ZC). Pour tenir compte de la fréquence du CaP dans ces 4 zones, ont été choisies 65 lésions dans la ZP, 30 lésions dans la ZT, 4 lésions dans le SFMA et 1 lésion dans la ZC (**Figure 24**).

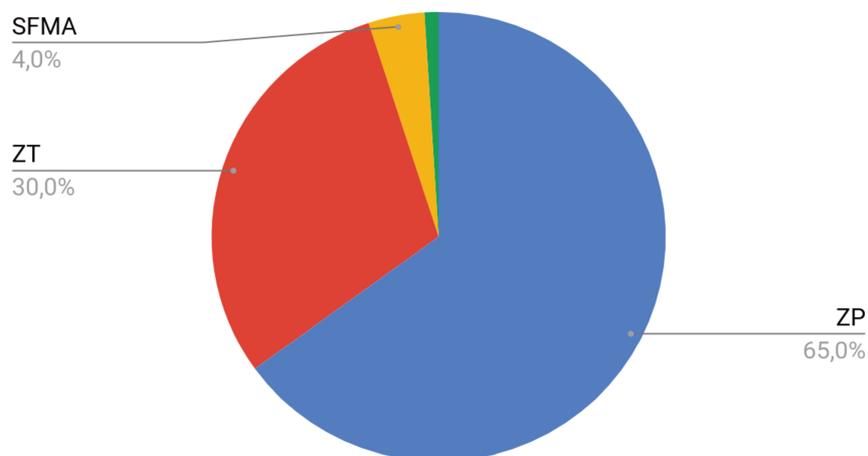


Figure 24 : Répartition des lésions « cible » dans les 4 zones prostatiques

Selon le score PI-RADS v2.1, nous avons sélectionné (**Figure 25**) :

- Dans la ZT : 8 lésions PI-RADS 1, 8 lésions PI-RADS 2, 9 lésions PI-RADS 3, 4 lésions PI-RADS 4 et 1 lésion PI-RADS 5.
- Dans la ZP : 10 lésions PI-RADS 1, 15 lésions PI-RADS 2, 11 lésions PI-RADS 3, 14 lésions PI-RADS 4 et 15 lésions PI-RADS 5.
- Dans le SFMA : 2 lésions PI-RADS 2, 1 lésion PI-RADS 3 et 1 lésion PI-RADS 4.
- Dans la ZC : 1 lésion PI-RADS 5.

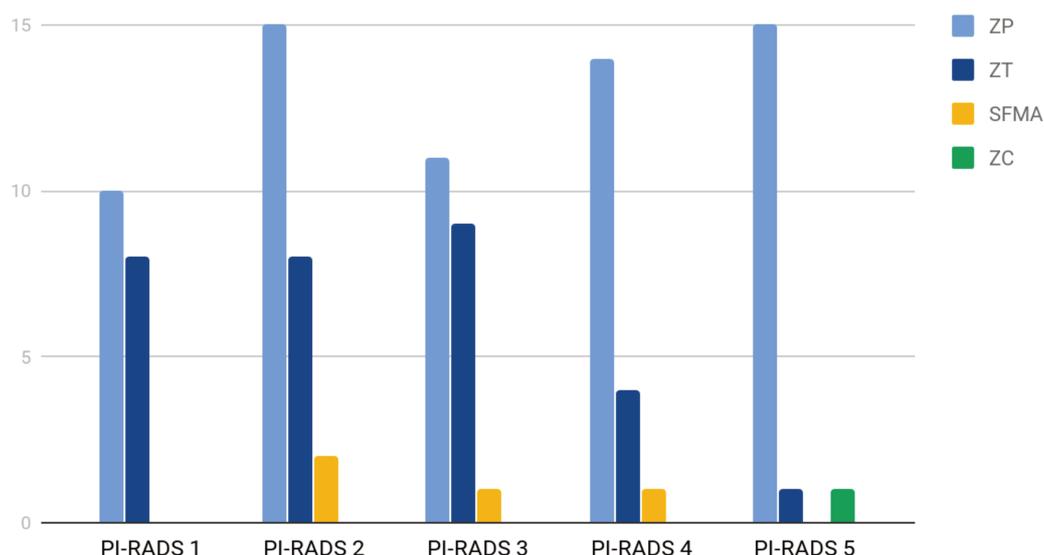


Figure 25 : Répartition des lésions cible selon PI-RADS v2.1

Préparation des images:

Les images de chaque IRM-mp ont été anonymisées et préparées entre Janvier et Juin 2019 par 3 lecteurs (Pr Philippe PUECH, Dr Jean-Philippe BRIDAULT et Julie BRUNEREAU, interne).

Après avoir été sélectionnées, les images ont été capturées au format JPEG à partir d'un serveur Portal (Philips Healthcare, Koninklijke, The Netherlands), et affichées de façon standardisée sur deux lignes de 3 images chacune (**Figure 26**) :

- Ligne supérieure :

a) 1 image T2 axiale, b) 1 image T2 sagittale, c) 9 images T1 DCE natives

- Ligne inférieure :

a) 1 image diffusion axiale, b) 1 image ADC axiale; c) 5 images T1 DCE avec soustraction.

Le fenêtrage et les niveaux de zoom étaient figés. Seule l'image T2 sagittale affichait le niveau de coupe. Toutes les images étaient calibrées, pour les mesures de distance en mm, en cliquant sur la lettre "d" du clavier.

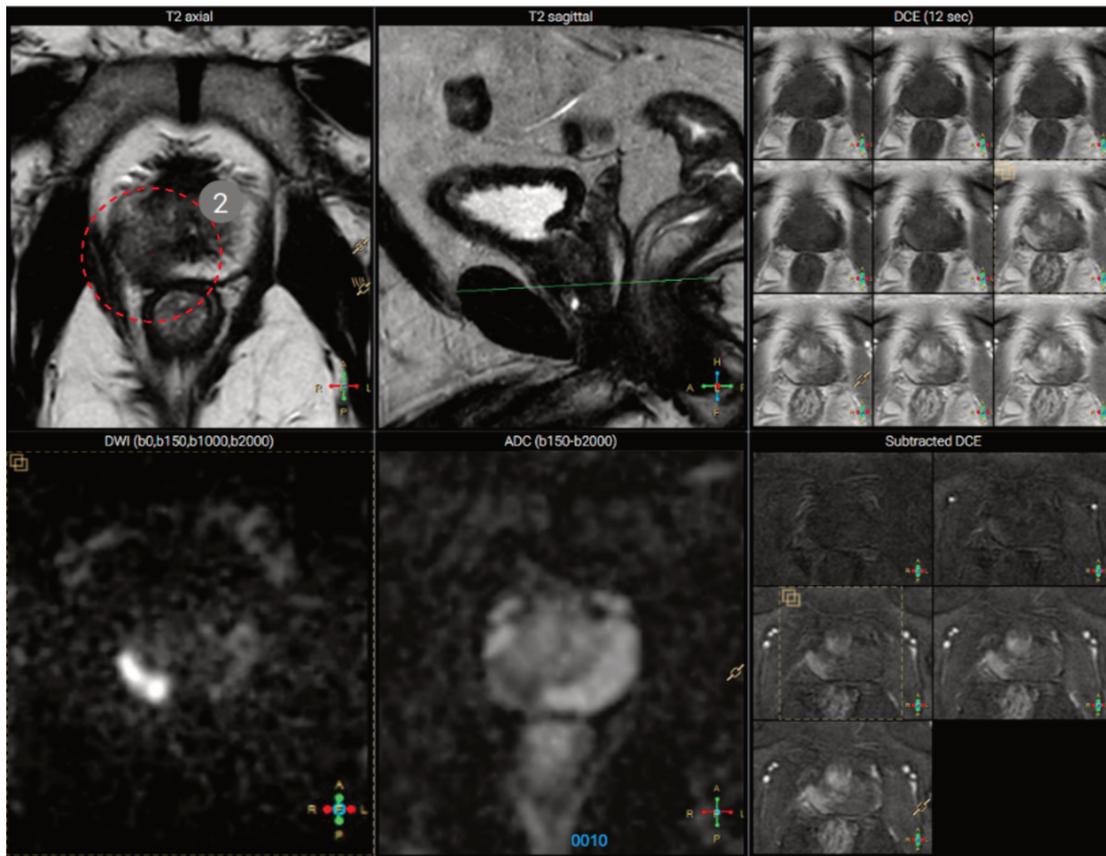


Figure 26 :

- Images de la ligne supérieure / de gauche à droite : image axiale T2, image sagittale T2, images DCE natives. A noter un cercle rouge ciblant une lésion à analyser sur l'image axiale T2.
- Images de la ligne inférieure / de gauche à droite : image axiale diffusion b2000, image axiale ADC, images DCE avec soustraction

Les données de l'étude (réponses) ont été conservées dans une seule table d'une base de données MySQL sécurisée par un login/mdp et cryptée. Elles n'étaient accessibles qu'aux investigateurs (Pr Philippe PUECH et Julie BRUNEREAU - interne), pendant la durée de l'étude.

