



UNIVERSITÉ DE LILLE
FACULTÉ DE MÉDECINE HENRI WAREMBOURG
Année : 2020

THÈSE POUR LE DIPLÔME D'ÉTAT
DE DOCTEUR EN MÉDECINE

Rôle de l'expérience et de la double lecture sur les performances de l'IRM biparamétrique dans le dépistage du cancer de la prostate cliniquement significatif chez les patients naïfs de biopsie.

Présentée et soutenue publiquement le 23 juin 2020 à 16h00
Au Pôle Formation

Par Louis COTTEN

JURY

Président :

Monsieur le Professeur Arnauld VILLERS

Assesseurs :

Monsieur le Professeur Xavier LEROY

Monsieur le Docteur Benoit RENARD

Monsieur Patrick DEVOS

Directeur de Thèse :

Monsieur le Professeur Philippe PUECH

Travail de la faculté de Médecine 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

IRM-mp: IRM multiparamétrique

IRM-bp: IRM bi paramétrique

PCa: cancer de prostate

CsPCa: cancer de prostate cliniquement significatif

DCE: imagerie de perfusion

DWI: séquences de diffusion

PBP: ponction biopsie prostatique

Se: sensibilité

Sp: spécificité

VPP: valeur prédictive positive

VPN: valeur prédictive négative

1.5T: 1.5 TESLA

3T: 3 TESLA

TABLE DES MATIÈRES

Avertissement	2
Remerciements	3
Président du jury	3
Membre du jury	4
Membre du jury	5
Membre du jury	6
Directeur de thèse	7
Liste des abréviations	8
Table des matières	9
Résumé	13
I. Introduction:	13
II. Matériel et Méthodes:	13
III. Résultats:	14
IV. Conclusion:	14
INTRODUCTION	14
I. Cancer de prostate	15
A. Epidémiologie	15

B.	Rappels anatomiques et histologiques	16
1.	Anatomie prostatique	16
2.	Cancer de prostate	17
3.	Score de Gleason	19
4.	Score ISUP	21
C.	Débat sur dépistage	21
D.	Diagnostic du cancer de prostate	24
1.	Diagnostic clinique	24
2.	Diagnostic biologique	25
3.	Diagnostic histologique	25
4.	Complications	28
II.	IRM de prostate	28
A.	IRM-mp	28
1.	Préparation	29
2.	Protocole	29
3.	Indications	30
B.	IRM bi paramétrique de prostate	33
III.	Standard PI-RADS2.1	34
	RATIONNEL DE L'ÉTUDE	38
	Matériels et méthodes	41

I. Population de patients	41
II. Protocole d'imagerie	43
III. Protocole de biopsies	44
IV. Population de lecteurs	44
V. Affectation des cas aux lecteurs	45
VI. Préparation des données	45
VII. Protocole de lecture des dossiers	46
VIII. Gold standard	48
IX. Analyse statistique des résultats	49
Résultats	51
I. Population de patients	51
II. Temps de lecture par dossiers	51
III. Taux de concordance entre les lecteurs	53
IV. Analyse de la simple lecture selon le type de lecteur	56
V. Analyse détaillée de la double lecture selon le binôme - hypothèse du worst case.	58
VI. Analyse détaillée de la double lecture selon le binôme - hypothèse du best case	61
Discussion	63
I. Design de l'étude	63
II. Choix des lecteurs et taux de bonnes réponses	64

III.	Concordance entre les lecteurs	65
IV.	Durée d'interprétation des dossiers	66
V.	Simple lecture	67
VI.	Choix du scénario "worst case" ou " <i>best case</i> " pour la double lecture	67
VII.	Performance de la double lecture	69
VIII.	Limites	70
IX.	Implications et perspectives	73
	Conclusion	75
	Bibliographie	76

RÉSUMÉ

I. Introduction:

Le cancer de la prostate est le cancer le plus fréquent chez l'homme de plus de 50 ans. Des outils sont nécessaires pour améliorer le diagnostic du cancer de prostate cliniquement significatif et éviter ainsi le risque de surtraitement. L'IRM multiparamétrique de prostate est un examen dont l'intérêt n'est plus à démontrer dans cette indication. Les résultats restent cependant limités dans certains cas avec un recours encore large aux biopsies prostatiques. Notre étude cherche à déterminer si l'expérience des lecteurs et la double lecture en aveugle des IRMs avec un protocole biparamétrique influent sur les performances de l'examen chez les patients naïfs de biopsie.

II. Matériel et Méthodes:

Nous avons inclus 55 patients consécutifs ayant consulté pour une IRM prostatique suivie d'une première série de biopsies. Nous avons sélectionné 4 lecteurs "seniors" et 4 lecteurs "juniors" qui ont tous classé ces IRMs comme suspecte ou non de cancer significatif. Nous avons comparé les réponses des lecteurs au résultat histologique (gold standard). Nous avons ensuite créé des binômes de lecteurs dont les données d'interprétation ont été comparées au gold standard, en utilisant l'hypothèse du "best" et du "worst" case. Nous avons calculé les performances en

termes de sensibilité, spécificité, VPP et VPN de chaque lecteur en simple lecture, puis de chaque binôme en double lecture.

III. Résultats:

Le coefficient alpha de Krippendorff était de 0.53 (0.36 ; 0.69) entre les différents lecteurs.

La médiane des sensibilités en lecture simple était de 0.92, la spécificité de 0.70, la VPP de 0.72 et la VPN de 0.91.

En prenant en compte l'ensemble des couples possibles de lecteurs, étudiés selon l'hypothèse du "worst case", la médiane de sensibilité en double lecture était de 1 [0.96 ; 1], la spécificité de 0.57, la VPP de 0.64 et la VPN de 1.

Avec l'hypothèse du "best case", la médiane de sensibilité en double lecture était de 0.84, la spécificité de 0.86, la VPP de 0.84 et la VPN de 0.86.

IV. Conclusion:

Une double lecture des IRM de prostate en protocole biparamétrique avec l'hypothèse du "worst" case permet d'obtenir une excellente valeur prédictive négative, pouvant potentiellement éviter un recours aux biopsies prostatiques chez les patients dont l'IRM est négative. L'expérience des lecteurs ne semble pas modifier les performances globales de l'examen.

INTRODUCTION

I. Cancer de prostate

A. Epidémiologie

Le cancer de la prostate est le cancer le plus fréquent chez l'homme, représentant 28% des cancers masculins (6).

Il survient dans 66% des cas chez les hommes âgés de plus de 65 ans.

On évalue à environ 53000 les nouveaux cas par an.

Il représente 7,5 % de l'ensemble des décès par cancer et se situe au troisième rang de décès par cancer chez l'homme

Le taux de mortalité du cancer de prostate est en diminution régulière depuis les années 90, du fait de l'amélioration des techniques chirurgicales et du dépistage précoce.

Chez les patients dont la tumeur est localisée, la survie à 10 ans est d'environ 90 % après chirurgie.

Il existe deux facteurs de risque reconnus de cancer de prostate: les patients afro-antillais et les patients présentant un antécédent familial de cancer de prostate.

B. Rappels anatomiques et histologiques

1. Anatomie prostatique

La prostate est un organe musculo-fibro-glandulaire, développé au contact de la portion proximale de l'urètre.

Elle peut être divisée en 3 zones principales, selon le modèle de Mac Neal: la prostate périphérique et la prostate centrale, correspondant à 95% de la masse prostatique chez le sujet jeune. Les 5% restant correspondent à la zone de transition, le tissu péri urétral et le stroma fibromusculaire antérieur (**Figure 1**).

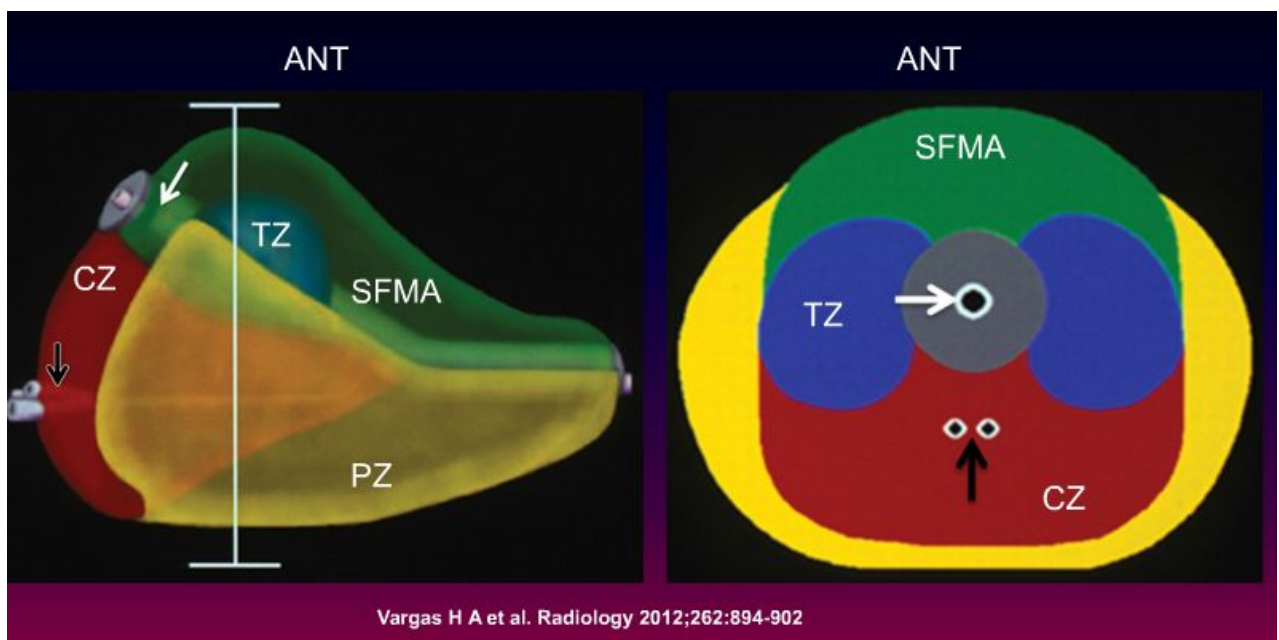


Figure 1: schématisation de l'anatomie zonale de la prostate. (1)

(TZ: zone de transition ; CZ: zone centrale ; SFMA: stroma fibromusculaire antérieur ; PZ: zone périphérique)

Son volume normal chez l'adulte jeune est évalué à 15-20cc.

Avec l'âge, il se développe progressivement une hypertrophie de la zone de transition, correspondant à l'hypertrophie bénigne de prostate.

2. Cancer de prostate

Le cancer de prostate ne provient pas de la dégénérescence de l'hypertrophie bénigne de prostate.

Il se développe primitivement au sein de la zone périphérique dans 70% des cas, plus rarement au sein de la zone de transition (25% des cas) du stroma fibromusculaire antérieur (5% des cas).

Il s'agit presque exclusivement d'adénocarcinomes prostatiques (>95%) même si on rapporte des rares cas de carcinome neuroendocrine, de tumeurs mésoenchymateuses malignes ou de lymphomes (exceptionnels).

Les tumeurs de la zone de transition d'un volume inférieur à 4 cm³ sont localisées préférentiellement au niveau du tiers antérieur et de la moitié inférieure de la zone de transition. (**Figure 2**) (2)

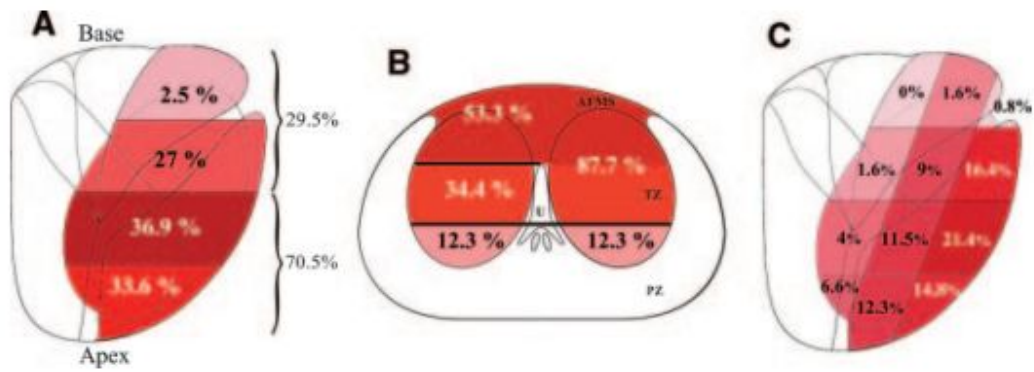


Figure 2: localisation préférentielle des tumeurs de moins de 4cm³ au sein de la zone de transition (2)

A,C : Représentation de la prostate en coupe sagittale ; B : Représentation schématique de la prostate en coupe axiale

Les tumeurs de la zone périphérique de moins de 4cm³ sont localisées préférentiellement au niveau de la moitié inférieure et dans les segments postérieurs (Figure 3).

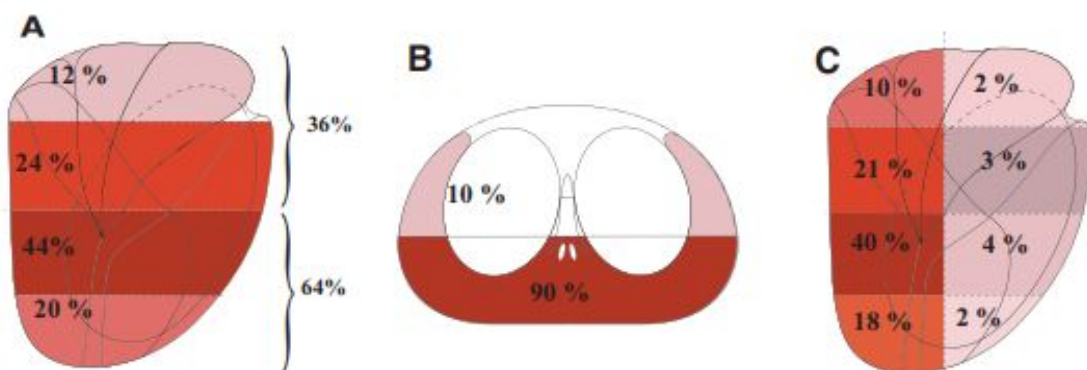


Figure 3: localisation préférentielle des tumeurs de moins de 2cm³ de la zone périphérique (2)

A,C : Représentation de la prostate en coupe sagittale ; B : Représentation schématique de la prostate en coupe axiale

L'évolution de l'adénocarcinome est lente - plusieurs années - dans la majorité des cas. L'extension tumorale est d'abord locale avec franchissement de la capsule prostatique et/ou extension aux vésicules séminales, puis locorégionale avec un envahissement ganglionnaire. L'extension à distance est majoritairement osseuse, alors que les métastases pulmonaires et hépatiques sont rares.

