

UNIVERSITÉ DE LILLE
FACULTÉ DE MÉDECINE HENRI WAREMBOURG
Année : (2022)

THÈSE POUR LE DIPLÔME D'ÉTAT
DE DOCTEUR EN MÉDECINE

**Reproductibilité inter-observateur et performance de l'évaluation par IRM de
l'extension extra-prostatique du cancer de la prostate : étude rétrospective
sur une série de 80 patients.**

Présentée et soutenue publiquement le 17 octobre 2022 à 16h
au Pôle Formation

par Rémi GOSSE

JURY

Président :

Monsieur le Professeur Philippe PUECH

Assesseurs :

Monsieur le Professeur Xavier LEROY

Monsieur le Docteur Jonathan OLIVIER

Monsieur le Docteur Louis COTTEN

Directeur de thèse :

Monsieur le Professeur Philippe PUECH

AVERTISSEMENT

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

LISTE DES ABRÉVIATIONS

CaP : Cancer de prostate

EEP : Extension extra-prostatique

BNV : Bandelette neuro-vasculaire

IRM : Imagerie par résonance magnétique

mm : Millimètre

LCC : Longueur de contact capsulaire

PIRADS : Prostate Imaging Reporting and Data System

VS : Vésicule séminale

VPP : Valeur prédictive positive

VPN : Valeur prédictive négative

RIO : Reproductibilité inter-observateur

Se : Sensibilité

Sp : Spécificité

CHU : Centre hospitalier universitaire

AUC : Aire sous la courbe

IC : Intervalle de confiance

ICC : Coefficient de corrélation inter-classe

RÉSUMÉ	1
I. INTRODUCTION	3
A. Cancer de prostate	3
1. Généralités	3
2. Importance de l'évaluation de l'extension extra-prostatique	3
3. IRM et extension extra-prostatique	4
4. Grade EPE et extension extra-prostatique	7
B. Objectifs de l'étude	10
II. MATÉRIEL ET MÉTHODES	10
A. Type d'étude	10
B. Base de données utilisée	10
C. Critères d'inclusion et d'exclusion	11
D. Protocole utilisé	12
E. Analyses statistiques	13
III. RÉSULTATS	14
A. Population de l'étude	14
B. Performances diagnostiques	15
C. Reproductibilité inter-observateur	19
IV. DISCUSSION	22
A. Performances diagnostiques du grade EPE et du score Likert	22
B. Reproductibilité inter-observateur des scores EPE et Likert	24
C. Reproductibilité inter-observateur de la mesure de la longueur de contact capsulaire	24
D. Limites et atouts de l'étude	26
E. Perspectives	27
V. CONCLUSION	29
VI. BIBLIOGRAPHIE	30
VII. ANNEXES	38

RÉSUMÉ

Contexte : Le cancer de prostate est le cancer le plus fréquent chez l'homme. La mise en évidence d'une extension extra-prostatique (EEP) en préopératoire est un élément pronostique majeur, modifiant la prise en charge thérapeutique. Plusieurs nomogrammes cliniques et scores IRM ont été proposés afin d'évaluer cette EEP, sans recommandations officielles. Parmi eux le score Likert et le grade EPE proposé par Mehralivand et al en 2019 permettant de standardiser la détection d'une EEP sur l'IRM.

Objectif : L'objectif principal était d'évaluer les performances diagnostiques du grade EPE par la relecture de 80 IRM par 3 lecteurs d'expérience différente, de les comparer à l'étude princeps et au score Likert et d'évaluer sa reproductibilité inter-observateur (RIO). L'objectif secondaire est d'évaluer la RIO de la mesure de la longueur de contact capsulaire (LCC), critère du grade EPE.

Matériel et méthode : une étude rétrospective, monocentrique a été réalisée au CHU de Lille, intégrant 80 patients ayant eu une IRM avant prostatectomie, relue par 3 radiologues d'expérience différente (1, 3 et > 20 ans) n'ayant pas connaissance du statut anatomopathologique. Les performances du grade EPE et du score Likert ont été mesurées pour chaque lecteur par l'aire sous la courbe (AUC), les sensibilité (Se) et spécificité (Sp) et les RIO par le coefficient alpha de Krippendorff.

Résultats : Notre étude montre des performances diagnostiques du grade EPE aussi bonnes que celles retrouvées dans l'étude de Mehralivand et celles du score Likert. En effet, il n'existe pas de différence significative ($p > 0,05$) des AUC du grade EPE et du score Likert pour chaque lecteur (AUC respectivement de 0,84, 0,75 et 0,79 pour le grade EPE et de 0,81, 0,69 et 0,78 pour le score Likert). La RIO de ces scores est modérée, avec une tendance à être meilleure pour le grade EPE et entre les lecteurs plus expérimentés, sans différence significative. La RIO de la mesure de LCC est quant à elle modérée à bonne, fonction des lecteurs.