3) Analyse statistiques:

A l'issue de cette étude, nous disposons des réponses des 52 lecteurs pour les 100 lésions, ce qui représentait 5200 réponses pour le score PI-RADS v2.1 et 5200 réponses pour la localisation lésionnelle.

Pour chaque lésion et chaque lecteur, nous disposons des paramètres suivants :

Le numéro de lecteur

Le numéro de cible

Le rang de lecture (utilisé pour tester l'effet d'apprentissage)

La topographie de la lésion : ZP ou ZT (en incluant SFMA et ZC dans cette zone)₄₀

- TopoOK: variable binaire précisant si le lecteur a positionné la lésion dans la bonne zone
- PI-RADSRef : PI-RADS de référence et considéré comme « Gold Standard »
- PI-RADS fourni par le lecteur

On peut donc évaluer si le lecteur a fourni la bonne réponse (bonrep) et la différence entre le PI-RADS fourni par le lecteur et le PI-RADS de référence (PI-RADSdiff).

Deux autres variables, le diamètre et le temps de passation, ont été recueillies mais non traitées dans l'analyse statistique.

En analysant, pour chaque lecteur, leurs réponses aux 100 lésions, nous avons pu calculer, 4 indicateurs de performance :

- PTopoOK : pourcentage de lésions localisées dans la bonne zone
- PbonREP : pourcentage de bonnes réponses (PI-RADS fourni par le lecteur identique au PI-RADS de référence)
- RMSE : erreur quadratique moyenne, écart normalisé au PI-RADS de référence, soit la racine carrée de la somme des différences (PI-RADS observateur - PI-RADS référence) au carrée sur le nombre de réponse n :
$$\sqrt{\frac{1}{100} \sum_1^{100} PIRADSdiff^2}$$
- Le Weighted Kappa de Cohen, indice de concordance, calculé pour des tableaux de contingence de taille 3x3 ou plus (33).

Ces différents indicateurs ont été utilisés pour l'évaluation du score PI-RADS.

Nous avons évalué ces paramètres en fonction des 5 grades du PI-RADS (PI-RADS 1, 2, 3, 4 et 5) et en regroupant ces 5 grades en 3 catégories (PI-RADS 1&2, PI-RADS 3, PI-RADS 4&5).

Une analyse descriptive a ensuite été réalisée. Les variables étant fortement asymétriques, elles ont été résumées par les quartiles (Q1, médiane, Q3).

Nous avons également cherché à répondre à 4 autres questions :

1) Existe-t-il plusieurs profils de lecteurs ? Pour ce faire, nous avons réalisé une classification hiérarchique ascendante sur les 4 indicateurs calculés précédemment, puis avons comparé les 4 paramètres selon les clusters obtenus (box-plot, test de

Wilcoxon ou Kruskal-Wallis)

2) Existe-t-il des groupes de lésions différentes, notamment en termes de difficulté de lecture. Nous avons utilisé la même méthodologie que pour la question 1.

3) Les indicateurs de performance sont-ils différents selon les lésions ZP et ZT ? Les indicateurs ont été comparés selon les groupes ZP et ZT à l'aide du test de Wilcoxon

4) Existe-t-il un effet « apprentissage » ? Pour ce faire, nous avons calculé, pour chaque lecteur, les indicateurs sur les 20 premières lésions analysées, puis les 40, puis les 60, les 80 et enfin les 100 lésions. Nous avons également pris en compte les différents clusters de lecteurs mis en évidence par la classification hiérarchique. L'évolution des paramètres selon les 5 conditions et les 2 clusters a été testée à l'aide d'un modèle linéaire mixte à 2 facteurs, avec test de l'interaction.

Toutes les analyses ont été réalisées à l'aide du logiciel SAS V9.4 (Cary, NC, USA).

IV- Résultats:

1) Population des lecteurs

Les caractéristiques de la population des 52 lecteurs sont détaillées dans le **tableau 1**.

Sexe	Féminin 24(46%)	Masculin 28(54%)			
Niveau d'Études	Internes 23(44%)	CCA 3 (6%)	PH 26(50%)		
Activité	Hospitalière 32(61%)	Mixte 4 (8%)	Privée 16 (31%)		
Niveau de spécialisation en imagerie Génito Urinaire	Activité Généraliste 28(54%)	Partiellement spécialisée 18(35%)	Spécialisée en imagerie Génito Urinaire 6(11%)		
Début IRMp	< 1an 21(40%)	> 1an 10(19%)	> 3ans 7(14%)	> 5ans 6(11%)	> 10 ans 8(16%)
Nombre d'IRMp interprété / an	< 50 24(46%)	50-100 9(17%)	100-200 11(21%)	> 200 8(16%)	
Dernière formation en Imagerie Prostatique	< 1an 31(60%)	> 1 an 15(28%)	> 3ans 4(8%)	> 5ans 2(4%)	
Participation RCP Génito Urinaire	Oui 22(42%)	Non 30(58%)			

Tableau 1: Caractéristiques de la population de lecteurs

2) Concordance lecteurs/référence pour le score PI-RADS :

La reproductibilité du PI-RADS v2.1 a été évaluée grâce aux valeurs WK et PbonREP les réponses concernant ce scoring sont détaillées dans les **tableaux 2 et 3**:

- Avec les 5 grades : WK = 0,58 (IC 0,57- 0,60) et PbonREP = 51,69%

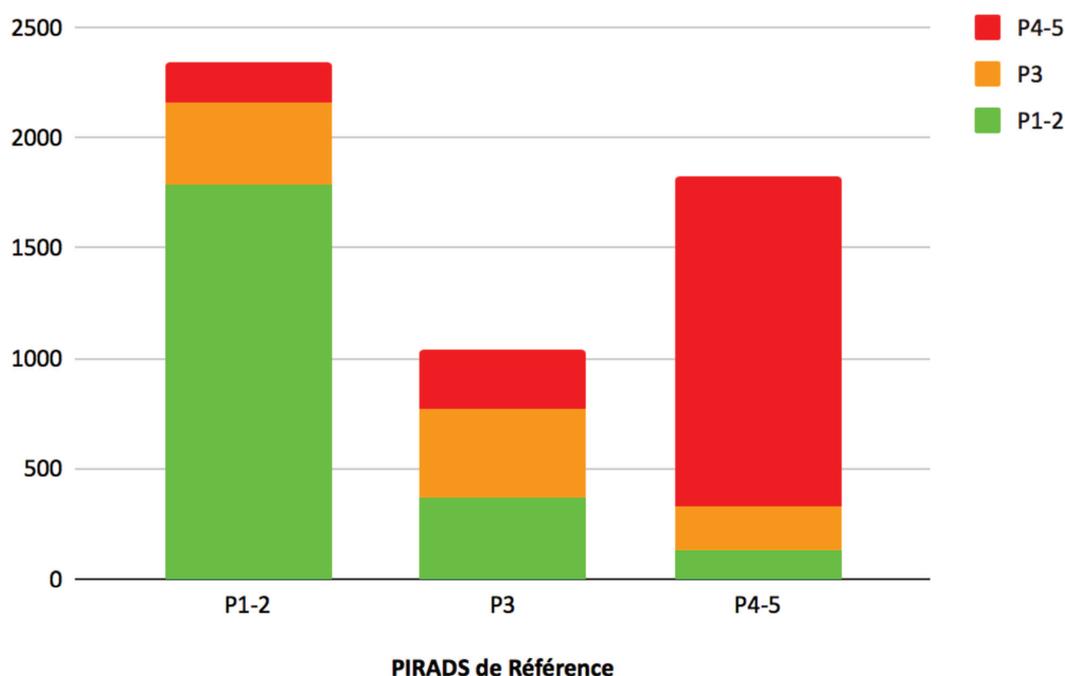
- Avec les 3 catégories : WK = 0,63 (IC 0,62-0,65) et PbonREP = 70,77%