Il s'agit d'une tumeur hormonodépendante qui synthétise, sous l'effet des androgènes, du PSA.

3. Score de Gleason

Le score de Gleason est le système de classification histologique le plus couramment utilisé pour le cancer de la prostate. Il évalue l'agressivité du cancer de la prostate.

Le degré d'atypie des cellules était classé initialement sur une échelle de 1 à 5. Les scores 1 et 2 ne sont plus utilisés en pratique sur les prélèvements biopsiques.

Plus le nombre est élevé, plus cela correspond à des cellules cancéreuses anormales, des foyers mal limités, en rapport avec un degré d'agressivité plus important (**Figure 4**).

Représentation du grade de Gleason

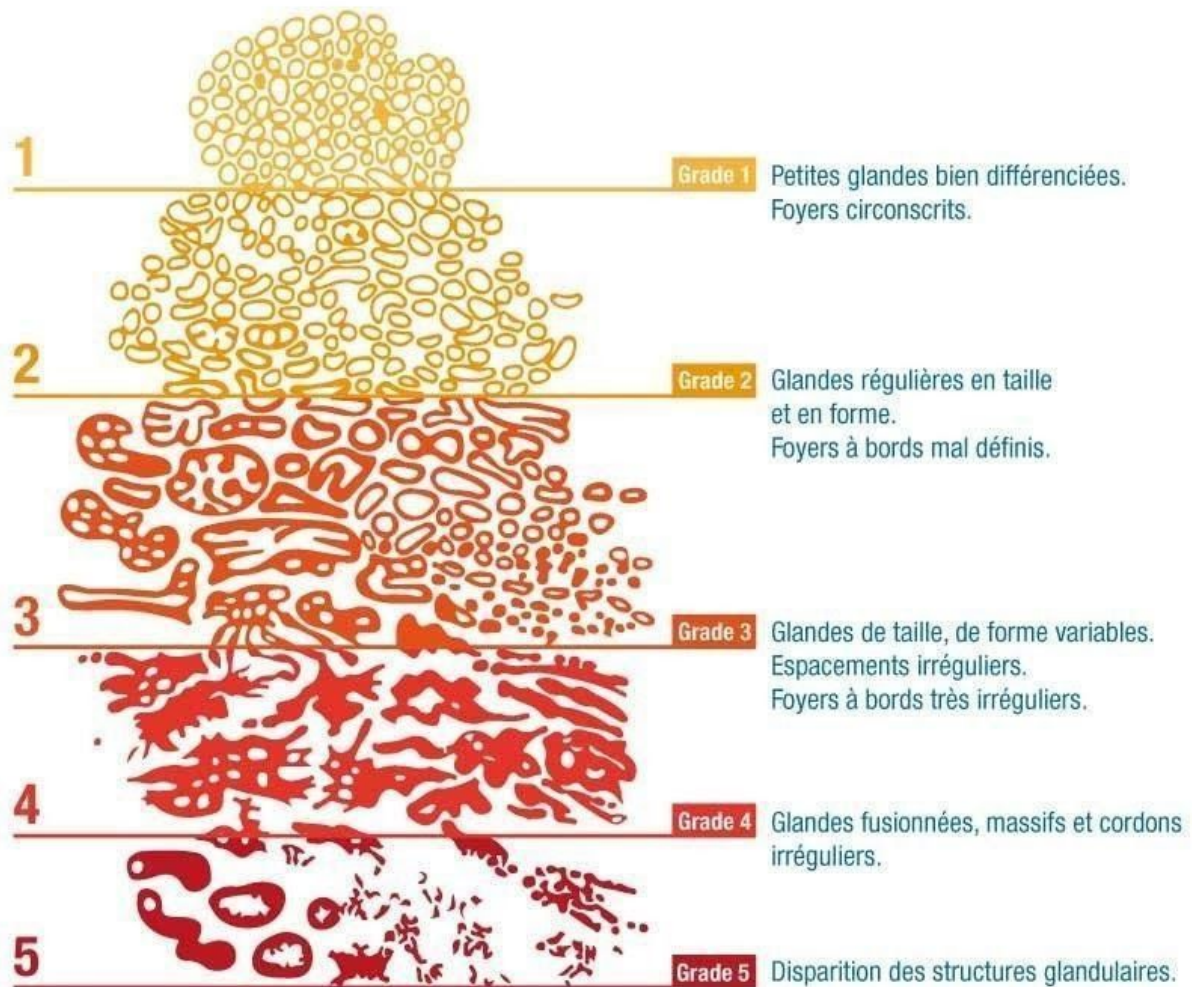


Figure 4 : Schématisation du score de Gleason (3)

Le degré d'agressivité des différentes cellules tumorales peut différer au sein même de la lésion prélevée. Ainsi, l'anatomocytopathologiste assigne un grade de 3 à 5 au niveau de la tumeur ; s'agissant de prélèvements issus de biopsies, on additionne

ensuite le grade le plus fréquent et le grade le plus élevé pour obtenir le score de Gleason.

L'ordre des grades à son importance : le premier chiffre représente le grade le plus représenté dans la tumeur. Ainsi, un cancer qui présente des grades 3 + 4 (score de Gleason 7) sera moins agressif qu'un cancer qui présente des grades de 4 + 3 (score de Gleason 7).

Généralement, plus les grades et les scores sont élevés, plus la tumeur est agressive et croît rapidement.

4. Score ISUP

Depuis 2014, une nouvelle classification est utilisée, affinant l'analyse et surtout utilisant un code plus facile à expliquer. Le grade de Gleason est toujours utilisé, mais la notation changée :

- Groupe ISUP 1 : Gleason 3+3. Correspond à une tumeur bien différenciée (meilleur pronostic)
- Groupe ISUP 2 : Gleason 3+4 (majorité de grade 3). Tumeur bien à moyennement différenciée.
- Groupe ISUP 3 : Gleason 4+3 (majorité de grade 4). Tumeur moyennement différenciée
- Groupe ISUP 4 : Gleason 4+4. Tumeur peu différenciée.
- Groupe ISUP 5 : Gleason 9 ou 10, comprenant des cellules de grade 5, dites indifférenciées : il n'y a plus de ressemblance avec la prostate normale (moins bon pronostic).

C. Débat sur dépistage

Le dépistage du cancer de prostate est très débattu depuis plusieurs années.

S'il est largement démontré qu'un dépistage de masse n'est pas indiqué en raison d'un sur-diagnostic et donc un sur-traitement des cancers non significatifs, un dépistage individuel reste d'actualité. (4)

En effet, l'étude multicentrique ERSPC (5) sur le dépistage systématique du cancer par PSA versus non-dépistage démontre que la première attitude permet une réduction de la mortalité spécifique par cancer de la prostate (**Figure 5**) et une réduction des formes métastatiques diagnostiquées. Ce bénéfice commence à apparaître à partir de 10 ans de suivi.

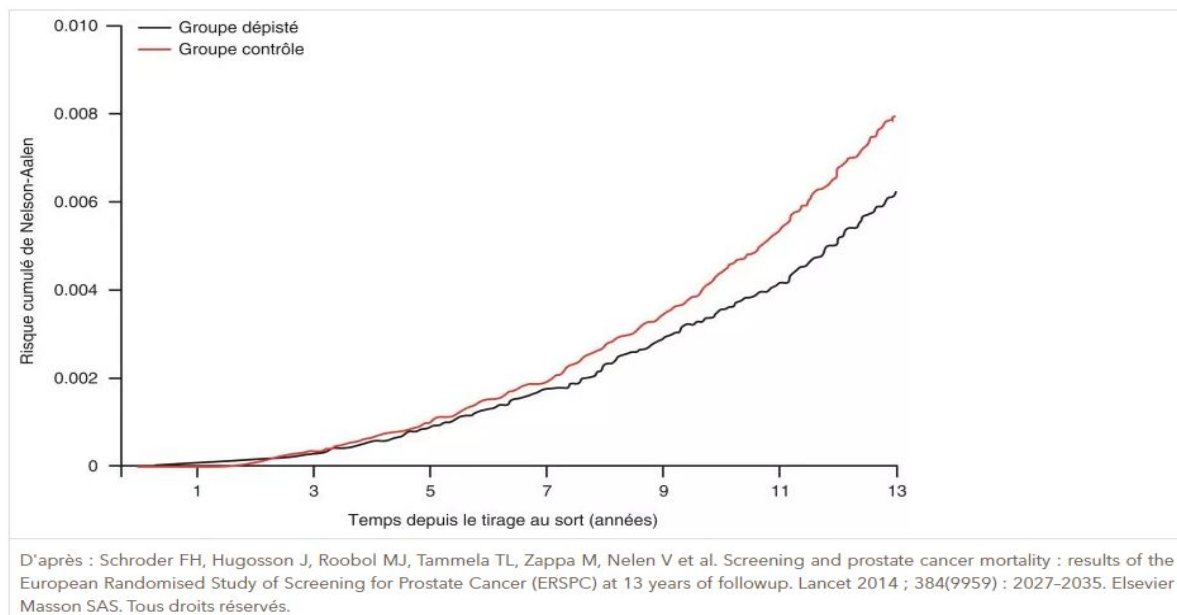


Figure 5: Courbes de décès des hommes inclus dans le programme de dépistage européen (*homme dépisté = « intervention group » randomisé contre non-dépisté = « control group »*) montrant un avantage au-delà de 10 ans de suivi. (5)

Dans ce contexte, l'association Française d'urologie a proposé en 2013 des recommandations (**Figure 6**) concernant un dépistage personnalisé:

Recommandations de l'Association française d'urologie (2013)

- Pas de dépistage de masse systématique du cancer de la prostate.
- Dépistage individuel (diagnostic précoce) après information objective pour ne pas laisser méconnaître et laisser évoluer un éventuel cancer agressif de la prostate.
- Le dépistage individuel repose sur le toucher rectal et sur le dosage du PSA.
- Le dépistage individuel pourrait être recommandé à partir de 45 ans chez les hommes à haut risque de développer un cancer de la prostate : les patients afro-antillais et les patients ayant au moins deux cas collatéraux ou de survenue avant 55 ans.
- Le dépistage individuel peut sinon être recommandé chez les hommes âgés de 50 ans à 74 ans, si l'espérance de vie est > 10 ans.
- Le dépistage individuel n'est pas recommandé chez les hommes dont l'espérance de vie est estimée inférieure à 10 ans en raison d'un âge avancé ou de comorbidités sévères.
- Si le PSA est > 4 ng/mL ou le TR est suspect une consultation urologique est recommandée pour avis, en vue de l'indication d'une biopsie prostatique échoguidée (12 prélèvements).

Figure 6: recommandations de l'AFU concernant le dépistage du cancer de prostate

(6)

D. Diagnostic du cancer de prostate

1. Diagnostic clinique

Dans la plupart des cas, le diagnostic est posé alors que le patient est asymptomatique.

Une symptomatologie fonctionnelle (troubles urinaires irritatifs ou obstructifs, hématurie, altération de l'état général, douleurs osseuses, signes neurologiques) peut faire suspecter un cancer localement avancé ou métastatique.

L'examen clinique repose principalement sur le toucher rectal qui recherche un nodule dur, indolore et évalue le cas échéant l'extension à la capsule ou aux organes de voisinage.

2. Diagnostic biologique

Le dosage du PSA total sérique permet de suspecter le cancer de prostate lorsque le taux est augmenté (plus de 4 ng/mL).

Les valeurs du PSA peuvent être modifiées par l'hypertrophie bénigne de prostate ou par une prostatite, ce qui rend son interprétation parfois difficile.

La densité de PSA permet de s'affranchir de certains obstacles en rapportant le taux de PSA au volume prostatique, permettant d'améliorer les performances du test pour les taux faibles (PSA entre 2.5 et 10 ng/mL). Dans ces conditions, le taux de PSA ne doit pas excéder 1/10ème du volume prostatique (en mL).

Dans les cas limites (PSA entre 4 et 10 ng/mL avec biopsies négatives), le rapport PSA libre / PSA total permet également de lever certains doutes, avec un rapport diminué dans le cancer comparativement à l'adénome de prostate.

Au final, plus qu'une valeur fixe, c'est réellement la cinétique du PSA qui est à prendre en compte: toute élévation progressive du taux de PSA doit conduire à une consultation de principe avec un urologue.

3. Diagnostic histologique

Tout patient présentant un taux de PSA élevé ou un examen clinique anormal doit bénéficier de prélèvements histologiques.

La méthode de référence correspond aux biopsies prostatiques dont la technique est standardisée.

Il est recommandé de réaliser les biopsies après information du patient et antibioprophylaxie orale par prise unique de fluoroquinolone.

Les biopsies sont réalisées par voie transrectale et sous guidage échographique.

Après une anesthésie locale de type bloc péri-prostatique, 12 prélèvements sont réalisés à la partie latérale et médiale de la base, du milieu et de l'apex prostatique, à droite comme à gauche (**Figure 7**).

Le débattement de l'aiguille à biopsier est de 23 mm, expliquant que certaines lésions très antérieures, très apicales ou basales, puissent être sous-échantillonnées par ce mode de prélèvement.