Conclusion : Le grade EPE est un bon outil clinique, permettant de standardiser et de graduer le risque d'EEP en IRM de manière simple, tout en limitant l'impact de l'expérience du radiologue.

I. INTRODUCTION

A. Cancer de prostate

1. Généralités

Le cancer de la prostate est le 2^{ème} cancer le plus fréquent dans le monde avec 1.414.000 nouveaux cas et 375.000 décès en 2020 (1). En France, il reste le cancer le plus fréquent chez l'homme (50 000 nouveaux cas en France en 2015 d'après santé public France) et constitue donc un problème de santé public majeur.

Ces dernières années, le taux de mortalité est en diminution régulière grâce au dépistage précoce et à l'amélioration de sa prise en charge (2).

2. Importance de l'évaluation de l'extension extra-prostatique

La présence d'une extension extra-prostatique (EEP) est un élément pronostique majeur dans la prise en charge thérapeutique. En effet, elle est associée à un taux plus important de marges chirurgicales positives, de récurrence biologique, de maladie métastatique ainsi qu'à une survie inférieure après prostatectomie (3–6).

Son évaluation pré-thérapeutique par l'imagerie est donc essentielle, afin d'adapter au mieux la prise en charge et sélectionner le meilleur traitement (7).

En effet, le staging tumoral préopératoire peut modifier la technique chirurgicale ; en présence d'une forte suspicion d'EEP, il peut être décidé de réaliser une prostatectomie sans épargne des bandelettes neuro-vasculaires (BNV), situées à la partie postéro-latérale de la prostate, jouant un rôle dans l'érection et la continence

urinaire. Il a été prouvé que cette technique de non préservation des BNV augmente les chances d'obtenir des marges chirurgicales saines (8–11), au prix d'un taux de complications plus important à type de dysfonction érectile et d'incontinence urinaire (12–14).

Afin d'estimer ce risque d'EEP, des nomogrammes ont été développés, comme le score d'AMICO (Annexe 1) ou le score CAPRA (Annexe 2) intégrant des données cliniques, biologiques et histologiques importantes dans la graduation du risque tels que le toucher rectal, le taux de PSA, le score ISUP (15), ainsi que l'imagerie par résonance magnétique (IRM) (16).

3. IRM et extension extra-prostatique

L'IRM est un examen central dans la prise en charge des cancers de la prostate. Elle sert notamment à détecter des lésions suspectes afin de guider les biopsies devant une élévation du taux de PSA ou un toucher rectal suspect (15).

Elle joue également un rôle dans l'évaluation de l'EEP du CaP.

Plusieurs études ont démontré la supériorité de cet examen comparativement aux nomogrammes utilisés traditionnellement dans la littérature (17,18). Toutefois, l'intégration des données de l'IRM dans ces scores reste à ce jour le meilleur moyen d'estimer le risque d'EEP avant la chirurgie (19,20).

Selon une méta-analyse publiée en 2015 par De Rooij et al (21), intégrant 9796 patients, l'évaluation de l'extension extra-prostatique par l'IRM possède une forte spécificité (88%) mais une sensibilité faible et variable (61%) dépendante des centres et de l'expérience des lecteurs.

Dans la littérature de nombreuses caractéristiques IRM sont décrites comme évocatrices d'une EEP :

- La longueur du contact capsulaire (LCC) variant de 6 (22) à 20 millimètres (mm) (23) selon les études (24–28),
- Le renflement et l'irrégularité capsulaire,
- L'oblitération de l'angle recto-prostatique,
- L'asymétrie des bandelettes neuro-vasculaires,
- L'invasion de la graisse péri-prostatique,
- L'invasion des vésicules séminales (VS).

Celles-ci sont notamment présentes dans la version 2 du Prostate Imaging Reporting and Data System (PIRADS) (25), mais peu sont validées dans la littérature (22,23,29).

Aucune de ces caractéristiques n'est totalement spécifique d'une EEP. Les précédentes études utilisant une classification binaire EEP + ou EPP - selon l'observation de ces caractéristiques étaient donc soumises à une forte variabilité inter-observateur (29,30).

L'IRM ne permet donc pas de confirmer l'EEP, qui reste un diagnostic anatomopathologique, mais permet de graduer son risque.

Dans cette optique, plusieurs scores IRM ont été développés.