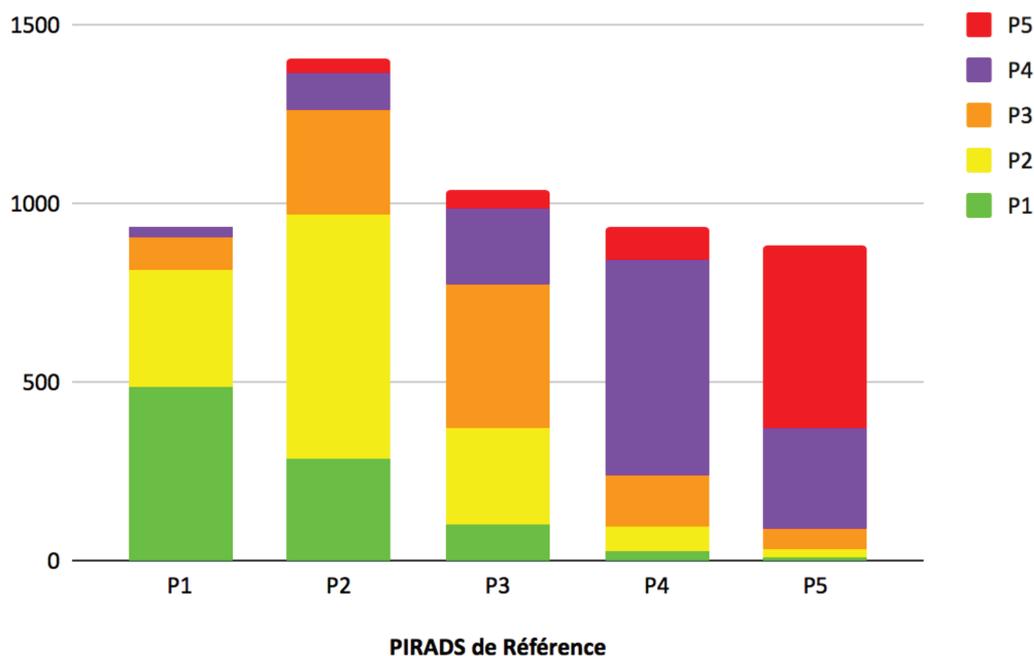
PIRADS	P1	P2	P3	P4	P5	Total
P1	485	287	98	23	9	902
P2	329	683	272	71	25	1380
P3	91	289	402	143	55	980
P4	28	104	211	606	283	1232
P5	3	41	57	93	512	706
Total	936	1404	1040	936	884	5200

PIRADS	P1-2	P3	P4-5	Total
P1-2	1784	370	128	2282
P3	380	402	198	980
P4-5	176	268	1494	1938
Total	2340	1040	1820	5200

Tableau 2 et 3: Réponses par catégories PI-RADS en 5 classes et 3 classes.

On a noté que la concordance des réponses était bonne pour les PI-RADS 1, PI-RADS 2, PI-RADS 4 et PI-RADS 5 mais plus aléatoire pour les PI-RADS 3 (tableaux 4 et 5).





Tableaux 4 et 5 : répartition des réponses en 3 ou 5 classes avec en abscisse le score PI-RADS de référence, en ordonné les scores PI-RADS des lecteurs.

3) Concordance lecteurs/référence pour la localisation lésionnelle:

Le PTopoOK des lésions était évalué à 77,46% (tableau 6).

topook	Frequency	Percent
0	1172	22,54
1	4028	77,46

Tableau 6: PTopoOK = pourcentage de concordance pour la localisation lésionnelle (en rouge)

4) Analyse en groupe de lecteurs:

En fonction de valeurs de WK, PbonREP et RMSE, on séparait la population des lecteurs en 2 catégories significativement différentes ($p < 0,001$) (Tableau 7) :

- une catégorie 1 de 16 lecteurs avec une bonne performance.
- une catégorie 2 de 36 avec une performance modérée.

CLUSTER	N Obs	Variable	N	Mean	Std Dev	Min	Lower Quartile	Median	Upper Quartile	Max
1	16	PtopoOK	16	0,82	0,05	0,73	0,79	0,82	0,86	0,89
		pbonrep	16	0,58	0,04	0,52	0,56	0,57	0,61	0,66
		RMSE	16	0,85	0,06	0,73	0,82	0,86	0,88	0,94
		wkappa	16	0,66	0,03	0,61	0,64	0,66	0,67	0,73
2	36	PtopoOK	36	0,76	0,08	0,54	0,70	0,78	0,81	0,89
		pbonrep	36	0,49	0,05	0,35	0,46	0,49	0,52	0,58
		RMSE	36	1,04	0,10	0,87	0,96	1,02	1,08	1,31
		wkappa	36	0,55	0,05	0,43	0,52	0,55	0,58	0,66

Tableau 7 : Performance des lecteurs dans le groupe 1 et le groupe 2

Il n'y avait pas de différence significative entre ces 2 groupes en fonction des caractéristiques des lecteurs citées dans le tableau 1 (notamment en fonction de leur expérience en IRM-mp).

5) Analyse en groupe des lésions cible :

Grâce aux 3 paramètres PbonREP, RMSE et PtopoOK, nous avons défini 4 groupes de lésions prostatiques (cluster) (**Tableau 8**) :

- les lésions du groupe C1 avec une bonne reproductibilité = 28 lésions.
- les lésions du groupe C2 avec une reproductibilité moyenne = 48 lésions.
- les lésions du groupe C3 avec une mauvaise reproductibilité = 13 lésions.
- les lésions du groupe C4 avec une très mauvaise reproductibilité = 13 lésions.

Cluster	N	PtopoOK	pbonrep	RMSE
C1	26	0,831	0,817	0,473
C2	48	0,867	0,434	0,989
C3	13	0,438	0,402	0,939
C4	13	0,657	0,336	1,498

Tableau 8: Groupes C1, C2, C3 et C4 en fonction de la pbonrep, PtopoOK et du RMSE.

6) Analyse en fonction de la topographie ZP/ZT:

En classant la topographie des lésions en 2 groupes simplifiés en fonction de leur localisation soit en ZP ou en ZT (= ZT+SFMA+ZC), on trouvait que les lésions de la ZP avaient un PTopoOK significativement plus élevé que dans la ZT ($p < 0.0001$).

L'analyse en fonction de la topographie ZP/ZT des 4 sous-groupes de lésions cibles montrait que le taux de lésions situées dans la ZT augmente de façon quasi linéaire du sous-groupe C1 au sous-groupe C4, (**Tableau 9**).

CLUSTER	ZP	ZT	%ZT
1	23	3	11,5
2	35	13	27,1
3	5	8	61,5
4	2	11	84,6

Tableau 9 : Nombre de lésions dans la ZP et la ZT en fonction des 4 sous-groupes de lésions (cluster).

A droite, pourcentage des lésions dans la ZT.

En outre, il existait une différence significative entre la ZP et la ZT avec une PbonREP significativement plus basse dans la ZT que dans la ZP ($p < 0.0097$) et un RMSE plus grand pour la ZT que pour la ZP ($p < 0.0012$) (**Tableau 10**).

toporef	N Obs	Variable	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
ZP	65	PTopoOK	0,23	0,79	0,90	0,94	1,00
		pbonrep	0,12	0,38	0,54	0,77	1,00
		RMSE	0,00	0,64	0,88	1,04	1,49
ZT	35	PTopoOK	0,27	0,48	0,65	0,83	0,98
		pbonrep	0,15	0,31	0,42	0,58	0,79
		RMSE	0,57	0,84	1,07	1,27	2,16

Tableau 10 : Etude de la PtopoOK, PbonREP et RMSE des lésions en fonction de leur topographie ZP ou ZT

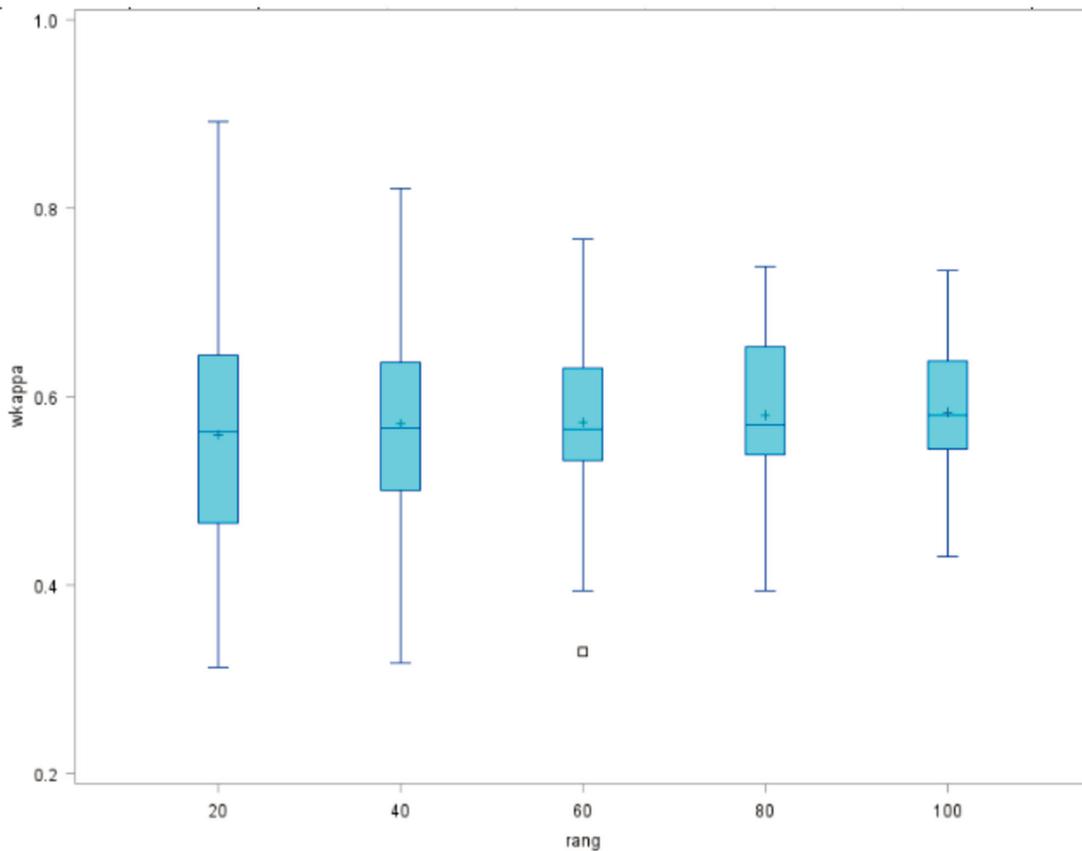
7) Analyse de l'apprentissage au fur et à mesure des cas:

Les paramètres WK, PbonREP, RMSE et PTopoOK de tous les lecteurs confondus, n'étaient pas significativement différents si on analysait les résultats tous les 20 cas.