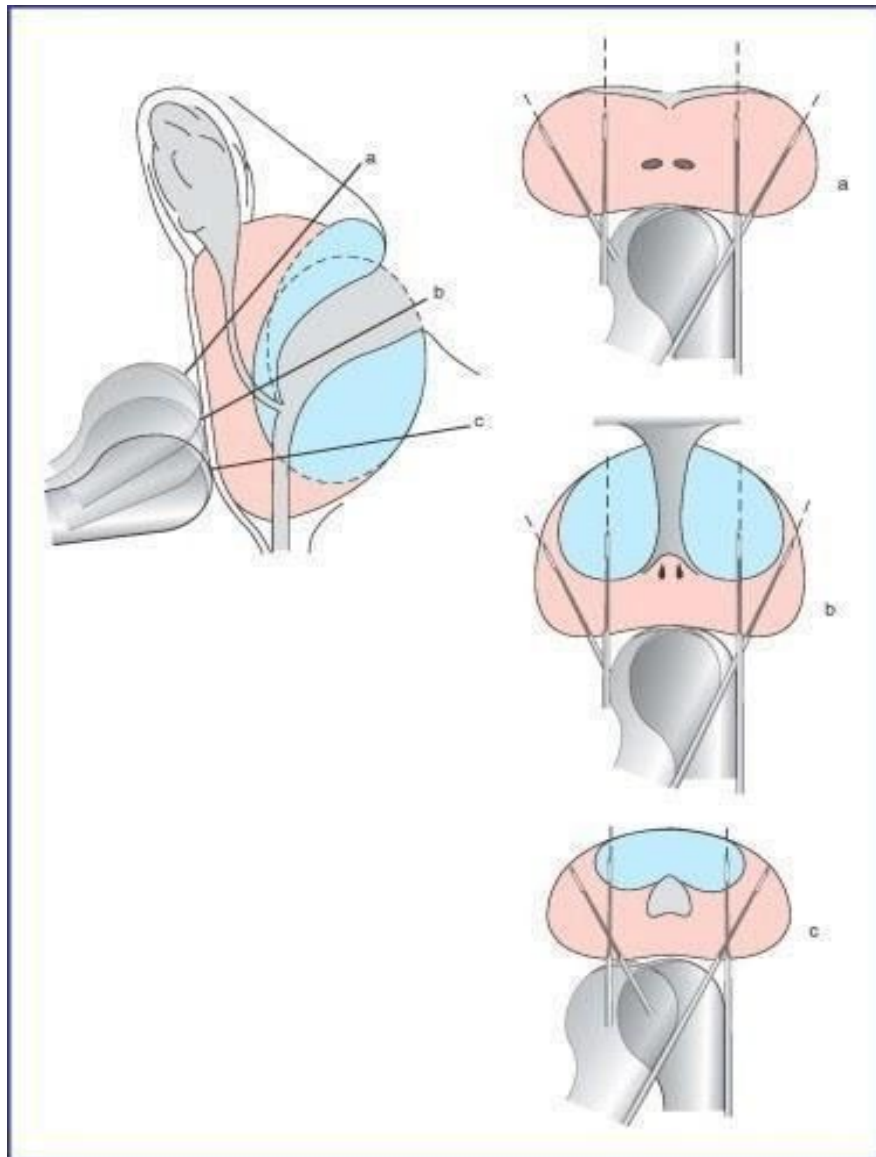


Figure 7: illustration du schéma des 12 biopsies prostatiques (coupes sagittales et axiales).(6)

Représentation schématique de la prostate en coupe sagittale (à gauche) et en coupes axiales (sur la droite)

4. Complications

Les complications des biopsies prostatiques les plus fréquentes sont les saignements (hématurie, rectorragie, hématospermie), souvent transitoires et peu importants.

Les complications infectieuses sont classiques mais beaucoup moins fréquentes (2-4% selon les études) depuis l'instauration de l'antibioprophylaxie systématique. (7)

Les troubles du bas appareil urinaire sont également rapportés; le taux de rétention d'urine est évalué entre 0.2 et 1.7%, tandis qu'un inconfort global et une majoration de la symptomatologie urinaire est rapportée dans 6 à 25% des cas (7).

La prise en charge optimale des complications, notamment infectieuses, repose sur une information du patient, une hospitalisation et une antibiothérapie adaptée.

II. IRM de prostate

A. IRM-mp

L'IRM de prostate a une place de plus en plus importante dans la prise en charge diagnostique du cancer de prostate.

Le protocole d'acquisition est standardisé.

L'IRM doit être réalisée avant les biopsies ou au moins deux mois après pour s'affranchir des saignements post biopsie.

1. Préparation

Il est recommandé de préparer les patients avant l'examen.

L'administration d'un antispasmodique permet de réduire le péristaltisme rectal pour limiter les artéfacts cinétiques.

Un lavement rectal permet de diminuer la présence d'air et de selles dans le rectum, limitant de ce fait les artéfacts de distorsion au contact.

2. Protocole

L'IRM multiparamétrique de prostate comprend des séquences en pondération T2, des séquences de diffusion, des séquences de perfusion en pondération T1 avant et après injection de gadolinium.

On peut réaliser une IRM de prostate avec un champ magnétique de 1,5 Tesla (1,5T) ou 3 Tesla (3T).

L'IRM 3T a l'avantage d'une augmentation du rapport signal sur bruit, permettant une amélioration de la résolution spatiale, de la résolution temporelle, ou des deux.

Les séquences T2 permettent une analyse morphologique. Elles comprennent des séquences axiales et des séquences sagittales ou coronales pour localiser les anomalies dans le plan cranio caudal. On recommande une épaisseur de coupe de 3mm et un FOV (champ de vue) de 12 à 20cm pour couvrir les vésicules séminales.

Les séquences de diffusion doivent comprendre des valeurs de B élevées $b > 1400 \text{ s/mm}^2$ et reconstruction de cartographie ADC (coefficient apparent de diffusion).

Les séquences de perfusion correspondent à des séquences T1 écho de gradient acquises avant, pendant et après injection de gadolinium. La résolution temporelle doit être inférieure à 7 secondes et le temps d'acquisition de plus de 2 minutes.

Les séquences réalisées avant injection permettent de dépister des remaniements hémorragiques.

Les séquences de perfusion sont discutables chez les patients naïfs de biopsie mais indispensables dans les cas difficiles, notamment chez les patients aux antécédents de néoplasie prostatique opérée ou traitée par radiothérapie.

Une séquence T2 "rapide" (ex: BTFE) sur un champ plus étendu permet d'évaluer l'intégrité des chaînes ganglionnaires jusqu'aux hiles rénaux.

3. Indications

L'IRM ne doit pas être considérée à elle seule comme un examen de dépistage du cancer de prostate.

Les recommandations actuelles préconisent de réaliser une IRM prostatique avant toute série de biopsies prostatiques.

L'intérêt majeur est de permettre la réalisation de biopsies dirigées soit "cognitives", c'est à dire en prenant en compte mentalement les données de l'IRM, soit avec des méthodes de fusion d'image IRM/Échographie.

Certaines lésions étant sous-échantillonnées par la méthode des 12 biopsies systématisées, les biopsies dirigées permettent dans certains cas de rattraper quelques faux négatifs.

L'IRM apporte également, en complément de l'examen clinique, des informations cruciales dans le cadre du bilan d'extension locorégional ; en plus de la mesure des lésions, l'examen permet de se prononcer sur l'extension extra prostatique le cas échéant.

En effet, le franchissement de la capsule (T3a) ou l'envahissement des vésicules séminales (T3b) sont des critères importants dans l'évaluation du TNM du cancer de prostate (**Figure 8**), avec une influence sur les modalités thérapeutiques.

- **T : Tumeur primitive**
 - **T0** : Absence de tumeur
 - **T1** : Tumeur non palpable ou non visible en imagerie
 - **T1a** < 5 % du tissu réséqué* et score de Gleason 6
 - **T1b** > 5 % du tissu réséqué* et/ou score de Gleason 7
 - **T1c** : découverte par élévation du PSA et réalisation de biopsies
 - **T2** : Tumeur limitée à la prostate (apex et capsule compris)
 - **T2a** : Atteinte de la moitié d'un lobe ou moins
 - **T2b** : Atteinte de plus de la moitié d'un lobe sans atteinte de l'autre lobe
 - **T2c** : Atteinte des deux lobes
 - **T3** : Extension au-delà de la capsule
 - **T3a** : Extension extra-capsulaire
 - **T3b** : Extension aux vésicules séminales
 - **T4** : Extension aux organes adjacents (sphincter urétral, rectum, paroi pelvienne) ou tumeur fixée
- **N : Ganglions régionaux**
 - **Nx** : ganglions régionaux non évalués
 - **N0** : Absence de métastase ganglionnaire
 - **N1** : Atteinte ganglionnaire(s) régionale(s)
 - **N1mi** : Métastase ganglionnaire < 0,2 cm
- **M : Métastases à distance**
 - **M0** : Absence de métastases à distance
 - **M1** : Métastases à distance
 - **M1a** : Ganglions non régionaux
 - **M1b** : Os
 - **M1c** : Autres sites

Figure 8: score TNM du cancer de la prostate (6)

Lorsque la tumeur reste limitée à la prostate, l'urologue peut dans certains cas réaliser une dissection intra faciale, permettant de préserver les bandelettes nerveuses, avec moins de risque d'influer sur la fonction érectile future. Lorsque la tumeur envahi la capsule, une dissection extra faciale est nécessaire, compromettant la fonction érectile ultérieure.

Enfin, l'IRM permet de compléter le bilan locorégional en dépistant d'éventuelles adénomégalies iliaques ou lombo aortiques.

L'IRM a également de plus en plus son importance dans le cadre de la documentation des récives après traitements (prostatectomie, radiothérapie, curiethérapie, thérapie focale...), en permettant de localiser les anomalies lorsqu'une réélévation du PSA fait suspecter une récive.

B. IRM bi paramétrique de prostate

L'IRM bi paramétrique de prostate correspond à un examen réalisé sans injection de produit de contraste.

Plus rapide, moins chère et plus simple à réaliser, l'IRM bi paramétrique présente des avantages certains.

L'injection de produit de contraste permet d'aider dans les cas difficiles l'interprétation. Cependant, il existe un réel débat quant à son intérêt réel en termes de discrimination lésionnelle. En effet, le rôle des séquences de perfusion est mineur dans les critères du standard PI-RADS v2.1.

En outre, des études récentes ont montré que l'IRM bi paramétrique était équivalente à l'IRM multiparamétrique chez les patients naïfs de biopsie prostatique (8).

La place et le rôle exact de chaque modalité reste donc à définir dans les années à venir.

III. Standard PI-RADS2.1

Plusieurs systèmes de classification des lésions ont vu le jour depuis l'avènement de l'IRM de prostate. Le score le plus utilisé en France et en Europe est le score PIRADS v2.1, dont le but est d'homogénéiser l'interprétation et améliorer la concordance entre lecteurs.

Ce score classe de 1 à 5 les lésions, avec un risque de cancer croissant progressivement avec le score (**Figure 9**).

Les IRMs classées 1 correspondent à des IRM normales.

Les lésions classées 5 correspondent à des lésions fortement suspectes de cancer.

Ce score repose principalement sur les séquences T2 (notamment pour la zone de transition) et diffusion (notamment pour la zone périphérique). (**Figures 10,11**)

Les séquences de perfusion ont un rôle moins important dans la classification PIRADS mais permettent dans certains cas de surclasser ou sous classer certaines lésions de la zone périphérique

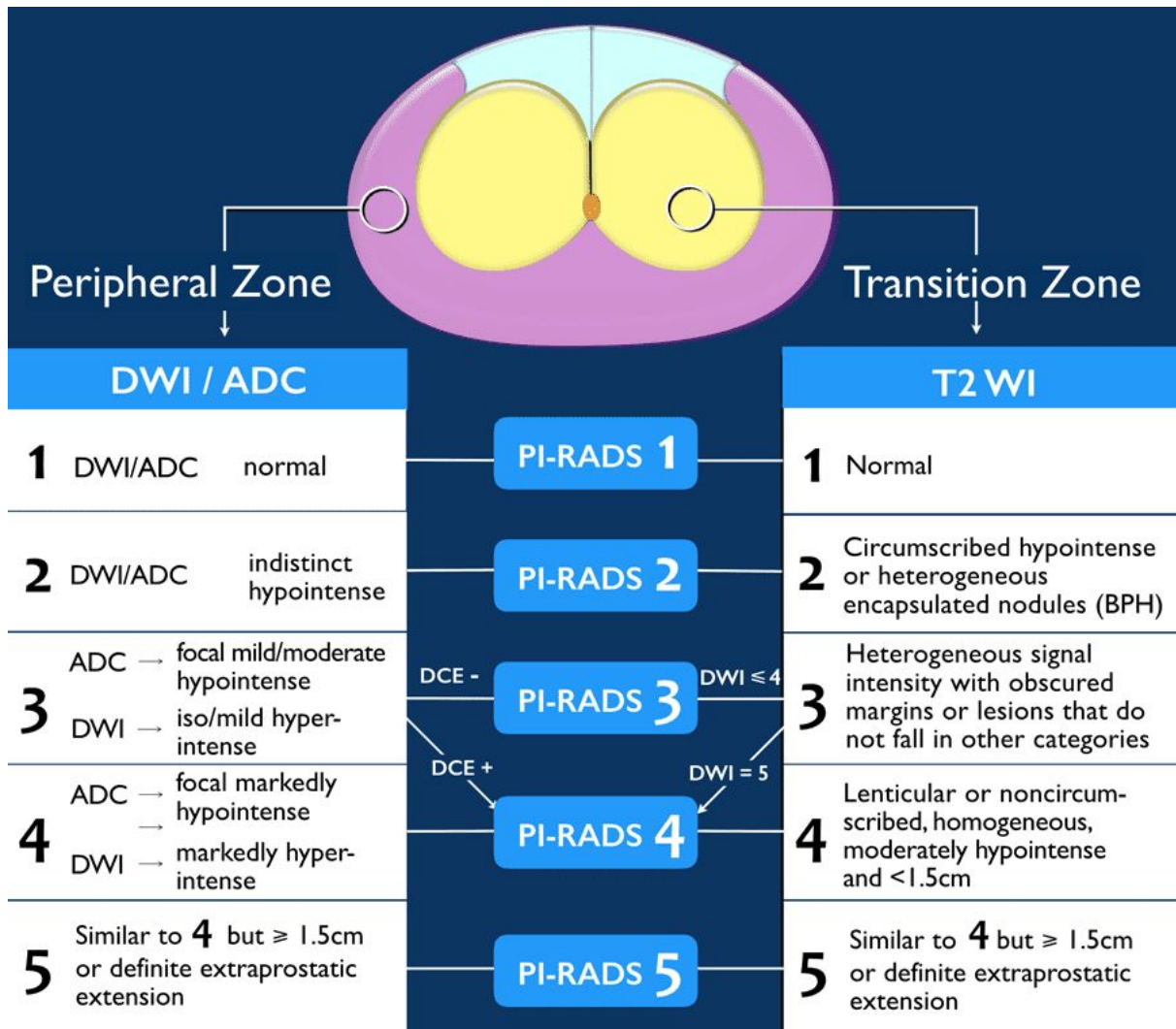


Figure 9: PI-RADS2.1:

Classement des lésions de la zone périphérique ou de la zone de transition selon les anomalies en diffusion (DWI), Perfusion (DCE) et sur la séquence T2. (9)

Peripheral Zone			
ADC	ADC	DWI	T2WI
1 Normal			
2 ADC: Indistinct hypointense			
3 ADC: focal mild/moderate hypointense DWI: iso/mild hyperintense < 1.5 cm			
4 ADC: focal markedly hypointense DWI: markedly hyperintense < 1.5 cm			
5 Similar to 4 but ≥ 1.5cm or definite extraprostatic extension			

Figure 10: Score Pi-RADS des lésions de la zone périphérique (9)

Anomalies de la zone périphérique retrouvées pour des scores PIRADS de 1 à 5. Coupes axiales de la prostate sur des séquences T2, diffusion (DWI) et cartographie ADC

Au sein de la zone périphérique, les lésions les plus suspectes correspondent à des images focales, en hypersignal diffusion, hyposignal ADC, avec prise de contraste précoce et intense.