Le score de Likert côté de 1 à 5 (1 = EEP non présente, 2 = EEP probablement non présente, 3 = EEP incertaine, 4 = EEP probablement présente, 5 = EEP présente), donnant l'impression générale du radiologue sur la probabilité d'EEP. Celui-ci a démontré son efficacité dans plusieurs études. L'étude de Freifeld Y et al. (2019) montre une valeur prédictive positive (VPP) de 64% pour les Likert 4 et 5, de 90% pour

l'atteinte des VS, une valeur prédictive négative (VPN) de 87% pour les Likert 1 et 2 et de 93% pour l'atteinte des VS (31). Ce score semble toutefois présenter un accord inter-observateur modéré (32).

En 2012, la société européenne de radiologie uro-génitale (ESUR) propose un score (*Figure 1*) contenant de nombreuses caractéristiques IRM qualitatives (33). Une méta-analyse récente de Wei Li et al. (34) portant sur 10 études et 1698 patients a montré une sensibilité de 71% et une spécificité de 72%, associée à des résultats variables et à une reproductibilité inter-observateur (RIO) moyenne, expliquées notamment par l'utilisation de nombreuses caractéristiques IRM uniquement qualitatives (comme le bombement ou l'irrégularité capsulaire et l'asymétrie des BNV).

Criteria	Findings	Score
Extra-capsular extension	Abutment	1
	Irregularity	3
	Neurovascular bundle thickening	4
	Bulge, loss of capsule	4
	Measurable extra-capsular disease	5
Seminal vesicles	Expansion	1
	Low T2 signal	2
	Filling in of angle	3
	Enhancement and impeded diffusion	4
Distal sphincter	Adjacent tumour	3
	Effacement of low signal sphincter muscle	3
	Abnormal enhancement extending into sphincter	4
Bladder neck	Adjacent tumour	2
	Loss of low T2 signal in bladder muscle	3
	Abnormal enhancement extending into bladder neck	4

Figure 1 : Score ESUR (33)

Enfin, la LCC prise individuellement semble être un marqueur sensible d'EEP, mais le cut-off à utiliser comme marqueur d'EEP semble discordant dans la littérature (24). Toutefois dans la dernière version du PIRADS (2.1), le cut-off retenu est de 10 mm (35).

Ces trois différentes méthodes ne semblent pas standardiser et graduer de manière simple et reproductible le risque d'EEP.

C'est dans cette optique qu'en 2019 Mehralivand et al (36) réalisa une étude prospective, monocentrique, incluant 553 patients, évaluant de manière individuelle la relation entre les 6 caractéristiques IRM décrites dans la version 2 du PI-RADS et l'EEP (*Figure 2*), dans le but de créer un score de prédiction de l'EEP se voulant fiable et simple d'utilisation, le *grade EPE*.

Individual MRI features	
Curvilinear contact length greater than 1.5 cm	88/208 (42) [35, 49]
Capsular bulge	78/175 (45) [38, 52]
Obliteration of rectoprostatic angle	31/51 (61) [48, 75]
Neurovascular bundle asymmetry	19/25 (76) [59, 91]
EPE at MRI	37/56 (66) [53, 78]
Seminal vesicle invasion at MRI	12/15 (80) [57, 100]

Figure 2. Figure extraite de l'article de Mehralivand et al (36) décrivant les 6 caractéristiques IRM de l'EEP. Le numérateur correspondant au nombre d'EEP, le dénominateur au nombre de fois où la caractéristique a été retrouvée dans la série et la parenthèse au pourcentage.

4. Grade EPE et extension extra-prostatique

Ce grade EPE est subdivisé en trois :

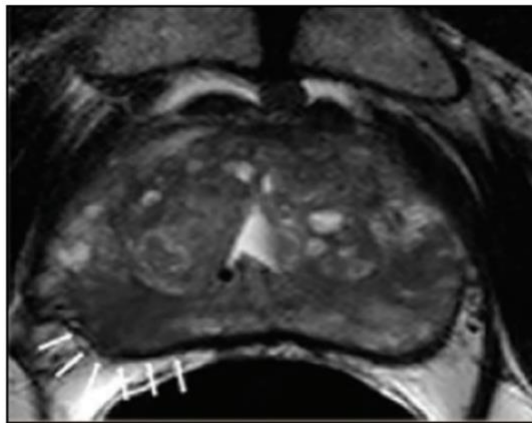
- **Grade 1** : présence d'une longueur tumorale de contact capsulaire ≥ 15 mm ou d'un renflement / une irrégularité capsulaire,

- **Grade 2** : présence d'une longueur tumorale de contact capsulaire ≥ 15 mm **et** d'un renflement / une irrégularité capsulaire,
- **Grade 3** : présence d'une extension extra-capsulaire franche ou d'une invasion des structures anatomiques adjacentes.