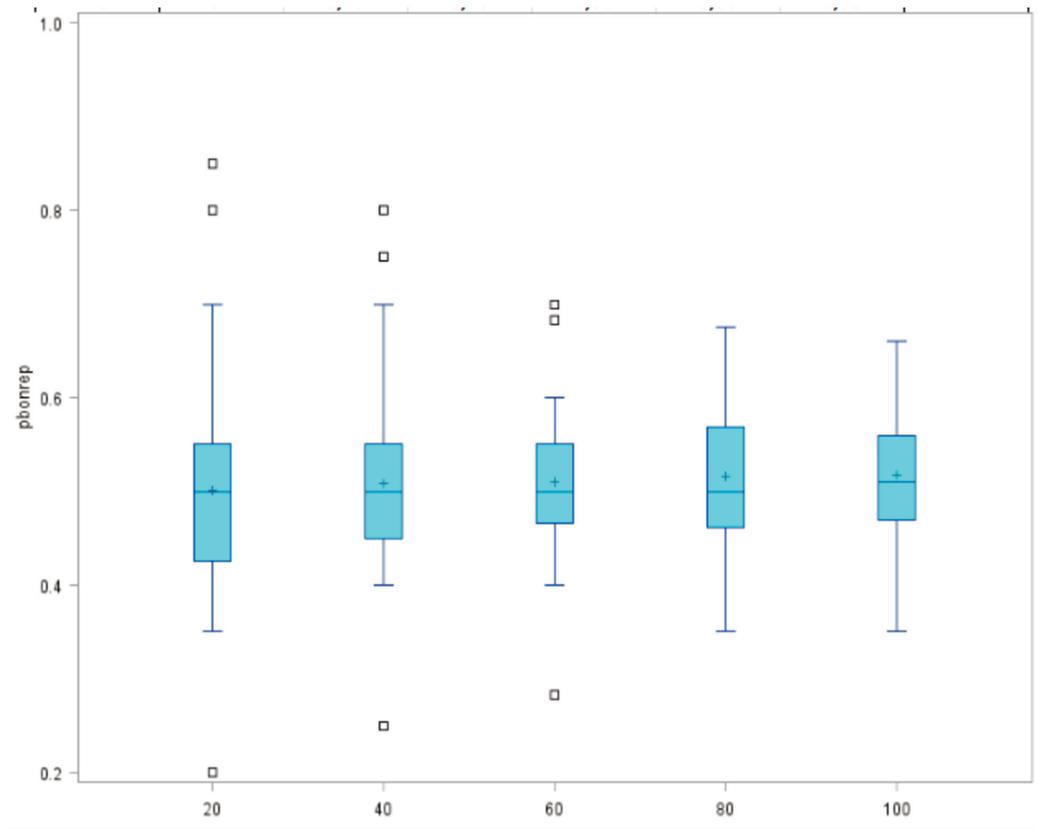
Il existait cependant une tendance générale à une moins grande dispersion des réponses à partir de 80 lésions analysées (**Figure 27**).

Il n'y avait pas non plus de différence de progression au fur et à mesure des cas, entre les deux sous-groupes de lecteurs (**Figure 28**).

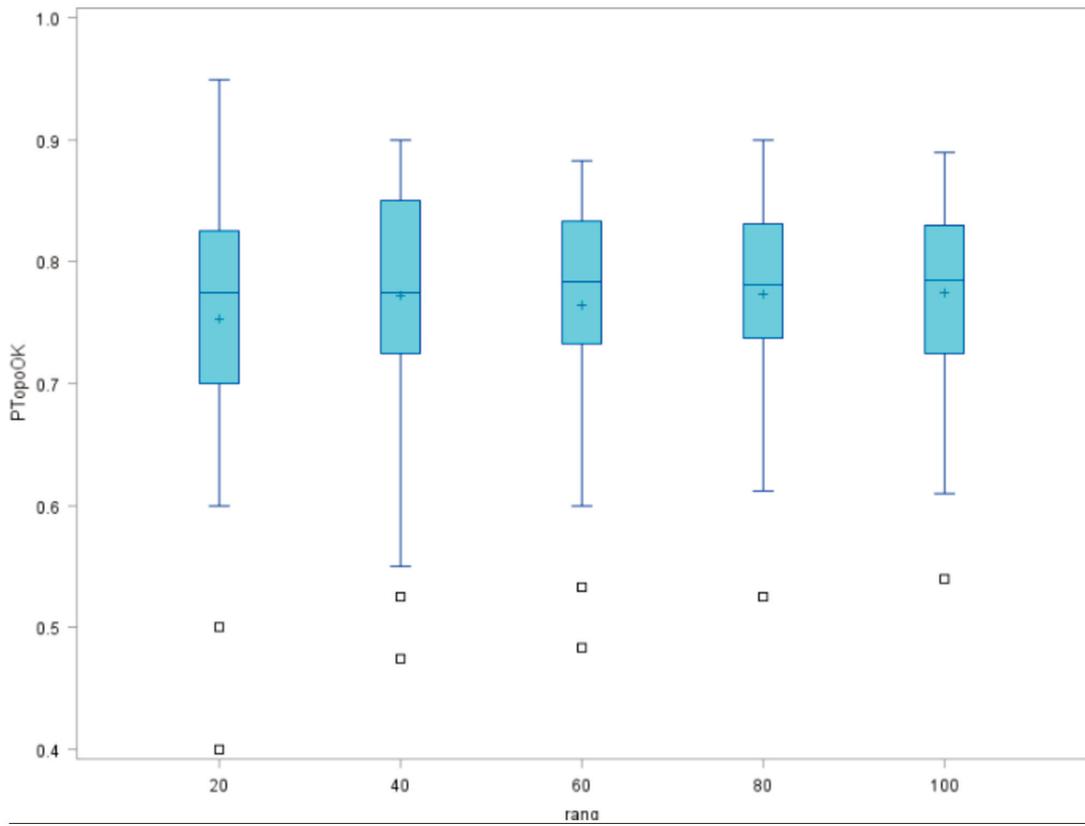
Figure 27 : Graphiques Box plot, représentant l'évolution tous les 20 cas (en abscisse) avec minimum, maximum, médiane et quartile (ordonnée) du WK (graphique 1), de la PbonREP (graphique 2), de la PTopoOK (graphique 3) et du RMSE (graphique 4).