Ces lésions sont d'autant plus suspectes que leur taille est importante (>1.5cm) ou qu'il existe des signes d'extension extra prostatique.

Transition Zone			
T2WI	T2WI	ADC	DWI
1 Normal			
2 Circumscribed hypointense or heterogeneous encapsulated nodules (BPH)			
3 Heterogeneous signal intensity with obscured margins or lesions that do not fall in other categories			
4 Lenticular or noncircum-scribed, homogeneous, moderately hypointense and <1.5cm			
5 Similar to 4 but \geq 1.5cm or definite extraprostatic extension			

Figure 11: Score Pi-RADS des lésions de la zone de transition.(9)

Anomalies de la zone de transition retrouvées pour des scores PIRADS de 1 à 5. Coupes axiales de la prostate sur des séquences T2, diffusion (DWI) et cartographie ADC

Au sein de la zone de transition, les lésions les plus suspectes correspondent à des images focales, en hyposignal T2 franc, non ou mal circonscrites.

Ces lésions sont d'autant plus suspectes que leur taille est importante (>1.5cm) ou qu'il existe des signes d'extension extra prostatique.

RATIONNEL DE L'ÉTUDE

Le cancer de la prostate est le cancer le plus fréquent chez l'homme de plus de 50 ans, et représente la seconde cause de décès par cancer chez l'homme en France.

(10)

Il s'agit d'un réel problème de santé publique.

Il existe encore à l'heure d'aujourd'hui un débat sur l'intérêt du dépistage du cancer de prostate (11,12). Il est clairement établi qu'un dépistage individuel est préféré au dépistage de masse (8), en raison d'un risque lié au sur traitement des cancers diagnostiqués à un stade trop précoce.

Pourtant, il y a toujours une proportion non négligeable d'hommes diagnostiqués avec une maladie avancée au moment du diagnostic, et les algorithmes de diagnostic actuels peuvent être améliorés, avec des outils permettant de distinguer les cancers de bas grade, sans risque à court ou moyen terme pour le patient, de plus haut lésions de grade (cancer de la prostate cliniquement significatif) avec un risque à court ou moyen terme pour le patient.

L'IRM multiparamétrique de la prostate (IRM-mp) en est l'exemple. Elle apparaît aujourd'hui comme un outil essentiel pour identifier le cancer de la prostate, aider à cibler les biopsies, évaluer son agressivité et suivre ses récives.

Depuis 2012, la standardisation des protocoles d'acquisition, des techniques de lecture et de compte-rendu standardisé selon le PI-RADS a permis de développer l'utilisation de l'imagerie prostatique.

Pourtant, on reproche encore souvent à l'IRM prostatique une importante variabilité inter observateurs (21,22,23)

Plusieurs études récentes (13,14) ont montré l'apport majeur de l'IRM lorsqu'elle est réalisée avant la première série de biopsies, en augmentant le nombre de cancers significatifs dépistés, en diminuant le nombre de cancers non significatifs dépistés, et en permettant dans certains cas de surseoir aux biopsies lorsque l'examen est rassurant.

Ces avancées ont mené en 2019 à des recommandations généralisant l'utilisation de l'IRM de prostate avant toute biopsie prostatique. (15,16)

Récemment, plusieurs études ont montré l'intérêt d'une variante de la mp-IRM, dite «bi-paramétrique» IRM de la prostate (bp-IRM), chez des patients naïfs de biopsie (17). Avec uniquement des séquences T2-W et DWI, et aucune utilisation de l'imagerie de perfusion (DCE). Cet examen court (10 min) et moins cher, ne nécessitant aucune injection intraveineuse de produit de contraste, permettrait un meilleur triage des patients atteints d'un cancer de la prostate cliniquement significatif (CsPCa) au moment de l'IRM, grâce à une spécificité et une valeur prédictive négative plus élevées, avec une perte de sensibilité mineure.

Sa place au sein du dépistage du cancer de la prostate reste donc aujourd'hui à définir clairement.

Néanmoins, même en utilisant le dernier référentiel PI-RADS2.1 (18,19) la sensibilité et la spécificité de l'IRM-mp et de l'IRM-bp sont variables en fonction de l'expérience

des lecteurs (21,22), et sa VPP et VPN dépendent fortement de la prévalence du cancer dans la population.

Sur la base d'études récentes et de travaux antérieurs (non publiés) dans notre service(20), nous avons conçu cette étude pour évaluer les performances d'un protocole d'IRM biparamétrique de prostate pour la sélection de patients ayant un cancer de prostate significatif à la biopsie, dans une population de patients consécutifs naïfs de biopsie, en tenant compte d'une variation potentielle de l'expérience des lecteurs.

Nous avons étudié les performances de la double lecture en utilisant une analyse du «best case» par rapport au «worst case».

MATÉRIELS ET MÉTHODES

I. Population de patients

Nos critères d'inclusion étaient : patients consécutifs consultant pour une IRM prostatique suivie d'une première série de biopsie prostate, avec IRM prescrite pour suspicion de néoplasie prostatique sur élévation du PSA et/ou anomalie au toucher rectal.

Nos critères d'exclusion étaient : les antécédents de biopsie prostatique, les antécédents de traitement chirurgical, de thérapie focale ou de radiothérapie prostatique. IRM réalisée pour une autre indication que la recherche de néoplasie. IRM ininterprétable sur l'une des séquences.

Nous avons inclus rétrospectivement tous les patients consécutifs ayant bénéficié d'une IRM prostatique suivie de biopsies de la prostate entre le 5 janvier 2016 et le 9 février 2016.

45 patients ont été exclus pour les raisons suivantes : absence de données histologiques après l'IRM (n =9), antécédent de biopsies réalisées avant l'IRM (n=27), examen fait hors contexte oncologique (post-embolisation prostatique,

recherche d'abcès) (n=1), examen présentant une des 3 séquences non interprétables (n=2), examen IRM réalisé après prostatectomie, radiothérapie (n=5), curie thérapie ou HIFU (n=1).

Au total, nous avons inclus 55 examens d'IRM de prostate, réalisés chez 55 patients (Figure 11).

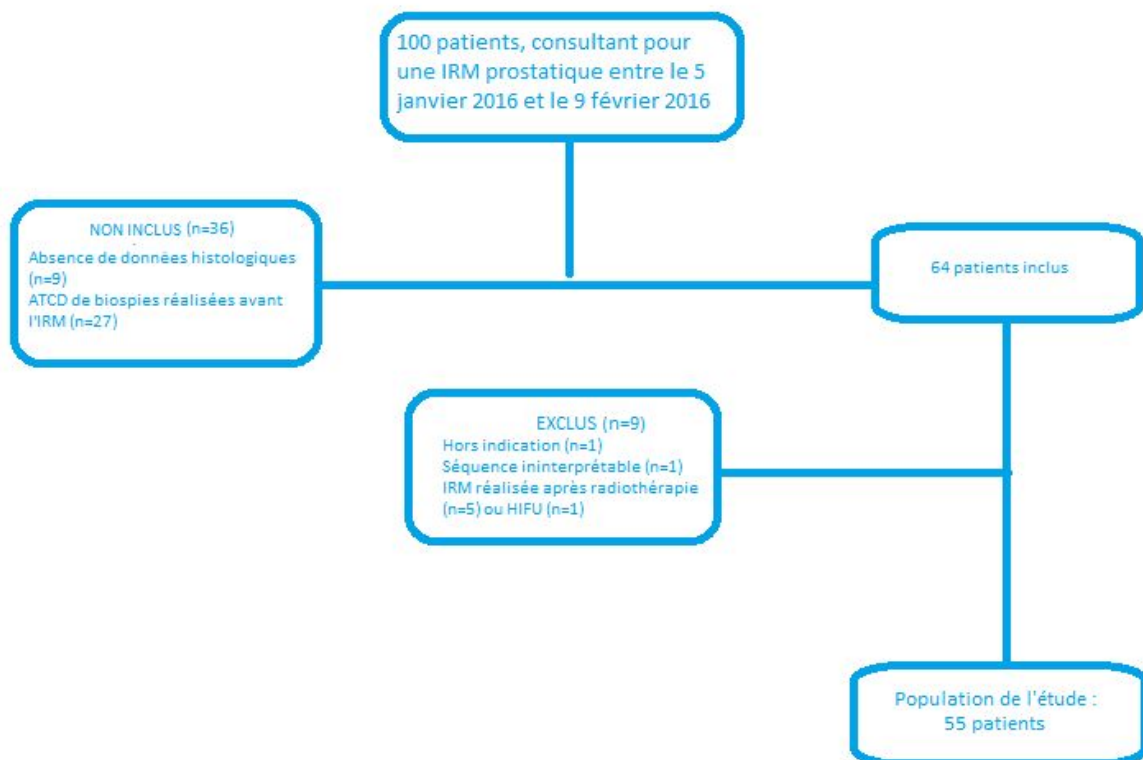


Figure 11 : Diagramme de flux de la population de patients
Après application des critères d'inclusion et d'exclusion, 55 patients ont pu participer à l'étude, sur les 100 patients recrutés.

Chaque patient a été informé par courrier de la tenue de l'étude, et que les données de leur IRM allaient pouvoir être utilisées à des fins de recherche. Les patients avaient la possibilité de s'y opposer.

Nous n'avons relevé aucun refus. Six patients ont cependant demandé d'être informés des résultats de l'étude.

Les données ont été analysées rétrospectivement après information des patients et validation par le comité hospitalier en charge de la gestion des données (attestation n°DEC18-366).

II. Protocole d'imagerie

Tous les examens IRM ont été réalisés dans le centre d'imagerie génito-urinaire du CHRU de Lille (Hôpital Huriez) sur une IRM Philips Ingenia de type 1,5 Tesla (Philips Healthcare, Koninklijke, NL). Tous les patients ont bénéficié d'un protocole standardisé à l'aide d'une antenne abdominale à 32 éléments: séquences axiales et sagittales T2 (TE = 110 ms, TR = 3000 ms, matrice : 224 x 220, FOV : 180 x 180 mm, épaisseur de coupe : 3 mm), séquences axiales en diffusion avec multiples valeurs de gradient b (TE = 95 ms, TR = 4347 ms, matrice : 68 x 59, FOV : 200 x 178 mm, épaisseur de coupe : 4 mm, valeurs de b : 0, 150, 1000, 2000), séquences axiales T1 en écho de gradient dynamiques (TE = 4.2 ms, TR = 8.9 ms, angle de bascule : 17°, épaisseur de coupe : 3 mm, FOV : 160 x 150 mm, matrice : 128 x 114, résolution temporelle : 15.8 secondes) avant et après injection de gadolinium (Dotarem, Guerbet, Villepinte, FR ; vitesse d'injection : 2 mL/sec).

III. Protocole de biopsies

Tous les patients inclus ont bénéficié, dans les suites de l'IRM, d'une série de biopsies prostatique par voie transrectale, selon le schéma de 12 biopsies systématiques. Après information des patients et recueil de leur consentement écrit, une anesthésie locale par bloc péri prostatique était réalisée. Les prélèvements étaient effectués sous guidage échographique, à l'aide d'une aiguille BARD 18 Gauges (Bard Biopsy Systems, Tempe, Arizona, US) selon un schéma de 12 biopsies systématiques auxquelles s'ajoutaient éventuellement, selon les résultats de l'imagerie, des biopsies dirigées complémentaires (chez 31 patients). Les opérateurs (radiologues ou urologues) disposaient des données de l'imagerie IRM avant le geste. Une antibioprophylaxie était réalisée par une prise unique de fluoroquinolone systémique orale.

IV. Population de lecteurs

Huit lecteurs ont été sélectionnés sur la base du volontariat et répartis en deux groupes, en fonction de leur expérience professionnelle spécifique en IRM prostatique :

a) un groupe "juniors" de quatre lecteurs internes en radiologie ayant validé au moins 6 semestres de radiologie, dont au minimum un semestre spécialisé en imagerie

génito-urinaire. Durant ce semestre, les internes recevaient 3 cours de 2h sur l'IRM prostatique et assistaient à environ 10 vacations de 4 heures d'IRM prostatique. Ces internes étaient formés exclusivement au PI-RADS 2.1

b) un groupe "seniors" de quatre radiologues ayant une expérience d'interprétation d'au moins 500 dossiers relus annuellement par chacun, dont 3 membres du groupe ESUR (European Society of Urogenital Radiology) prostate, utilisant le PI-RADS 2.1 depuis sa publication.

Aucun lecteur n'a participé au design de l'étude ou n'a eu accès aux examens avant le début de l'étude.

V. Affectation des cas aux lecteurs

Les examens des patients ont été anonymisés pour lecture en aveugle, puis classés dans un ordre aléatoire par le logiciel d'interprétation.

Chaque lecteur junior ou sénior s'est ainsi vu attribuer les mêmes 55 examens anonymisés dans un ordre de lecture différent.

Le tableau de correspondance de l'anonymisation n'était connu que de l'investigateur principal (n'appartenant pas au groupe des lecteurs) et non communiqué aux lecteurs.

VI. Préparation des données

Les images de chaque examen IRM ont été préparées de façon standardisée en capturant les images d'une console d'interprétation Portal (Philips Healthcare,

Koninklijke, NL) affichant l'examen de la même façon en 2 lignes de 2 images incluant a) T2 axial, b) T2 sagittal, c) 4 images de diffusion natives, d) carte ADC.

Les lecteurs ne disposaient pas, par choix des investigateurs de l'étude, des données de la perfusion après injection.

Pour chaque examen, une vingtaine d'images étaient capturées automatiquement, allant de l'apex prostatique à la partie moyenne des vésicules séminales dans le plan cranio-caudal.

Le niveau de coupe de la séquence axial T2 était représenté par une ligne sur la séquence sagittale. La mesure en taille des lésions était réalisable sur le logiciel.

Le fenêtrage des niveaux de gris et le zoom étaient quant à eux figés.

VII. Protocole de lecture des dossiers

Chaque lecteur disposait d'une session de travail personnalisée incluant les 55 dossiers qui lui étaient assignés.