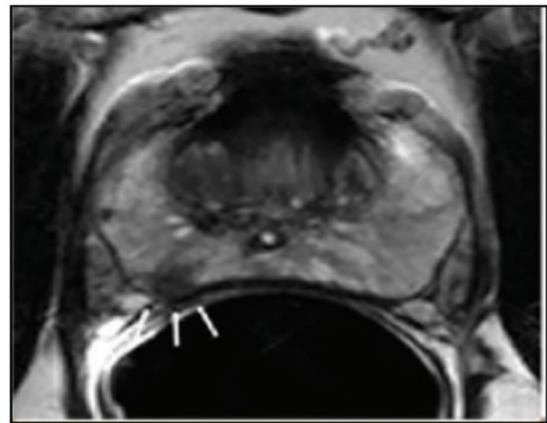
De ce score ont été retirés - dans une optique de simplification - l'asymétrie des BNV et l'atteinte des VS, jugées trop rares, et l'oblitération de l'angle recto-prostatique ne s'intéressant qu'aux tumeurs postérieures.

Le grade EPE permet une augmentation progressive du taux de détection de l'EEP comme le montre la *Figure 3* tirée de l'article de Mehralivand et al.

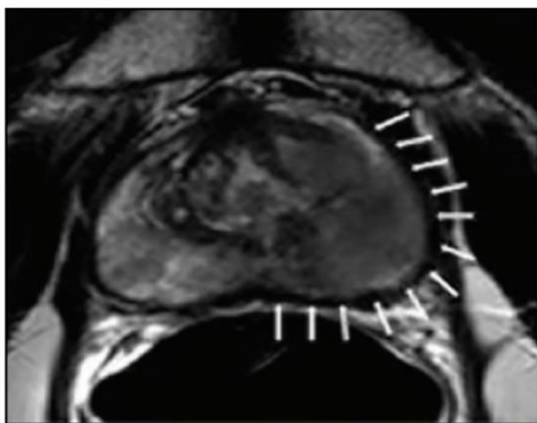
Cette étude retrouve en effet une graduation du risque d'EEP qui est de 24,3% pour le grade 1, 38% pour le grade 2 et de 66,1% pour le grade 3.



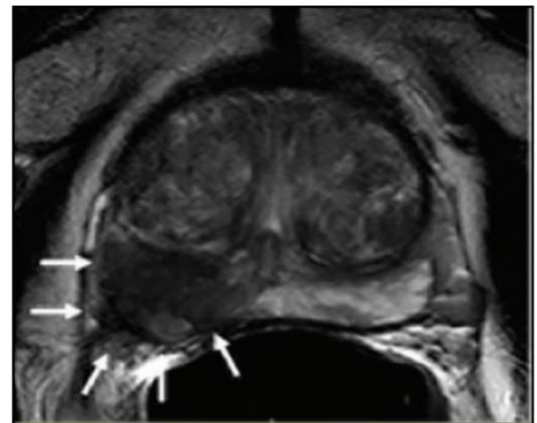
GRADE EPE 1
 Risque d'extension extra-prostatique **24,3%**
Longueur de contact curviligne



GRADE EPE 1
 Risque d'extension extra-prostatique **24,3%**
Irrégularité ou bombement capsulaire



GRADE EPE 2
 Risque d'extension extra-prostatique **38%**
Longueur de contact curviligne ET
irrégularité ou bombement capsulaire



GRADE EPE 3
 Risque d'extension extra-prostatique **66,1%**
Rupture de la capsule prostatique OU
invasion des structures adjacentes

Figure 3. Images IRM tirées de l'article de Mehralivand et al (36) Illustrant les différentes caractéristiques IRM du grade EPE.

Ce score semble être simple d'utilisation (n'utilisant que trois caractéristiques dont une quantitative), il standardise et gradue le risque d'EEP avec des performances diagnostiques semblant comparables à la littérature.

Toutefois, l'étude princeps de Mehralivand et al est monocentrique, basée sur l'interprétation d'un radiologue expérimenté sur une IRM à haut champ magnétique (3T).

B. Objectifs de l'étude

L'objectif principal de cette étude est d'évaluer les performances du grade EPE pour le diagnostic de l'extension extra-prostatique grâce à la relecture par 3 lecteurs d'expérience différente de 80 IRM réalisées à 1,5T en les comparant aux résultats de l'étude princeps et du score Likert, et d'évaluer sa reproductibilité inter-observateur.