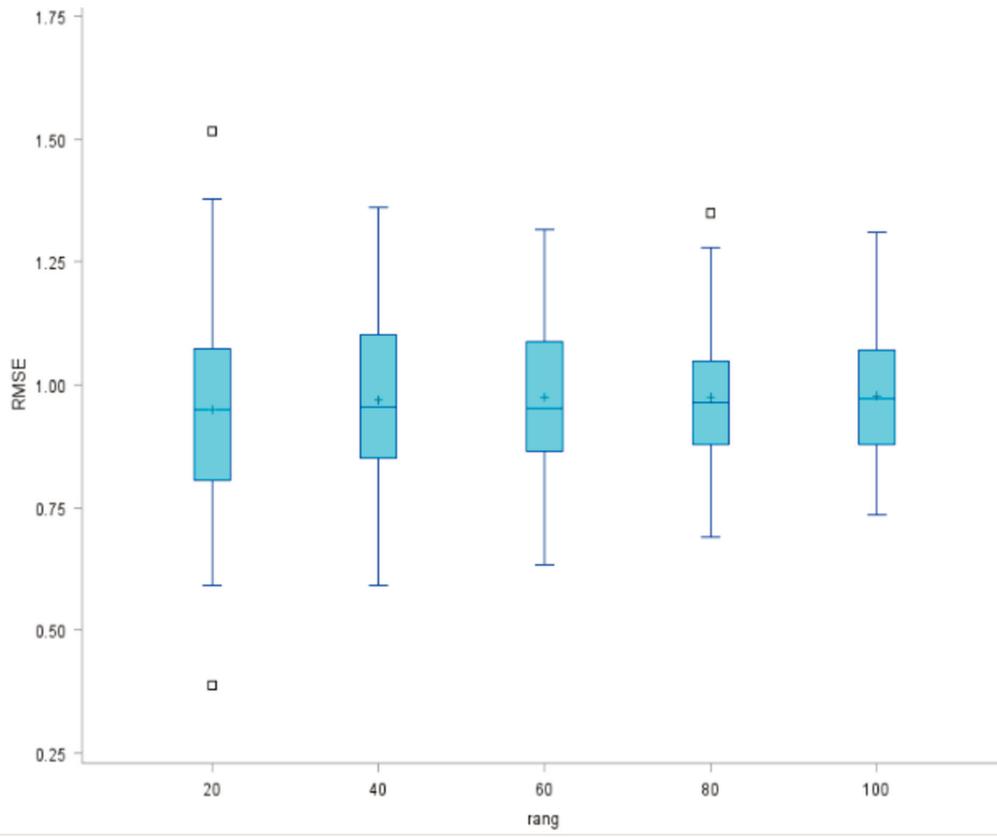
Graphique 1



Graphique 2

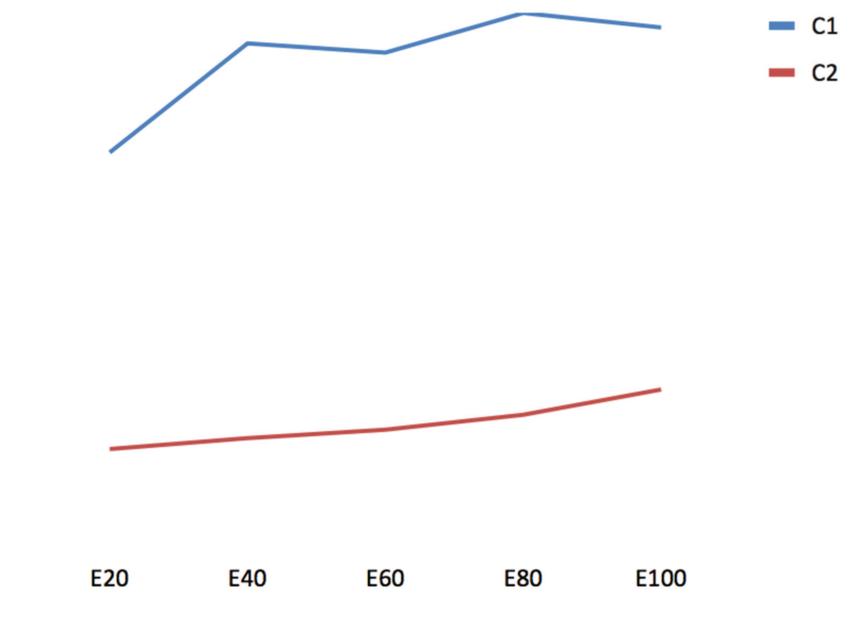


Graphique 3

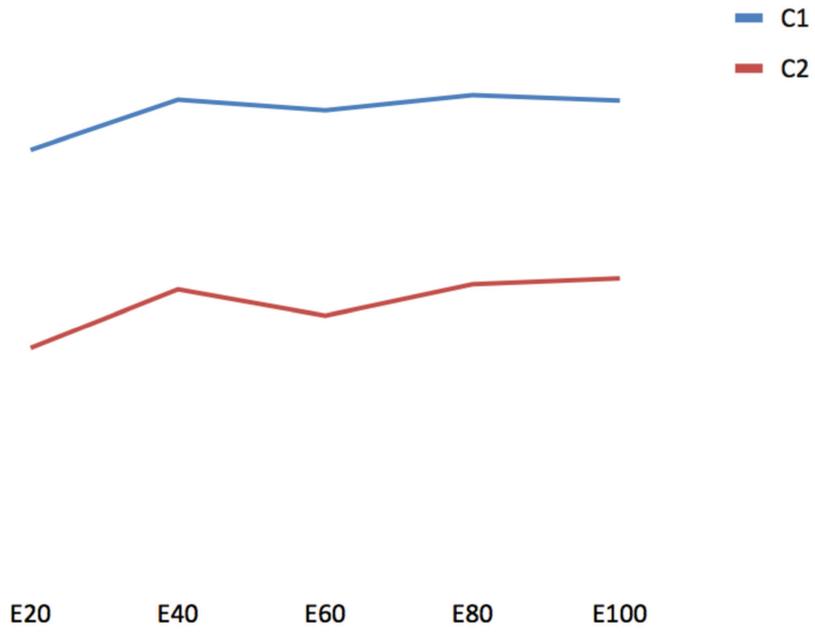


Graphique 4

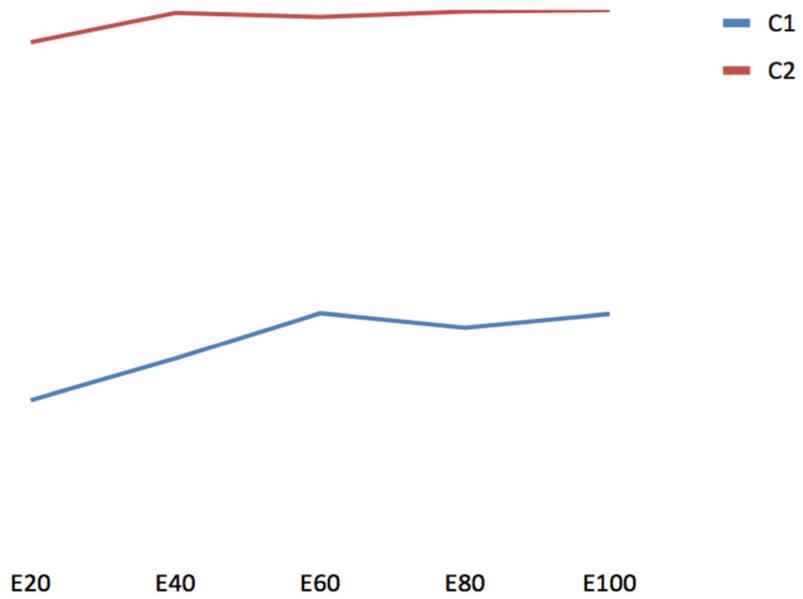
Figure 28 : Evolution par groupe de lecteur C1 et C2 tous les 20 cas (abscisse), en fonction (ordonnée) du WK (Graphique 1), de la PTopoOK (Graphique 2), du RMSE (Graphique 3) et de la PbonREP (Graphique 4) :



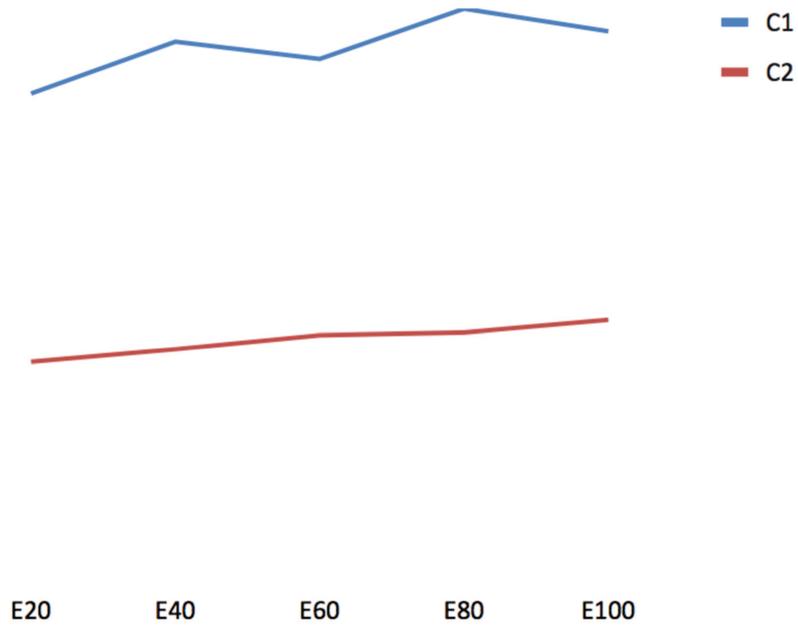
Graphique 1



Graphique 2



Graphique 3



Graphique 4

V- Discussion:

Nous avons voulu tester la reproductibilité inter observateurs de la dernière version (v2.1) du score PI-RADS. Ce paramètre a été évalué dans la littérature, avec plusieurs autres scores dont la finalité était similaire à celle du PI-RADS. Dans le domaine de l'imagerie oncologique, le « scoring » a pour objectif de standardiser et d'homogénéiser l'interprétation d'examen d'imagerie réalisés fréquemment ou systématiquement pour les patients, par des lecteurs nombreux et d'expérience variable, et nécessitant une transmission d'informations à d'autres spécialistes, pouvant modifier la prise en charge des patients en fonction du résultat de l'examen. Ont été initialement mis au point le score BI-RADS (Breast Imaging Reporting and Data System) pour la cancérologie mammaire et le score TI-RADS (Thyroid Imaging Reporting and Data System) pour la cancérologie thyroïdienne. D'autres scores standardisés et moins développés existent pour les lésions hépatiques (LI-RADS), ou encore les cancers de l'ovaire (O-RADS). Une importante variabilité inter observateurs a été publiée dans l'utilisation des scores les plus anciens:

- Pour le score BI-RADS, ont été rapportés une valeur de k à 0,45 en mammographie en 2002 (34), une valeur de k à 0,3 en échographie en 2008 (35) et un coefficient alpha de Krippendorff à 0,38 en IRM pour des lésions autres que des masses en 2015 (36).
- Pour le score TI-RADS, une valeur de k à 0,636 a été rapportée dans une récente étude (37).