La lecture des dossiers pour les "juniors" et les "seniors" s'est déroulée entre le 27/12/2019 et le 18/01/2020. Chaque lecteur pouvait accéder depuis n'importe quel PC avec connexion internet à sa session de travail personnalisée. Cet accès était sécurisé par un mot de passe, différent pour chaque lecteur, qui était attribué par le responsable de l'étude.

La lecture de l'ensemble des dossiers pouvait être effectuée en un ou plusieurs temps.

Toutes les lectures se sont faites de façon individuelle, sans communication entre les lecteurs.

Par chaque examen, le radiologue disposait de l'âge du patient, des éventuels antécédents personnels ou familiaux de cancer, de son taux de PSA et des données du toucher rectal si disponibles.

Un questionnaire standardisé permettait de classer l'examen dans un des deux groupes suivants:

- **Groupe A:** IRM normale ou IRM présentant une anomalie considérée comme bénigne, ou des anomalies de signal non spécifiques, c'est à dire aucune cible suspecte de cancer (PIRADS 1 ou 2)
- **Groupe B:** présence d'une ou plus d'une lésion cible suspecte de cancer (PIRADS 3,4 ou 5).

Après validation de la réponse, chaque lecteur avait accès au résultat anatomopathologique (biopsies et éventuellement prostatectomie). Il n'était pas possible de modifier sa réponse après validation.

Chaque examen a donc été classé par huit opérateurs différents (quatre juniors et quatre seniors) avec des lectures réalisées en aveugle les unes des autres, et classé par chaque lecteur selon deux groupes A ou B en IRM.

Le temps de lecture pour chaque dossier et par lecteur était mesuré automatiquement par le logiciel. Le temps de lecture par dossier a été comparé entre les différents lecteurs à l'aide du test de Kruskal-Wallis.

Nous avons également comparé le temps d'interprétation entre les dossiers de cancer significatif et de cancer non significatifs au moyen d'un test de Wilcoxon.

VIII. Gold standard

Pour chaque patient, le gold standard était défini par le résultat des biopsies qui étaient réalisées immédiatement après l'IRM lorsque cela était réalisable (dans la majorité des cas). Dans le cas contraire, elles étaient réalisées au plus tard dans les 7 jours suivants. On disposait pour chaque examen du nombre de carottes (de 12 à 16), de la longueur de la carotte, du résultat de l'analyse anatomopathologique (Bénin, PIN de haut grade, adénocarcinome).

Le cas échéant, on disposait également du score de Gleason (ISUP), de la longueur de cancer observée sur la carotte, d'un éventuel envahissement péri nerveux.

Chaque patient était ensuite classé en deux groupes:

- a) pas de cancer de prostate ou cancer de prostate non cliniquement significatif (groupe A) ;
- b) cancer de prostate cliniquement significatif (« groupe pathologique B »).

Le cancer de prostate significatif était défini par un score ISUP ≥ 2 ou la présence d'un score ISUP = 1 avec une MCCL ≥ 6 mm (17)

IX. Analyse statistique des résultats

La performance de chaque lecteur pour classer les examens en groupe A ou B a été calculée en comparant les réponses au gold standard (Se, Sp, VPP, VPN).

La concordance entre les différents lecteurs a été analysée par un test de Krippendorff.

Les réponses des lecteurs ont ensuite été combinées en simulant tous les binômes de lecteurs possibles (28 binômes au total: 16 binômes junior-sénior, 6 binômes junior-junior et 6 binômes sénior-sénior) afin d'évaluer les performances d'une double lecture.

Pour s'affranchir des discordances entre deux lecteurs, nous avons testé deux hypothèses :

La première, dite du "best case":

- Si l'un des deux lecteurs classait l'IRM en groupe A (absence de lésion significative), le classement du binôme de lecteurs était A.
- L'IRM était donc classée A si la réponse du binôme était A/A, B/A ou A/B.
- Pour que l'IRM soit classée B, il fallait ainsi que les deux lecteurs l'aient classée dans le groupe B.

La seconde hypothèse, dite du “worst case” :

- Si l'un des deux lecteurs classait l'IRM en groupe B (suspecte de lésion significative), le classement du binôme de lecteurs était B.
- L'IRM était donc classée B si la réponse du binôme était B/B, B/A ou A/B.
- Pour que l'IRM soit classée A, il fallait ainsi que les deux lecteurs l'aient classée dans le groupe A.

Deux nouvelles réponses “combinées” (“worst case” et “best case”) étaient obtenues, et comparées au gold standard.

Au total, nous avons été en mesure de calculer Sensibilité (Se), Spécificité (Spé), Valeur Prédicative Positive (VPP) et Valeur Prédicative Négative (VPN), pour chaque binôme en double lecture, pour le "worst" et pour le "best" case.

Les performances globales de la double lecture ont ensuite été mesurées selon la médiane des différents paramètres statistiques pour l'ensemble des couples, puis par expérience (junior/senior).

Les concordances au sein de chaque binôme ont été analysées à l'aide d'un test Kappa.

Toutes les statistiques ont été réalisées dans le logiciel SAS (SAS Software V9.4, SAS Institute, CARY, NC, USA)

RÉSULTATS

I. Population de patients

Les patients inclus dans l'étude étaient des hommes âgés de 50 à 82 ans (médiane: 64,4 ans) adressés par un urologue pour la réalisation d'une IRM prostatique.

Le taux moyen de PSA était de 7.26ng/mL.

Les biopsies retrouvaient un cancer significatif chez 25 (45%) des 55 patients.

Parmi eux, on retrouvait de l'adénocarcinome de type Gleason 6 chez 6 patients, Gleason 7 chez 19 patients (12 Gleason 3+4 et 7 Gleason 4+3).

Chez les 30 patients restants, les biopsies retrouvaient du cancer non significatif (Gleason 6) chez 6 patients et sont revenues normales chez 24 patients.

Le score ISUP était de 1 chez 36 patients (65%), de 2 chez 12 patients (22%) et 3 chez 7 patients (13%)

II. Temps de lecture par dossiers

Le temps de lecture par dossier s'étendait de 67 à 247 secondes, avec un temps de lecture moyen par dossier de 136 secondes (**figure 13**).

Le temps de lecture moyen pour les juniors était de 99 secondes contre 172 secondes chez les seniors (**figure 12**)

La durée moyenne de lecture la plus courte était retrouvée chez les seniors, avec 67 secondes par dossier.

La durée moyenne de lecture la plus longue était également retrouvée chez les seniors, avec 325 secondes en moyenne par dossier.

Le temps de lecture par dossier était significativement différent entre les différents lecteurs ($p < 0.001$).

Il n'existait pas de différence significative en termes de temps de lecture entre les dossiers de cancers significatifs ou non significatifs ($p = 0.95$)

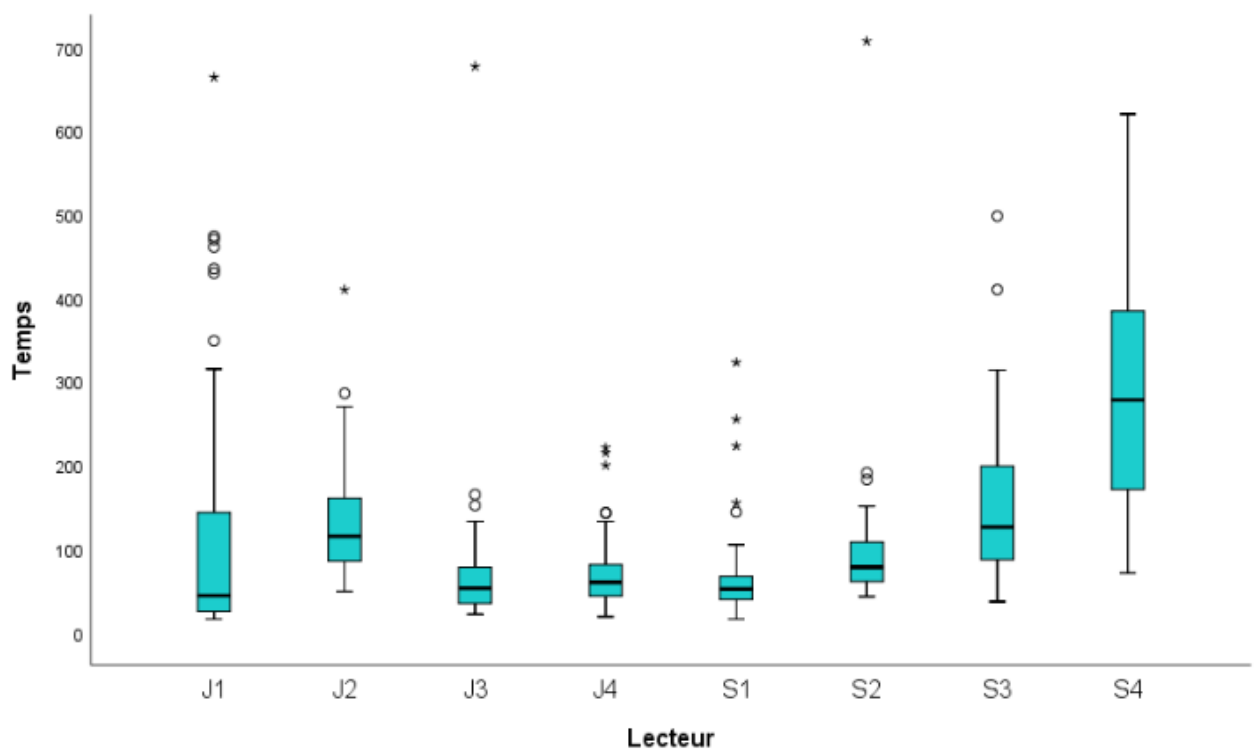


Figure 12: temps de lecture moyen par lecteur:

Moyenne et écart type du temps de lecture (en secondes) des 55 dossiers, pour chacun des lecteurs juniors (J1 à J4) et seniors (S1 à S4)

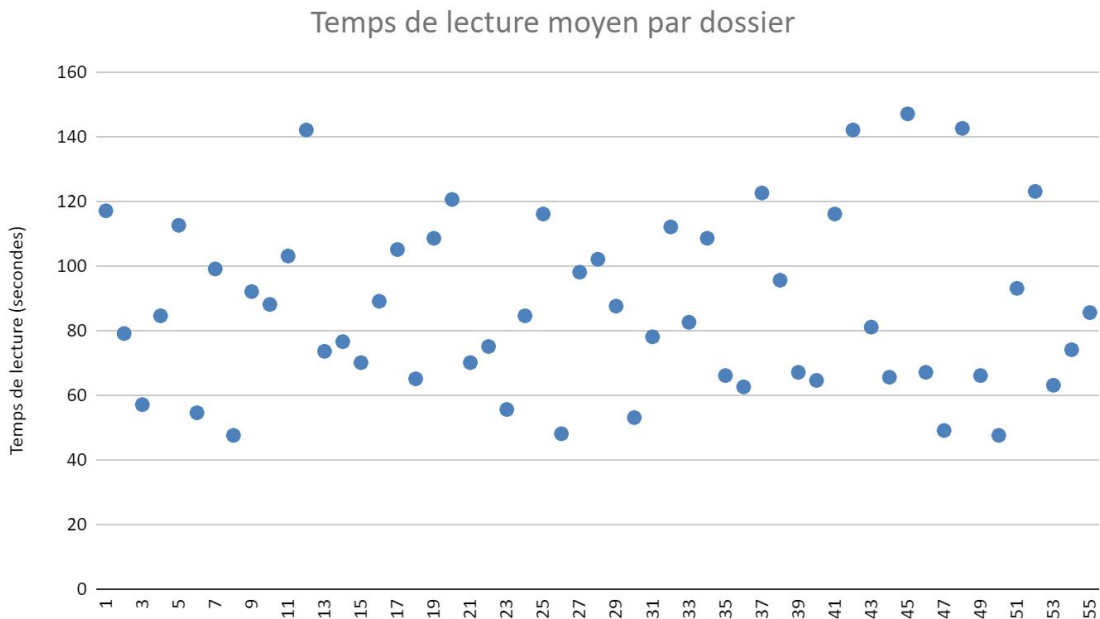


Figure 13: temps de lecture moyen par dossier
Moyenne du temps de lecture (en secondes) des 8 lecteurs pour chacun des 55 dossiers

III. Taux de concordance entre les lecteurs

La concordance entre les huit lecteurs a été étudiée au moyen d'un test de Krippendorff.

Le coefficient alpha de Krippendorff était de 0.53 [0.36 ; 0.69] entre les différents lecteurs.

Il était de 0.45 [0.26 ; 0.61] entre les quatre lecteurs juniors et de 0,60 [0,41;0,76] entre les quatre lecteurs seniors. (**Figure 14**)

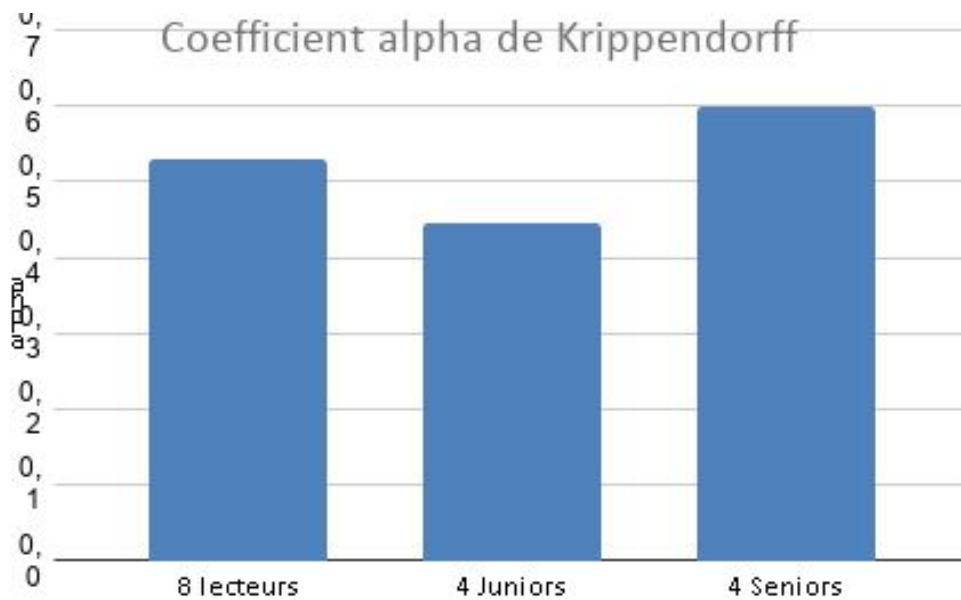


Figure 14: Coefficient alpha de Krippendorff

Concordance inter observateur évaluée par le calcul du coefficient alpha chez les 8 lecteurs, chez les 4 lecteurs juniors et les 4 lecteurs seniors.