Depuis la création du PI-RADS en 2012, la reproductibilité de ce score a été testée à maintes reprises dans différentes études (38, 39, 40, 41), qui ont inspiré notre méthodologie. a) Nous avons réalisé une étude multicentrique alors que de nombreux articles publiés sur ce thème, étaient monocentriques (39, 40, 41). b) Notre effectif de lecteurs ($n=52$) était beaucoup plus conséquent que ceux d'études précédentes (6 lecteurs au maximum) (38). c) Nous avons décidé d'inclure un panel de radiologues seniors et juniors plus ou moins expérimentés dans la pratique de l'IRM-mp, à la différence d'études antérieures qui faisaient uniquement appel à des radiologues expérimentés (38, 39, 40, 41).

Notre choix avait pour but d'évaluer la reproductibilité du score PI-RADS v2.1, en réduisant autant que possible, le poids de l'expérience du lecteur qui était jugé trop important avec « l'échelle subjective de Likert » utilisée avant le score PI-RADS.

Nous rapportons une reproductibilité inter observateurs modérée avec une valeur du k à

0,58 mais ce résultat est comparable à ceux rapportés antérieurement :

- Renard-Penna et al (39) ont publié un k à 0,73 et Vaché et al (40) entre 0,38 et 0,44 dans des études monocentriques évaluant le score PI-RADS v1.
- Muller et al ont rapporté un k à 0,46 dans une étude monocentrique évaluant le score PI-RADS v2 (41).
- Rosenkrantz et al (38) ont publié une étude multicentrique évaluant le score PI-RADS v2, et ont rapporté un k global à 0,55 pour des lésions de grade PI-RADS ≥ 4 .

Dans notre étude, nous avons aussi mis en évidence une concordance tout à fait acceptable entre les lecteurs et l'expert concernant la localisation des lésions au sein de la cartographie prostatique en 41 secteurs (PTopoOK = 0,77). Cette notion n'avait jusqu'à présent pas été rapportée dans la littérature en utilisant la cartographie prostatique la plus récente et en faisant appel à un effectif important de lecteurs.

En utilisant les paramètres nous ayant permis d'évaluer le score PI-RADS (PbonREP et RMSE) et la localisation lésionnelle (PTopoOK), nous avons séparé les lésions prostatiques en 4 sous-groupes de reproductibilité décroissante. Dans les sous-groupes de lésions à mauvaise voire à très mauvaise reproductibilité (groupes 3 et 4), les lésions de la ZT étaient prédominantes et les reproductibilités inter observateurs dans l'analyse des lésions de la ZT et celle de la ZP étaient statistiquement différentes.

Une différence de reproductibilité inter observateurs avait déjà été signalée par Rosenkrantz et al dans la détermination du grade PI-RADS des lésions de la ZT et celles de la ZP : k à 0,51 dans la ZT et k à 0,59 dans la ZP (38).

Cette variabilité inter observateurs est particulièrement nette pour les lésions de la ZT de grades PI-RADS faibles à intermédiaires (grades 1, 2, 3). Dans la ZT, les critères du PI-RADS se basent sur des notions subjectives de forme et d'encapsulation d'un nodule. Le choix pour cette zone, de séquences de référence en pondération T2, offre aux lecteurs une excellente résolution spatiale et une excellente résolution en contraste mais cela donne lieu à des interprétations très variables d'un lecteur à l'autre. En effet il semble fastidieux de coter dans les cas d'HBP l'intégralité des nodules de la ZT, de plus la notion d'encapsulation d'un nodule partielle ou complète, surajouté à la notion de nodule dans le nodule restent des paramètres subjectifs. Une influence plus importante des séquences de diffusion diminuerait certainement cette variabilité inter observateurs dans la détermination du grade PI-RADS des lésions de la ZT (42). C'est déjà ce que l'on constate dans la ZP où les critères de référence sont basés sur les séquences de

diffusion. En effet, le côté robuste et l'analyse relativement binaire des séquences de diffusion laissent moins de place à l'interprétation et à la subjectivité des lecteurs.

La variabilité inter lecteurs est aussi plus nette dans la ZT que dans la ZP pour la localisation des lésions. Le score PI-RADS propose une cartographie prostatique qui a progressivement évolué dans le temps, passant de 29 secteurs dans la version v1 à 39 secteurs dans la version v2 puis 41 dans la v2.1. Cette évolution a surtout concerné les parties antérieures (SFMA) et postérieures (ZP et ZC) de la prostate et peu la partie médiane (ZT).

La sectorisation de la prostate est donc plus simple dans la ZT que dans les autres zones. Dans la cartographie en 41 secteurs, la ZT ne représente que 30% des secteurs alors que chez le sujet âgé, son volume représente très souvent plus de 50% du volume prostatique global.

Nous pensons que le lecteur a souvent du mal à trouver la frontière entre les 3 étages prostatiques (base, milieu, apex) et à séparer les parties antérieures et les parties postérieures dans la ZT dont l'architecture et l'anatomie sont naturellement perturbées par l'hypertrophie bénigne de prostate.

En outre, le lobe médian n'apparaît pas dans la cartographie prostatique. Tous ces éléments expliquent probablement les fréquentes erreurs de localisation lésionnelle dans la ZT et la plus faible reproductibilité inter observateurs constatée dans notre étude. Ce constat plaide en faveur d'une nouvelle évolution de la cartographie prostatique comme cela a déjà pu être suggéré dans la littérature (43).

Notre étude n'a pas mis en évidence de lien entre l'expérience des lecteurs et la reproductibilité inter observateurs. Nous avons statistiquement séparé la population des 52 lecteurs en 2 groupes : un ayant une bonne concordance avec la référence (n = 16) et l'autre ayant une moins bonne concordance (n = 36). La répartition des radiologues expérimentés et non expérimentés n'était pas différente dans ces deux groupes. Cette notion avait déjà été évoqué dans d'autres études de faible effectif (41, 44) mais n'est pas étonnante puisque l'un des objectifs du PI-RADS est d'homogénéiser l'interprétation des IRM-mp quelle que soit l'expérience des lecteurs.

La plateforme PST semble un outil facile d'accès et d'utilisation pour les lecteurs. Le lecteur peut ainsi s'entraîner à la lecture d'IRM-mp et évaluer ses performances au décours. L'une des perspectives de notre travail serait de proposer le PST de façon plus généralisée aux radiologues souhaitant se former en IRM-mp et de donner à l'issue une

certification avec le coefficient kappa du lecteur.

Notre étude présentait plusieurs limites liées au recrutement des lecteurs, au choix des cas et à la plateforme utilisée :

- a) La participation à cette étude s'est faite sur la base du volontariat ce qui nous a conduit à recruter des individus motivés par l'IRM-mp, l'apprentissage du PI-RADS et/ou l'utilisation d'un nouvel outil d'auto-évaluation « en ligne ». Cette population ne représentait pas forcément celle de tous les radiologues.
- b) Nous avons sélectionné des cas provenant d'un seul centre d'imagerie et d'un seul appareil 1,5 T. Ils ne pouvaient donc pas illustrer la diversité des machines 1,5 T et 3 T utilisées au quotidien pour réaliser des IRM-mp.
- c) Nous avons choisi les cas de façon arbitraire en essayant de répartir au mieux le type et la localisation des lésions. Cet élément a pu introduire des biais de sélections.
- d) Les nombres de cas ($n = 50$) et de lésions ($n = 100$) pouvaient sembler élevés aux lecteurs et le conduire à en bâcler certains avec pour conséquence, une dégradation de la reproductibilité inter observateurs. Malgré cet effectif important, nous n'avons pas réussi à mettre en évidence une progression des lecteurs tout au long de leurs sessions de travail, et ce quel que soit leur sous-groupe.
- e) Enfin, en prédéterminant deux zones à analyser, nous n'avons pu tester chez le lecteur, sa capacité à dépister des lésions suspectes.
- f) Les lecteurs pouvaient accéder à la plateforme PST via un simple lien internet et utiliser un PC, une tablette ou un portable pour lire les 50 cas et étudier les 100 lésions. Ce mode de fonctionnement ne reflétait en rien, la réalité d'une interprétation sur une console PACS ou sur une station de travail professionnelle où le lecteur peut avoir accès aux données et outils DICOM et disposer d'écrans de taille et de résolution optimisées.
- g) Les images disponibles étaient limitées en nombre et extraites à la norme JPEG (comprimées). Seul le plan axial T2 offrait la possibilité d'un défilement, l'image

sagittale étant unique et ne servant qu'à déterminer le niveau de coupe. Les images DCE étaient petites et l'outil de mesure était peu fonctionnel. La sélection de deux lésions par cas a pu engendrer une inversion des lésions lors du remplissage des réponses et ainsi induire des erreurs. Tous ces éléments pourront être corrigés ou améliorés dans la future version.

VI- Conclusion:

Notre étude montre une reproductibilité inter observateurs du PI-RADS v2.1 comparable à celle du PI-RADS v2 dans une large population de radiologues séniors et juniors plus ou moins expérimentés dans la lecture des IRM-mp. L'expérience des lecteurs ne semble pas avoir une influence majeure, ce qui est en accord avec les objectifs du score PI-RADS. Malgré la révision récente de certains critères morphologiques du PI-RADS, il semble persister des difficultés dans le classement des lésions de la ZT que ce soit pour l'attribution du grade et pour la localisation lésionnelle dans la cartographie en 41 secteurs. Des études complémentaires semblent nécessaires pour valider ces conclusions et proposer d'éventuels axes d'amélioration. La plateforme PST semble être un outil efficace pour évaluer les performances du score PI-RADS en proposant à un effectif important de lecteurs, l'analyse d'un nombre élevé de lésions.

ANNEXES

Annexe 1 : Questionnaire préliminaire d'ouverture du PST

Votre formation	
Quelle est votre spécialité ?	Radiologue
Quel est votre niveau de formation ?	Interne
Qu'est ce qui décrit le mieux votre pratique actuelle ?	Public
Peut on considérer que vous êtes spécialisé(e) en imagerie génito-urinaire ?	Oui, partiellement
Votre pratique de l'imagerie prostatique	
Depuis combien d'années faites vous régulièrement de l'imagerie de la prostate ?	< 1 an
Combien d'IRM de prostate avez-vous interprété dans les 12 derniers mois ?	50-100
Quand s'est déroulée votre dernière participation à une formation à l'IRM de prostate (en ligne ou présentielle) ?	< 1 an
Participez-vous régulièrement à des RCPs en uro-oncologie ?	Non
Réalisez-vous personnellement et régulièrement des biopsies de la prostate échoguidées ?	Non
Votre connaissance de PI-RADS	
Comment qualifieriez vous votre connaissance de la version 2.1 de la classification PI-RADS ?	Bon(ne)
Quelle est la séquence dominante pour classer une lésion de la zone périphérique ?	T2
Quelle est la séquence dominante pour classer une lésion de la zone de transition ?	Diffusion
Quelle est la séquence dominante pour classer une lésion de le stroma fibromusculaire antérieur ?	T2
Quelle est la séquence dite "secondaire" qui peut modifier le score d'une lésion de la zone périphérique (pré)classée 3 dans la séquence dominante ?	Perfusion
Quelle est la séquence dite "secondaire" qui peut modifier le score d'une lésion de la zone de transition (pré)classée 3 dans la séquence dominante ?	Diffusion
D'après PI-RADS 2.1, un nodule de la zone de transition parfaitement cerclé mais restreignant la diffusion doit être classé...	PI-RADS 1
D'après PI-RADS 2.1, une lésion de la zone périphérique de 18 mm, localisée, restreignant faiblement la diffusion et hypervascularisée doit être classée...	PI-RADS 4
Comment considérez vous votre reproductibilité inter-observateur pour déterminer la TOPOGRAPHIE d'une lésion selon le schéma à 41 secteurs de PI-RADS 2.1, par rapport à un(e) expert(e) ?	Bon(ne)
Comment qualifieriez vous votre reproductibilité inter-observateur avec un expert, pour la classification des lésions selon PI-RADS 2.1 ?	Bon(ne)
Votre participation à l'étude	
Confirmez-vous que vous avez téléchargé le document " Formulaire de consentement pour participation à la recherche ", que vous l'avez lu complètement, et acceptez toutes les conditions de stockage, traitement et utilisation scientifique des données que vous allez produire pour cette étude ?	Oui

VII- Bibliographie :

1. Cowppli-Bony A, Colonna M, Ligier K, Jooste V, Defossez G, Monereau A, le réseau Francim. Épidémiologie descriptive des cancers en France métropolitaine : incidence, survie et prévalence. Bull Cancer 2019 ; 106 : 617-34.
2. Fournier G, Valeri A, Mangin P, Cussenot O. Cancer de la prostate. Épidémiologie. Facteurs de Risques. Anatomopathologie. Ann Urol 2004 ; 38 : 187-206.
3. Rozet F, Hennequin C, Beauval JB, et al. Recommandations nationales 2016 sur le cancer de prostate. Prog Urol 2016 ; 27 : 95-143.
4. McNeal JE, Redwine EA, Freiha FS, Stamey TA. Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction spread. Am J Surg Pathol 1988 ; 12 : 897-906.
5. Nevoux P, Ouzzane A, Ahmed HU, Emberton M, Montironi R, Presti JC Jr, Villers A. Quantitative tissue analyses of prostate cancer foci in an unselected cystoprostatectomy series. BJU int 2012 ; 110 : 517-23.
6. Cohen RJ, Shannon BA, Phillips M et al. Central zone carcinoma of the prostate gland: a distinct tumor type with poor prognostic features. J Urol 2008 ; 179 : 1762-7.
7. Olson MC, Atwell TD, Mynderse LA, King BF, Welch T, Goenka AH. CT-guided transgluteal biopsy for systematic sampling of the prostate in patients without rectal access : a 13-year single-center experience. Eur Radiol 2017; 27 : 3326-32.
8. Stamey TA, McNeal JE, Yemoto CM, Signal BM, Johnstone IM. Biological determinants of cancer progression in men with prostate cancer. JAMA 1999 ; 281 : 1395-400.
9. Epstein JL, Egevad L, Amin MB, Delahunt B, Srigley JR, Humphrey PA, Grading committee. The 2014 International Society of Urological Pathology (ISUP) Consensus conference on Gleason grading of prostatic carcinoma : definition of

- grading patterns and proposal for a new grading system. *Am J Surg Pathol* 2016 ; 40 : 244-52.
10. Tabayoyong W, Abouassaly R. Prostate Cancer Screening and the Associated Controversy. *Surg Clin North Am* 2015 ; 95 : 1023-39.
 11. Dépistage du cancer de la prostate. Ameli.fr. Le dépistage du cancer de la prostate : s'informer avant de décider. Avis HAS ; mars 2016.
 12. Dépistage du cancer de la prostate. Ameli.fr. La première prescription du PSA chez l'homme asymptomatique. Avis HAS ; mars 2016.
 13. Rozet F, Hennequin C, Beauval, et al. Cancer de la prostate. Diagnostic, traitement, recommandations. *Prog Urol* 2018 ; 28 : 81-132.
 14. Salomon L, Bastide C, Beuzeboc P, et al. Recommandations en onco-urologie 2013 du CCAFU : Cancer de la prostate. *Prog Urol* 2013 ; 23 : 69-101.
 15. Salomon L, Azria D, Bastide C, et al. Recommandations en Onco-Urologie 2010 : Cancer de la prostate. *Prog Urol* 2010 ; 20 : 17-51.
 16. Catalona WJ, Smith DS, Ratliff TL, Dodds KM, Coplen DE, Yuan JJ, Petros JA, Andriole GL. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *N Engl J Med* 1991 ; 324 : 1156-61.
 17. Eyraud R, Senechal C, Santy A, et al. Bénéfice d'inclure la densité du PSA<0,2 dans la surveillance active des patients Afro-caribéens. *Prog Urol* 2014 ; 24 : 871-2.
 18. Renard-Penna R, Rouviere O, Puech P, et al. Current practice and access to prostate MR imaging in France. *Diagn interv imaging* 2016 ; 97 : 1125-29.
 19. D'Amico AV, Whittington R, Malkowicz SB, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy or interstitial radiation therapy for clinically localized prostate cancer. *JAMA* 1998 ; 280 : 969-74.
 20. Kasivisvanathan V, Rannikko AS, Borghi M, et al. MRI-targeted or standard biopsy for prostate-cancer diagnosis. *N Engl J Med* 2018 ; 378 : 1767-77.