Dans le but d'éliminer un biais lié à un lecteur "déviant", nous avons procédé à un nouveau calcul du coefficient alpha sur sept des huit lecteurs (en excluant tour à tour chaque lecteur). (**Figure 15**)

Les résultats étaient sensiblement identiques quel que soit le lecteur exclu, avec un coefficient alpha de Krippendorff s'étendant de 0.51 [0.34 ; 0.69] à 0.56 [0.39 ; 0.72].

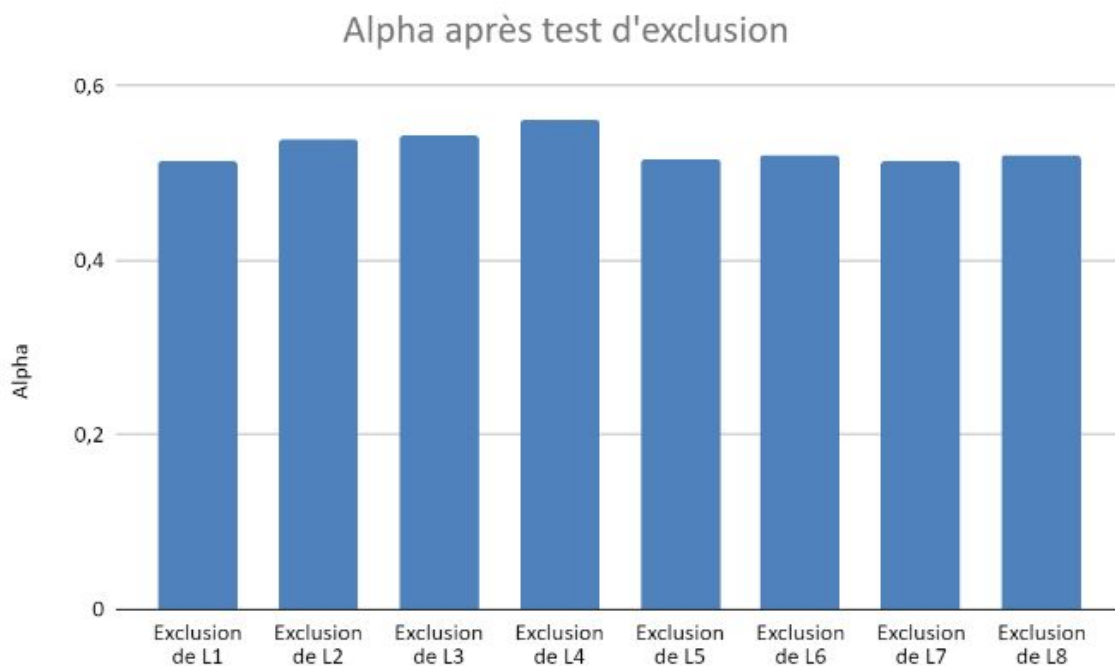


Figure 15: Coefficient alpha de Krippendorff avec test d'exclusion

Calcul du coefficient alpha avec exclusion tour à tour de chaque lecteur (L1 à L8), ne retrouvant pas de différence significative, traduisant l'absence de "lecteur déviant".

La concordance au sein de chaque binôme de lecteurs a été évaluée par le coefficient Kappa, pour chaque couple de lecteurs.

Ce coefficient Kappa s'étendait entre 0.30 [0.05 ; 0.55] et 0.67 [0.48 ; 0.86].

IV. Analyse de la simple lecture selon le type de lecteur

Chez les juniors, la médiane des sensibilités en lecture simple était de 0.90, la spécificité de 0.70, la VPP de 0.71 et la VPN de 0.88 (**Figures 16 à 18; Tableau 1**).

Chez les seniors, la médiane des sensibilités en lecture simple était de 0.94, la spécificité de 0.72, la VPP de 0.73 et la VPN de 0.93.

Sur l'ensemble des lecteurs, la moins bonne sensibilité était de 0.84 et la meilleure de 1.

La moins bonne spécificité était de 0.55, la meilleure de 0.83.

La moins bonne VPN était de 0.84, la meilleure de 1.

La moins bonne VPP était de 0.63, la meilleure de 0.92

Au total, en prenant en compte l'ensemble des lecteurs, la médiane des sensibilités en lecture simple était de 0.92, la spécificité de 0.70, la VPP de 0.72 et la VPN de 0.91.

Junior 1	Gold standard				Valeur
	A	B	Total		
				Sen	0,840
A	22	4	26	Spé	0,733
B	8	21	29	VPN	0,846
Total	30	25	55	VPP	0,724

Figure 16: réponses du lecteur junior 1 comparées au gold standard

Exemple du tableau de contingence reprenant les résultats des réponses du lecteur "Junior 1" ainsi que les performances diagnostiques correspondantes en terme de sensibilité, spécificité, VPP, VPN (avec intervalles de confiance).

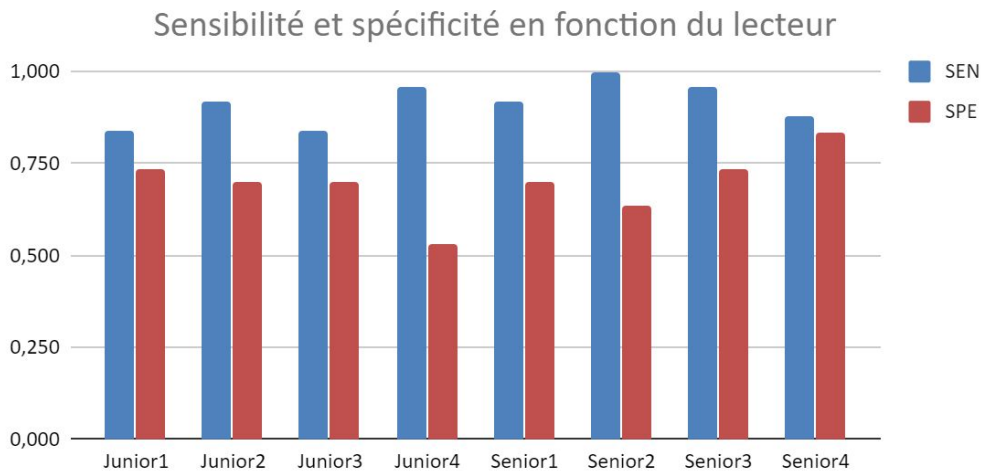


Figure 17: Performances diagnostiques en simple lecture
Sensibilité et spécificité en fonction du lecteur

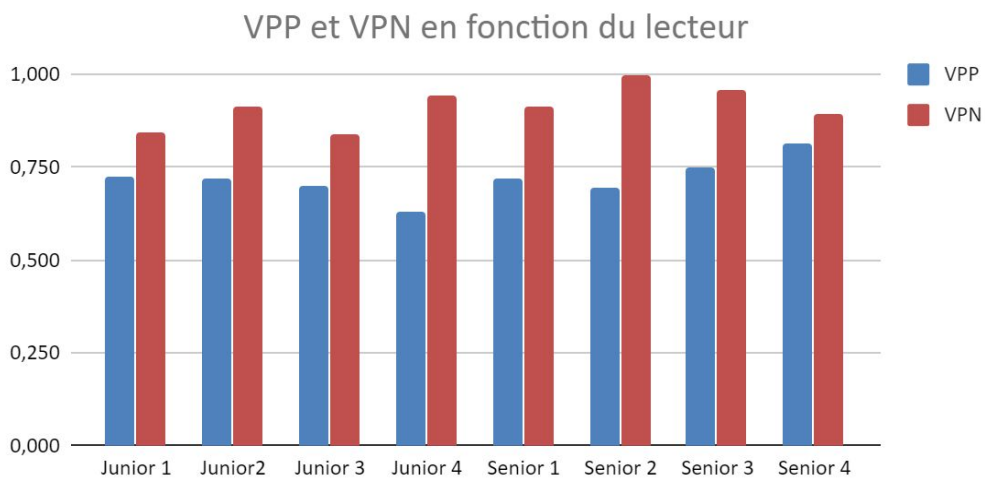


Figure 18: Performances diagnostiques en simple lecture
Valeur prédictive positive et valeur prédictive négative en fonction du lecteur

Tableau 1: Performances diagnostiques lecteur par lecteur en simple lecture

Tableau reprenant les valeurs exactes de la sensibilité, spécificité, valeur prédictive positive et valeur prédictive négative pour chaque lecteur en simple lecture

	SEN	SPE	VPP	VPN
Junior 1	0,840	0,733	0,724	0,846
Junior 2	0,920	0,700	0,719	0,913
Junior 3	0,840	0,700	0,700	0,840
Junior 4	0,960	0,533	0,632	0,941
Senior 1	0,920	0,700	0,719	0,913
Senior 2	1,000	0,633	0,694	1,000
Senior 3	0,960	0,733	0,750	0,957
Senior 4	0,880	0,833	0,815	0,893
Q1	0,87	0,68	0,70	0,88
Médiane	0,92	0,70	0,72	0,91
Q3	0,96	0,73	0,73	0,95

**v. Analyse détaillée de la double lecture selon le binôme
- hypothèse du worst case.**

En prenant en compte l'ensemble des couples possibles de lecteurs, étudiés selon l'hypothèse du "worst case", la médiane [Q1 ; Q3] de sensibilité en double lecture était de 1 [0.96 ; 1], la spécificité de 0.57 [0.50 ; 0.60], la VPP de 0.64 [0.63 ; 0.67] et la VPN de 1 [0.95 ; 1]. (**Tableau 2**)

Tableau 2: comparaison des performances des différents types de lecture avec l'hypothèse du worst case

Tableau comparant les performances diagnostiques en terme de sensibilité, spécificité, VPP et VPN selon le type de lecture (simple ou double) et selon le type de binôme (junior-junior, junior-senior ou senior-senior) avec l'hypothèse du worst case.

Worst Case				
Combinaison	SEN	SPE	VPP	VPN
simple lecture	92,00	70,00	71,88	91,30
J-J	98,00	48,00	62,00	97,00
J-S	100,00	55,00	64,48	100,00
S-S	100,00	56,67	65,79	100,00

Parmi les 3 couples de lecteurs, les meilleures performances étaient retrouvées pour les couples senior/senior avec une médiane (Q1 ; Q3) de sensibilité de 1 [0.97 ; 1], une spécificité de 0.57 [0.67 ; 0.63], une valeur prédictive positive de 0.66 [0.66 ; 0.68] et une valeur prédictive négative de 1 [0.96 ; 1]. (**Figure 21**)

Les performances des couples senior/juniors étaient presque identiques à celles des couples senior/senior.

Les performances des couples juniors/juniors étaient un peu moindres en termes de spécificité (0.48) et de VPP (0.62).

Performances diagnostiques selon le type de lecture et le type de binôme

Hypothèse du worst case

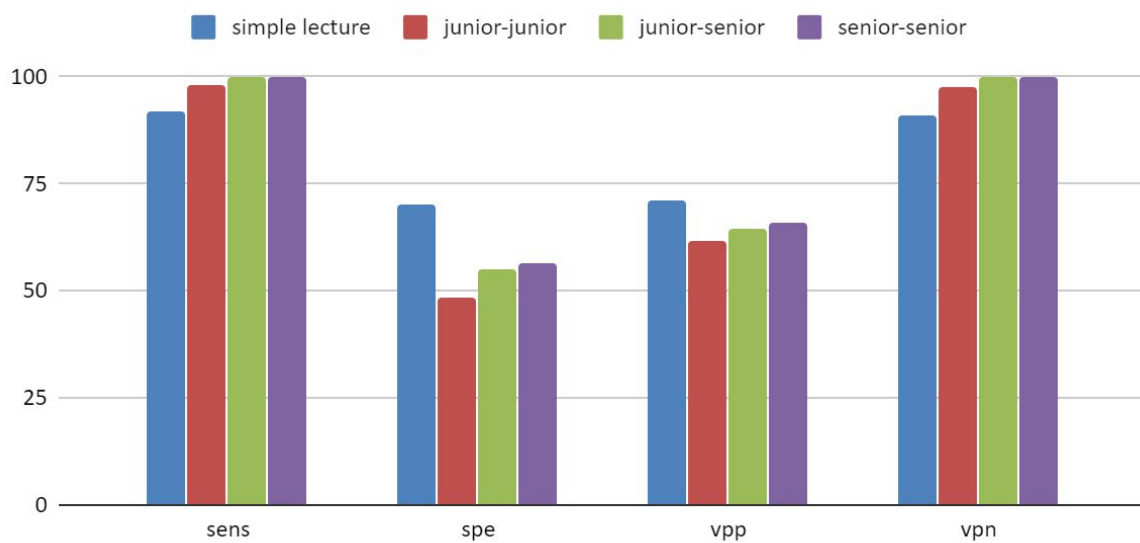


Figure 19: comparaison graphique des différents types de lecture avec l'hypothèse du worst case

Graphique comparant les performances diagnostiques selon le type de lecture (simple ou double) et selon le type de binôme avec l'hypothèse du worst case.

VI. Analyse détaillée de la double lecture selon le binôme

- hypothèse du best case

Avec l'hypothèse du "best case", la médiane [Q1 ; Q3] de sensibilité en double lecture était de 0.84 [0.8 ; 0.88], la spécificité de 0.86 [0.83;0.88], la VPP de 0.84 [0.79 ; 0.86] et la VPN de 0.86 [0.84 ; 0.90]. (**Figure 22**)

Tableau 3: comparaison des performances des différents types de lecture avec l'hypothèse du best case

Tableau comparant les performances diagnostiques en terme de sensibilité, spécificité, VPP et VPN selon le type de lecture (simple ou double) et selon le type de binôme (junior-junior, junior-senior ou senior-senior) avec l'hypothèse du best case.

Best Case				
Combinaison	SEN	SPE	VPP	VPN
simple lecture	92,00	70,00	71,88	91,30
J-J	82,00	85,00	81,69	84,48
J-S	84,00	85,00	83,05	86,67
S-S	88,00	88,33	86,86	90,16

Il existait de même que pour le "worst case" une tendance à des meilleurs performances pour les couples senior/senior (**Figure 23**) avec une sensibilité de 0.88 [0.88 ; 0.91], une spécificité de 0.90 [0.87 ; 0.90], une VPP de 0.88 [0.85 ; 0.88], une VPN de 0.90 [0.90 ; 0.90] que pour les couples junior/junior qui avaient une

sensibilité de 0.82 [0.77 ; 0.84], une spécificité de 0.85 [0.83 ; 0.87], une VPP de 0.84 [0.82 ; 0.87] et une VPN de 0.84 [0.82 ; 0.87].

Performances diagnostiques selon le type de lecture et le type de binôme

Scénario du best case

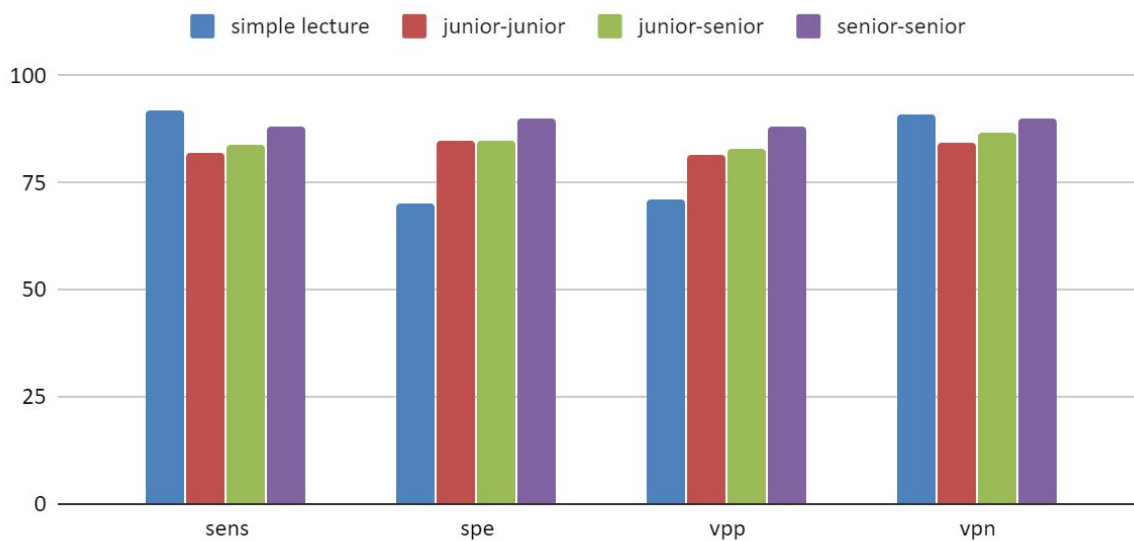


Figure 20: comparaison graphique des différents types de lecture avec l'hypothèse du best case

Graphique comparant les performances diagnostiques selon le type de lecture (simple ou double) et selon le type de binôme avec l'hypothèse du best case.

DISCUSSION

I. Design de l'étude

Il s'agit, à notre connaissance, de la première étude publiée traitant de l'intérêt de la double lecture appliquée à l'IRM bi paramétrique de prostate.

Nous avons fait le choix délibéré de sélectionner à la fois des lecteurs "juniors" dont l'expérience en imagerie prostatique était limitée à des connaissances théoriques avec une expérience très limitée, et des lecteurs "seniors" de grande expérience pour pouvoir analyser la meilleure configuration possible d'une double lecture.

Nous avons également décidé de proposer aux lecteurs un choix binaire (A ou B) un peu différent de la pratique réelle de l'interprétation en conditions réelle avec utilisation du standard PIRADS2.1, dans le but de se rapprocher le plus possible d'un examen de dépistage, permettant un triage rapide des patients.

Le choix de l'IRM bi paramétrique était cohérent avec cet objectif de triage, d'autant plus qu'une méta-analyse récente a prouvé sa supériorité face à l'IRM multiparamétrique, chez les patients naïfs de biopsie.

En outre, en recrutant des patients indemnes de toute biopsie prostatique ou traitement antérieur, nous proposons une approche de type dépistage, qui nécessite un examen simple, rapide et efficient.

Le rationnel du "Best" et "worst" case découle de la nécessité d'une réponse binaire.

Dans cette étude, nous avons choisi de simuler une double lecture en combinant deux évaluations indépendantes du bp-MRI, au lieu d'effectuer une nouvelle lecture de consensus avec un troisième juge. Nous pensons que cette méthode est plus simple, plus rapide et ne nécessite pas de temps pour l'examen des cas discordants ou une décision arbitraire. De plus, ces scénarios peuvent être automatisés. Enfin, la mise en pratique de ces deux hypothèses a déjà bénéficié d'une étude préliminaire (20) réalisée dans notre service, non publiée, qui montrait des résultats prometteurs.

II. Choix des lecteurs et taux de bonnes réponses

Il ne semble pas exister de différence significative en simple lecture entre les performances des lecteurs "seniors" comparativement aux lecteurs "juniors".

On note que l'utilisation homogène des critères PIRADS2.1 dans notre population, permet de limiter considérablement l'écart entre les lecteurs "seniors" et les lecteurs "juniors" en termes de discrimination lésionnelle.

En double lecture, les couples "seniors-juniors" sont très proches en termes de performance des couples "seniors-seniors".

Il paraît important, dans le cas où une double lecture est mise en place, de définir les caractéristiques du premier et du deuxième lecteur. Dans notre cas, il ne semble pas nécessaire de réunir deux lecteurs "experts" pour améliorer les performances de l'interprétation de l'examen, pour la recherche de cancer significatif.

III. Concordance entre les lecteurs

Nous avons étudié la concordance en les lecteurs à l'aide du test de Krippendorff pour la simple lecture car nous disposions de plusieurs lecteurs (8 en tout) pour interpréter un même examen.

En double lecture, le test Kappa était plus adapté car il s'appliquait à des binômes de lecteurs.

Le "test d'exclusion" nous a permis de vérifier qu'il n'existait pas un lecteur "déviant" qui aurait pu fausser nos résultats: cette analyse retrouve des valeurs stables du coefficient alpha de Krippendorff quel que soit le lecteur exclu.

La concordance mesurée entre les lecteurs par le test Kappa et test de Krippendorff était modérée dans notre étude, avec des valeurs pour le coefficient Kappa qui s'étendaient entre 0.30 et 0.67.

En effet, on considère souvent des valeurs de coefficient Kappa supérieures à 0.60 comme bonnes, les valeurs inférieures à 0.60 comme une concordance modérée et les valeurs inférieures à 0.40 comme faibles.

La concordance entre les séniors était un peu supérieure (0.60) que chez les juniors (0.45).

κ	Interpretation
< 0	Désaccord
0.0 — 0.20	Accord très faible
0.21 — 0.40	Accord faible
0.41 — 0.60	Accord modéré
0.61 — 0.80	Accord fort
0.81 — 1.00	Accord presque parfait

Figure 21: Interprétation du coefficient Kappa

Parmi les raisons pouvant expliquer de tels résultats, il est clair que la variabilité inter observateur du PI-RADS telle que décrite dans la littérature (14) joue un rôle important.

On note également les conditions d'interprétation sur logiciel, un peu différentes de l'interprétation en vacation (absence de zoom ou de fenêtrage possible) ont pu entraîner des difficultés d'interprétation pour certains dossiers.

Ces résultats peuvent également témoigner du caractère partiellement lecteur dépendant de l'IRM prostatique pour certains dossiers.

Quoi qu'il en soit, une faible concordance entre les lecteurs d'un même examen justifie d'autant plus une démarche de type "double lecture" comme dans notre étude.

IV. Durée d'interprétation des dossiers

Le temps d'interprétation était significativement différent entre les différents lecteurs, sans rôle évident de l'expérience sur une plus grande rapidité d'interprétation.

On notait un temps d'interprétation significativement très variable pour un même dossier entre les différents lecteurs, sans explication retrouvée.

Il n'existait pas de différence significative de temps d'interprétation entre les dossiers traitant de cancers significatifs et les IRM normales.

V. Simple lecture

On note que l'IRM prostatique présente, en simple lecture, une excellente sensibilité: 0.92 [0.87;0.96] ainsi qu'une excellente VPN: 0.91 [0.88;0.95].

La spécificité (0.70) et la VPP (0.72) sont parallèlement assez moyennes.

Ces données sont concordantes à celles de la littérature. (24)

VI. Choix du scénario "worst case" ou "best case" pour la double lecture

La réalisation d'une double lecture impose de faire un choix sur le traitement des discordances entre les deux lecteurs le cas échéant.

Dans le cas du "worst case", si l'un des lecteurs cote l'examen comme B (cible suspecte de cancer), alors la réponse du binôme de lecteurs sera forcément B.

Dans cette configuration, il faut que les deux lecteurs cotent l'examen comme normal pour que le résultat du binôme soit A.

La valeur prédictive négative et la sensibilité de l'examen sont donc améliorées par augmentation des vrais négatifs. (**Figure 22**)

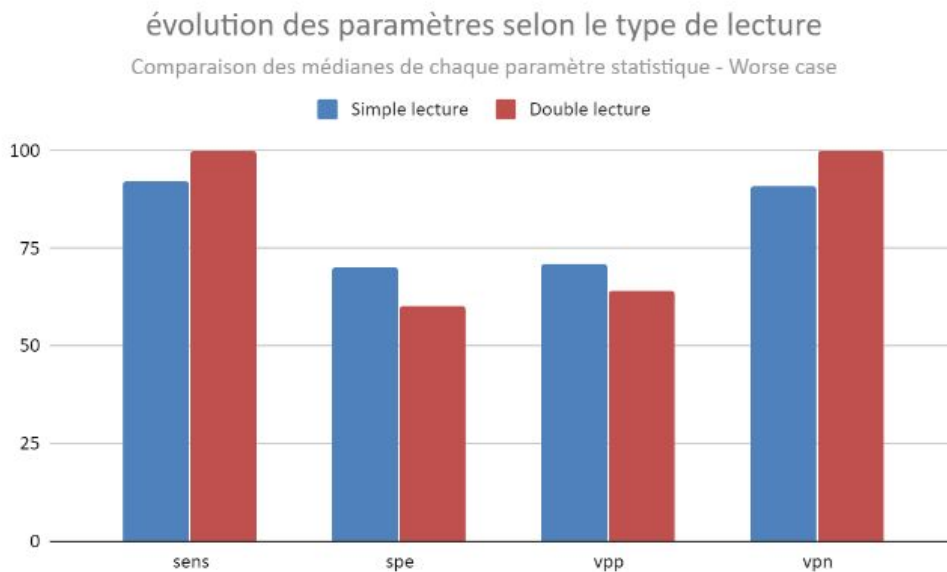


Figure 25: simple lecture contre double lecture avec le “worst” case scenario
On note dans ce scénario une amélioration significative de la sensibilité et de la VPN, au détriment de la spécificité et de la VPP.

À l'inverse, le choix du best case augmenterait le nombre de vrais positifs et donc la spécificité et la valeur prédictive positive du test (**Figure 26**).

Evolution des paramètres selon le type de lecture

Compaïson des médianes de chaque paramètre statistique - Best Case

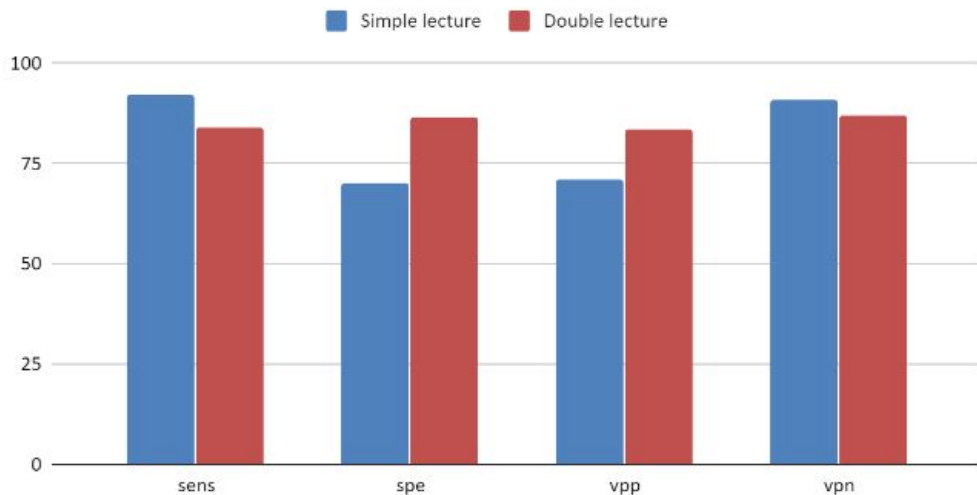


Figure 26: simple lecture contre double lecture avec le “best” case scenario
On note dans ce scénario une amélioration importante de la spécificité et de la VPP au détriment d’une diminution nette de la sensibilité et de la VPN.

En réflexion avec des urologues spécialisés en pathologie prostatique, nous avons défini que le paramètre statistique qui méritait le plus une amélioration était la valeur prédictive négative.

En effet, une excellente VPN pourrait permettre, dans certains cas, lorsque l'IRM est négative, de surseoir aux biopsies.

C’est donc l’hypothèse du “worst case” qui dans notre situation prend tout son intérêt.

VII. Performance de la double lecture

Dans notre étude, pour l'hypothèse du "worst case", la sensibilité et la valeur prédictive négative semblent se majorer lors d'une double lecture en aveugle, comparativement à une simple lecture (sensibilité et spécificité de 100% avec les couples sénior-sénior contre respectivement 92% et 91% en simple lecture).

La spécificité et la valeur prédictive positives sont quant à elles clairement diminuées en double lecture (respectivement 57% et 66% en double lecture avec les couples sénior-sénior contre 70% et 72% en simple lecture).

L'amélioration des performances diagnostiques de l'examen avec le modèle du "worst case", est cependant relativement peu marquée (+8 points de sensibilité et +9 points de VPN pour le "worst case" comparativement à la simple lecture)

Cela s'explique en partie par les valeurs déjà très élevées en termes de sensibilité, spécificité de la simple lecture, relevées dans notre étude.

Dans notre étude, 100% des cancers ont été dépistés en double lecture avec l'hypothèse du worst case, tandis que 8% auraient été manqués en simple lecture.

Inversement, lorsque les deux lecteurs étaient concordants pour dire que l'IRM était normale, ils avaient raison dans 100% des cas, contre 9% d'erreur en simple lecture.

L'application du scénario du "best case" améliore de manière significative la spécificité et la VPP qui passent respectivement à 88% (+18 points) et 88% (+14

points) en double lecture avec des lecteurs séniors comparativement à la simple lecture.

De telles performances en spécificité et VPP ne présentent cependant pas, dans le paradigme actuel, d'intérêt majeur.

VIII. Limites

Notre étude connaît plusieurs limites.

Premièrement, les conditions d'interprétation sur logiciel, plutôt que sur une console dédiée, ont pu induire des difficultés chez certains lecteurs, diminuant la concordance entre les lecteurs.