21. Uriburu-Pizarro F, Kasivisvanathan V, Puech P, Villers A. Pre-biopsy MRI as an adjunct for cancer detection in men with elevated PSA and no previous biopsy. *Transl Androl Urol* 2017 ; 6 : 387-94.
22. Fütterer 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 : 1045-53.
23. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. *Eur Radiol* 2012 ; 22 : 746-57.
24. Wilt TJ, Brawer MK, Jones KM, et al. Radical prostatectomy versus observation for localized prostate cancer. *N Engl J Med* 2012 ; 367 : 203-13.
25. Donovan JL, Hamdy FC, Lane JA, et al. Patient-reported outcomes after monitoring, surgery, or radiotherapy for prostate cancer. *N Engl J Med* 2016 ; 375 : 1425-37.
26. Hamdy FC, Donovan JL, Lane JA, et al. 10-year outcomes after monitoring, surgery, or radiotherapy for localized prostate cancer. *N Engl J Med* 2016 ; 375 : 1415-24.
27. Heidenreich A, Bastian PJ, Bellmunt J, et al. EAU guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent-update 2013. *Eur Urol* 2014 ; 65 : 124-37.
28. Ouazzane A, Renard-Penna R, Marliere F, et al. Magnetic Resonance Imaging targeted biopsy improves selection of patients considered for active surveillance for clinically low risk prostate cancer based on systematic biopsies. *J Urol* 2015 ; 194 : 350-6.
29. Turkbey B, Rosenkrantz AB, Haider MA, et al. PI-RADS Version 2.1 : 2019 Update of Prostate Imaging Reporting and Data System Version 2. *Eur Urol* 2019;76 : 340-51.
30. Likert R. A technique for the measurement of attitudes. *Arch Psychol.* 1932;140:1-55.

31. 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.
32. Bridault JP. Thèse de doctorat en médecine 2018 : Prostate Score Trainer : une plateforme en ligne d'entraînement et d'auto-évaluation à l'interprétation standardisée de l'IRM prostatique. Université de Lille 2. <http://pepite.univ-lille2.fr/>
33. Cohen J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 1960 : 20 ; 37-46.
34. Berg WA, D'Orsi CJ, Jackson VP, Basset LW, Beam CA, Lewis RS, Crewson PE. Does training in the Breast Imaging Reporting and Data System (BI-RADS) improve biopsy recommendations or feature analysis agreement with experienced breast imagers at mammography ? Radiology 2002 ; 224 : 871-80.
35. Lee HJ, Kim EK, Kim MJ, Youk JH, Lee JY, Kang DR, Oh KK. Observer variability of Breast Imaging Reporting and Data System (BI-RADS) for breast ultrasound. Eur J Radiol 2008 ; 65 : 293-8.
36. El Khoury M, Lalonde L, David J, Labelle M, Mesurolle B, Trop I. Breast imaging reporting and data system (BI-RADS) lexicon for breast MRI : interobserver variability in the description and assignment of BI-RADS category. Eur J Radiol 2015 ; 84 : 71-6.
37. Basha MAA, Alnaggar AA, Refaat R, et al. The validity and reproducibility of the thyroid imaging reporting and data system (TI-RADS) in categorization of thyroid nodules: Multicentre prospective study. Eur J Radiol 2019 ; 117 : 184-92.
38. Rosenkrantz AB, Ginocchio LA, Cornfeld D, et al. Interobserver Reproducibility of the PI-RADS Version 2 Lexicon : A Multicenter Study of Six Experienced Prostate Radiologists. Radiology 2016 ; 280 : 793-804.
39. Renard-Penna R, Mozer P, Cornud F, Barry- Delongchamps N, Bruguière E, Portalez D, Malavaud B. Prostate imaging reporting and data system and Likert scoring system : multiparametric MR imaging validation study to screen patients

for initial biopsy. *Radiology* 2015 ; 275 : 458-68.

40. Vaché T, Bratan F, Mège-Lechevallier F, Roche S, Rabilloud M, Rouvière O. Characterization of prostate lesions as benign or malignant at multiparametric MR imaging: comparison of three scoring systems in patients treated with radical prostatectomy. *Radiology* 2014 ; 272 : 446-55.
41. Muller BG, Shih JH, Sankineni S, et al. Prostate cancer: interobserver agreement and accuracy with the revised prostate imaging reporting and data system at multiparametric MR imaging. *Radiology* 2015 ; 277 : 741-50.
42. Rosenkrantz AB, Kim S, Campbell N, Gaing B, Deng FM, Taneja SS. Transition zone prostate cancer : Revisiting the role of multiparametric MRI at 3 T. *AJR Am J Roentgenol* 2015 ; 204 : 266-72.
43. Tavolaro S, Mozer P, Roupret M, et al. Transition zone and anterior stromal prostate cancers : Evaluation of discriminant location criteria using multi parametric fusion guided biopsy. *Diagn Interv Imaging* 2018 ; 99 : 403-11.
44. Greer MD, Shih JH, Lay. N, Choyke PL, Turkbey. B, et al. Interreader Variability of Prostate Imaging Reporting and Data System Version 2 in Detecting and Assessing Prostate Cancer Lesions at Prostate MRI. *AJR Am J Roentgenol* ;2019;212: 1197-1205.

Auteur : Nom : BRUNEREAU Prénom : Julie
Date de soutenance : 27/04/2020
Titre de la thèse : Évaluation de la reproductibilité inter-observateurs du score de suspicion de cancer de prostate PI-RADS version 2.1
Thèse - Médecine : Lille 2020.
Cadre de classement : Radiodiagnostic et Imagerie médicale.
Mots-clés : IRM, prostate, cancer, reproductibilité, PI-RADS, évaluation formative, pratiques professionnelles.

Résumé :

Objectifs : évaluer la reproductibilité inter observateurs du score de suspicion de cancer de prostate PI-RADS v2.1 au sein d'une population de 52 radiologues grâce à la lecture de 50 IRM de prostate (IRM-mp) sur la plateforme internet « Prostate Score Trainer » (PST).

Matériel et Méthodes : 52 lecteurs ont été inclus dans cette étude entre le 1^{er} novembre 2019 et le 24 janvier 2020. Il s'agissait de radiologues séniors et juniors provenant de multiples centres d'imagerie et ayant une expérience variable dans la lecture des IRM-mp. Sur la plateforme PST, les lecteurs ont analysé 100 lésions prostatiques provenant de 50 examens consécutifs réalisés sur un aimant 1,5 T entre le 5 janvier et le 12 février 2016. Pour chacune des 100 lésions, le grade PI-RADS (5200 réponses) et la localisation lésionnelle (5200 réponses) ont été déterminés. Les 10400 réponses ont été comparées à celles d'un expert considéré comme référence. Pour étudier la reproductibilité inter observateurs, nous avons calculé le coefficient k de Cohen (weighted kappa = wk), le taux de concordance avec la référence (P_{bonREP} pour le PI-RADS et P_{topoOK} pour la localisation) et l'écart normalisé à la référence ($RMSE$). Le score PI-RADS a été étudié en 5 et 3 classes. Les lésions ont été étudiées par zone. Les lecteurs et les lésions ont été séparés en sous-catégories (clusters) en fonction de leurs niveaux de reproductibilité.

Résultats: Pour le score PI-RADS en 5 classes / 3 classes, les valeurs des $wk - P_{bonREP}$ étaient respectivement : 0,58/0,63 - 0,51/0,70. Pour la localisation lésionnelle, la valeur P_{topoOK} était 0,77. Ces paramètres étaient significativement différents dans la zone de transition (ZT) et dans la zone périphérique (ZP). Grâce aux paramètres de reproductibilité et aux valeurs de $RMSE$, les lecteurs et les lésions ont été respectivement séparés en 2 et 4 clusters. Les lésions de la ZT étaient significativement sur-représentées dans les clusters à faible et très faible concordances. Il n'y avait pas de lien significatif entre l'expérience des lecteurs et leur répartition dans les 2 clusters.

Conclusion: notre étude montre une reproductibilité inter observateurs du PI-RADS v2.1 comparable à celle du PI-RADS v2. L'expérience des lecteurs ne semble pas avoir une influence majeure. Il semble persister des difficultés dans l'attribution du grade PI-RADS et la localisation lésionnelle pour les lésions de la ZT. La plateforme PST semble être un outil efficace pour évaluer les performances du score PI-RADS en proposant à un effectif important de lecteurs, une analyse d'un nombre élevé de lésions.

Composition du jury :

Président : Pr Arnaud VILLERS
Asseseurs : Pr Xavier LEROY, Dr Benoit RENARD, Mr Patrick DEVOS
Directeur de thèse : Pr Philippe PUECH