Deuxièmement, le design de l'étude faisait qu'il n'existait pas de test statistique adapté pour comparer la sensibilité et spécificité de la simple par rapport à la double lecture.

En effet, pour montrer avec certitude une telle différence, il aurait fallu faire analyser un grand nombre d'examens par deux lecteurs uniquement. On aurait alors pu étudier les différences de performances entre le lecteur "seul" et les performances du binôme. Ce type de protocole n'était malheureusement pas compatible avec notre besoin d'analyser dans le même temps le rôle de l'expérience dans l'interprétation en double lecture. Il semble néanmoins, à la vue des résultats retrouvés que la double lecture présente des performances supérieures à la simple lecture.

De même, il n'existait pas de manière fiable de comparer statistiquement les performances des lecteurs "juniors" des lecteurs "seniors", bien qu'il existait une tendance à des meilleures performances des binômes de "seniors" comparativement aux binômes de "juniors".

Le nombre de dossiers (55) était relativement faible, découlant du fait que chaque lecteur devait lire l'ensemble des dossiers. La raison en était que nous voulions être capables d'étudier la concordance entre les différents lecteurs.

On peut également apporter un biais potentiel lié au Gold Standard. En effet, si la preuve histologique est souvent considérée comme la référence dans la majorité des études, sa valeur dépend essentiellement de la fiabilité des prélèvements. On sait que la technique des biopsies systématiques de prostate risque de sous échantillonner certaines tumeurs, du fait des limitations en rapport avec la technique de prélèvement. Il existe donc possiblement parmi les prélèvements négatifs un certain nombre de faux négatifs pouvant biaiser nos mesures. Un recueil du devenir des patients deux ou trois ans plus tard aurait pu par exemple limiter considérablement ce biais.

Enfin, on notera que les performances diagnostiques retrouvées en termes de VPP et de VPN dépendent de la prévalence de cancer dans notre population. En effet, les patients inclus dans notre étude étaient représentatifs des hommes consultant pour une IRM avant une première série de biopsies dans notre centre, mais ne le sont pas forcément pour d'autres centres.

IX. Implications et perspectives

Il semble qu'une double lecture en aveugle de l'IRM de prostate, en utilisant le modèle du "worst case" améliore la valeur prédictive négative de l'examen. On peut se demander à partir de quel seuil de VPN l'IRM peut contribuer à trier les patients devant ou non bénéficier de biopsies prostatiques lorsqu'ils présentent une élévation significative du PSA.

Selon les résultats de notre étude, un patient respectant nos critères d'inclusion et dont l'IRM prostatique (interprétée en double lecture) serait négative, aurait moins de 4% de risque d'avoir tout de même un cancer significatif.

Les performances du binôme de lecteur semblent également dépendre de sa composition. Dans notre étude, même s'il n'existait pas de différence évidente en simple lecture entre les différents lecteurs, la double lecture retrouvait des performances un peu supérieures des binômes de type senior-senior et senior - junior, comparativement aux binômes composés uniquement de lecteurs juniors.

Pour faciliter encore la mise en place de cette double lecture, on pourrait aussi envisager, dans un futur proche, une double lecture automatisée, par un algorithme d'intelligence artificielle dont l'efficacité serait à démontrer.

L'attitude la plus prometteuse actuellement, dans les cas difficiles, semble être l'utilisation des algorithmes de calcul de risque de cancer (ex: ERSPC RC) qui permettent en incluant le taux de PSA, le volume prostatique, l'âge du patient, les

données du toucher rectal et de l'IRM, d'évaluer le risque pour le patient d'avoir un cancer de prostate et le risque d'avoir un cancer de prostate significatif.

Cette étude préliminaire préfigure une étude plus robuste où l'on pourrait étudier la double lecture avec le modèle du worst case, sur un plus grand nombre de dossiers, en ne faisant appel qu'à deux lecteurs.

CONCLUSION

Les performances diagnostiques de l'IRM prostatique semblent améliorées par une double lecture, par rapport à une lecture simple en utilisant le modèle du worst case.

En améliorant la valeur prédictive négative de l'examen, on pourrait diminuer de manière significative le recours aux biopsies prostatiques chez certains patients.

L'expérience des deux lecteurs ne semble pas avoir d'importance significative sur la performance globale du binôme.

La sensibilité de l'examen reste toujours trop faible pour faire de l'IRM prostatique un examen de dépistage, justifiant la poursuite d'une prise en charge multimodale.

Ces données méritent d'être validées lors d'une étude prospective sur un plus grand nombre de patients.

BIBLIOGRAPHIE

1. Vargas HA, Akin O, Franiel T, Goldman DA, Udo K, Touijer KA, et al. Normal central zone of the prostate and central zone involvement by prostate cancer: clinical and MR imaging implications. *Radiology*. mars 2012;262(3):894-902.
2. Haffner J, Potiron E, Bouyé S, Puech P, Leroy X, Lemaitre L, et al. Peripheral zone prostate cancers: location and intraprostatic patterns of spread at histopathology. *The Prostate*. 15 févr 2009;69(3):276-82.
3. L'anatomopathologie [Internet]. Prostanet. [cité 15 avr 2020]. Disponible sur: <https://prostanet.com/cancer-prostate/le-diagnostic/l-anatomopathologie/>
4. l'Urologie M de. Recommandations françaises du Comité de Cancérologie de l'AFU – Actualisation 2018–2020 : cancer de la prostate [Internet]. 2018 [cité 15 avr 2020]. Disponible sur: <https://www.urofrance.org/base-bibliographique/recommandations-francaises-du-comite-de-cancerologie-de-lafu-actualisation-10>
5. Schröder FH, Hugosson J, Roobol MJ, Tammela TLJ, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet Lond Engl*. 6 déc 2014;384(9959):2027-35.
6. l'Urologie M de. Chapitre 16 - Tumeurs de la prostate [Internet]. 2016 [cité 15 avr 2020]. Disponible sur: <https://www.urofrance.org/congres-et-formations/formation-initiale/referentiel-du-college/tumeurs-de-la-prostate.html>
7. Loeb S, Vellekoop A, Ahmed HU, Catto J, Emberton M, Nam R, et al. Systematic review of complications of prostate biopsy. *Eur Urol*. déc 2013;64(6):876-92.
8. Alabousi M, Salameh J-P, Gusenbauer K, Samoïlov L, Jafri A, Yu H, et al. Biparametric vs multiparametric prostate magnetic resonance imaging for the detection of prostate cancer in treatment-naïve patients: a diagnostic test accuracy systematic review and meta-analysis. *BJU Int*. 2019;124(2):209-20.
9. Rosenkrantz AB, Taneja SS. Radiologist, Be Aware: Ten Pitfalls That Confound the Interpretation of Multiparametric Prostate MRI. *Am J Roentgenol*. janv 2014;202(1):109-20.

10. Kasivisvanathan V, Rannikko AS, Borghi M, Panebianco V, Mynderse LA, Vaarala MH, et al. MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis. *N Engl J Med*. 10 mai 2018;378(19):1767-77.
11. Rouvière O, Puech P, Renard-Penna R, Claudon M, Roy C, Mège-Lechevallier F, et al. Use of prostate systematic and targeted biopsy on the basis of multiparametric MRI in biopsy-naïve patients (MRI-FIRST): a prospective, multicentre, paired diagnostic study. *Lancet Oncol*. 2019;20(1):100-9.
12. Radtke JP, Giganti F, Wiesenfarth M, Stabile A, Marengo J, Orczyk C, et al. Prediction of significant prostate cancer in biopsy-naïve men: Validation of a novel risk model combining MRI and clinical parameters and comparison to an ERSPC risk calculator and PI-RADS. *PloS One*. 2019;14(8):e0221350.
13. Kang Z, Min X, Weinreb J, Li Q, Feng Z, Wang L. Abbreviated Biparametric Versus Standard Multiparametric MRI for Diagnosis of Prostate Cancer: A Systematic Review and Meta-Analysis. *AJR Am J Roentgenol*. 2019;212(2):357-65.
14. Greer MD, Brown AM, Shih JH, Summers RM, Marko J, Law YM, et al. Accuracy and agreement of PI-RADSv2 for prostate cancer mpMRI: A multireader study. *J Magn Reson Imaging JMRI*. 2017;45(2):579-85.
15. Branger N, Maubon T, Traumann M, Thomassin-Piana J, Brandone N, Taix S, et al. Is negative multiparametric magnetic resonance imaging really able to exclude significant prostate cancer? The real-life experience. *BJU Int*. 2017;119(3):449-55.
16. Anastasiadis E, Charman SC, Arumainayagam N, Sohaib AS, Allen C, Freeman A, et al. What Burden of Prostate Cancer Can Radiologists Rule Out on Multiparametric Magnetic Resonance Imaging? A Sensitivity Analysis Based on Varying the Target Condition in Template Prostate Mapping Biopsies. *Urology*. sept 2015;86(3):544-51.
17. Ahmed HU, El-Shater Bosaily A, Brown LC, Gabe R, Kaplan R, Parmar MK, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *Lancet Lond Engl*. 25 2017;389(10071):815-22.
18. Padhani AR, Barentsz J, Villeirs G, Rosenkrantz AB, Margolis DJ, Turkbey B, et al. PI-RADS Steering Committee: The PI-RADS Multiparametric MRI and MRI-directed Biopsy Pathway. *Radiology*. 2019;292(2):464-74.
19. Knaapila J, Autio V, Jambor I, Ettala O, Verho J, Kiviniemi A, et al. Impact of biparametric prebiopsy prostate magnetic resonance imaging on the diagnostics of clinically significant prostate cancer in biopsy naïve men. *Scand J Urol*. févr 2020;54(1):7-13.

20. Double lecture de l'IRM multiparamétrique pour le diagnostic du cancer cliniquement significatif de la prostate. Doctorat de médecine / 2018 / Vincent-Jacques Anne-Sophie. Médecine. Radiodiagnostic et imagerie médicale
21. Kohestani K, Wallström J, Dehlfors N, Sponga OM, Månsson M, Josefsson A, Carlsson S, Hellström M, Hugosson J.Scand. Performance and inter-observer variability of prostate MRI (PI-RADS version 2) outside high-volume centres.J Urol. 2019 Oct;53(5):304-311.
22. Stabile A, Giganti F, Kasivisvanathan V, Giannarini G, Moore CM, Padhani AR, Panebianco V, Rosenkrantz AB, Salomon G et al. Factors Influencing Variability in the Performance of Multiparametric Magnetic Resonance Imaging in Detecting Clinically Significant Prostate Cancer: A Systematic Literature Review. Eur Urol Oncol. 2020 Mar 16. S2588-9311(20)30030-4.
23. Di Campli E, Delli Pizzi A, Seccia B, Cianci R, d'Annibale M, Colasante A, Cinalli S, Castellan P, Navarra R, Iantorno R et al. Diagnostic accuracy of biparametric vs multiparametric MRI in clinically significant prostate cancer: Comparison between readers with different experience. Eur J Radiol. 2018 Apr;101:17-23.
24. Boesen L, Nørgaard N, Løgager V, Balslev I, Bisbjerg R, Thestrup KC, Winther MD, Jakobsen H, Thomsen HS. Assessment of the Diagnostic Accuracy of Biparametric Magnetic Resonance Imaging for Prostate Cancer in Biopsy-Naive Men: The Biparametric MRI for Detection of Prostate Cancer (BIDOC) Study. JAMA Netw Open. 2018 Jun1;1(2):e180219.

AUTEUR : Nom : COTTEN **Prénom :** Louis

Date de Soutenance : 23/06/2020

Titre de la Thèse : Rôle de l'expérience et de la double lecture sur les performances de l'IRM biparamétrique dans le dépistage du cancer de la prostate cliniquement significatif chez les patients naïfs de biopsie

Thèse - Médecine - Lille 2020

Cadre de classement : Radiodiagnostic et Imagerie médicale

DES + spécialité : Radiodiagnostic et Imagerie médicale

Mots-clés :

Résumé :

Introduction:

Le cancer de la prostate est le cancer le plus fréquent chez l'homme de plus de 50 ans. Des outils sont nécessaires pour améliorer le diagnostic du cancer de prostate cliniquement significatif et éviter ainsi le risque de sur-traitement. L'IRM multiparamétrique de prostate est un examen dont l'intérêt n'est plus à démontrer dans cette indication. Les résultats restent cependant limités dans certains cas avec un recours encore large aux biopsies prostatiques. Notre étude cherche à déterminer si l'expérience des lecteurs et la double lecture en aveugle des IRMs avec un protocole biparamétrique influent sur les performances de l'examen chez les patients naïfs de biopsie.

Matériel et Méthodes:

Nous avons inclus 55 patients consécutifs ayant consulté pour une IRM prostatique suivie d'une première série de biopsies. Nous avons sélectionné 4 lecteurs "seniors" et 4 lecteurs "juniors" qui ont tous classé ces IRMs comme suspecte ou non de cancer significatif. Nous avons comparé les réponses des lecteurs au résultat histologique (gold standard). Nous avons ensuite créé des binômes de lecteurs dont les données d'interprétation ont été comparées au gold standard, en utilisant l'hypothèse du "best" et du "worst" case. Nous avons calculé les performances en termes de sensibilité, spécificité, VPP et VPN de chaque lecteur en simple lecture, puis de chaque binôme en double lecture.

Résultats:

Le coefficient alpha de Krippendorff était de 0.53 (0.36 ; 0.69) entre les différents lecteurs.

La médiane des sensibilités en lecture simple était de 0.92, la spécificité de 0.70, la VPP de 0.72 et la VPN de 0.91.

En prenant en compte l'ensemble des couples possibles de lecteurs, étudiés selon l'hypothèse du "worst case", la médiane de sensibilité en double lecture était de 1 [0.96 ; 1], la spécificité de 0.57, la VPP de 0.64 et la VPN de 1.

Avec l'hypothèse du "best case", la médiane de sensibilité en double lecture était de 0.84, la spécificité de 0.86, la VPP de 0.84 et la VPN de 0.86.

Conclusion:

Une double lecture des IRM de prostate en protocole bi-paramétrique avec l'hypothèse du "worst" case permet d'obtenir une excellente valeur prédictive négative, pouvant potentiellement éviter un recours aux biopsies prostatiques chez les patients dont l'IRM est négative. L'expérience des lecteurs ne semble pas modifier les performances globales de l'examen.

Composition du Jury :**Président :**

Monsieur le Professeur Arnauld VILLERS

Assesseurs :

Monsieur le Professeur Philippe PUECH

Monsieur le Professeur Xavier LEROY

Madame le Docteur Benoit RENARD

Monsieur Patrick DEVOS