L'objectif secondaire est d'évaluer la reproductibilité inter-observateur des mesures de la LCC, seule caractéristique quantitative du grade EPE.

II. MATÉRIEL ET MÉTHODES

A. Type d'étude

Il s'agit d'une étude observationnelle, rétrospective, monocentrique, réalisée au sein du service de radiologie génito-urinaire du centre hospitalier universitaire (CHU) de Lille

B. Base de données utilisée

La base de données IRM utilisée est celle d'une précédente thèse intitulée « Identification de biomarqueurs cliniques, d'imagerie et histologiques à risque d'extension extra-capsulaire du cancer prostatique (pT3) » qui recensait les données de 80 patients prostatectomisés au CHU de Lille entre mars 2017 et mai 2018.

C. Critères d'inclusion et d'exclusion

Les critères d'inclusion étaient la disponibilité des résultats du toucher rectal, du dosage du PSA, des images DICOM, des compte-rendus anatomopathologiques des biopsies et de la prostatectomie.

Les critères d'exclusion concernaient les patients présentant des données clinico-biologiques manquantes (dosage du PSA, toucher rectal et stade clinique) (n=2), l'absence de résultat anatomopathologique des biopsies (n=18) et des images DICOM indisponibles (n=52) venant principalement du fait que ces examens étaient réalisés dans un autre établissement. (Figure 4)

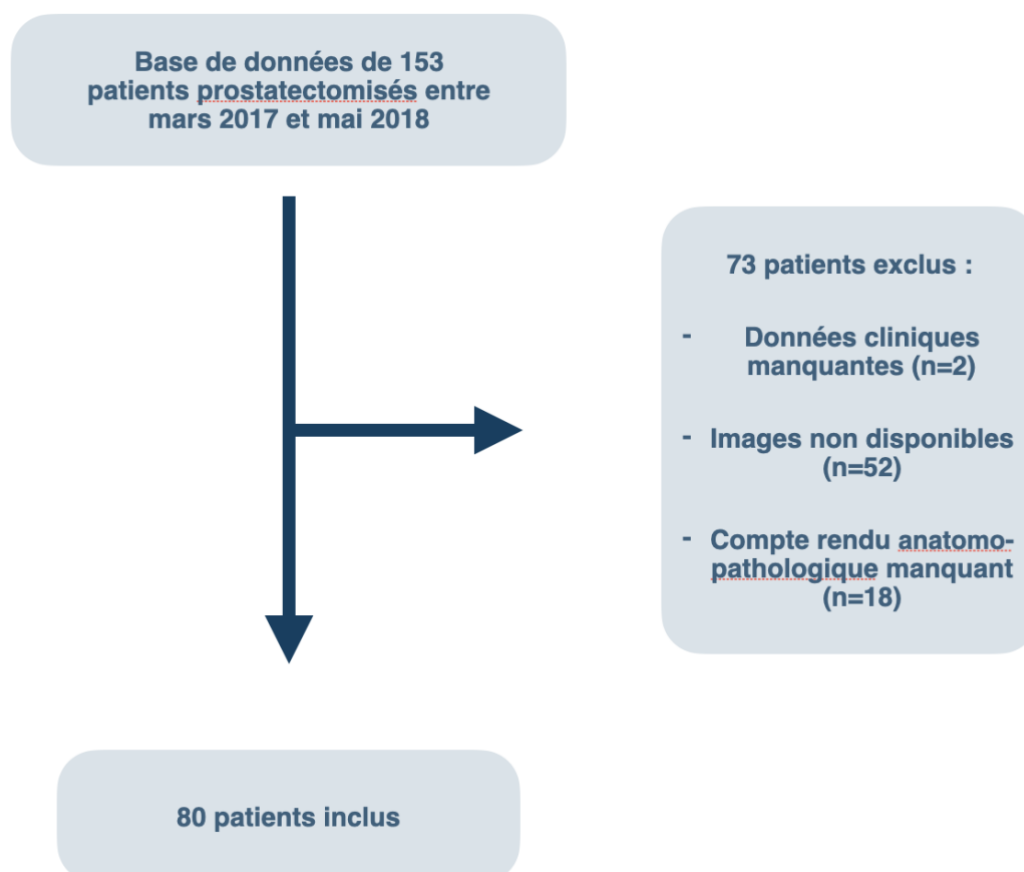


Figure 4. Flow-chart de l'étude

Au total, ont été inclus 153 patients et exclus 73 patients. Chaque patient a bénéficié d'une IRM multiparamétrique réalisée à 1,5T avec une antenne de surface.

D. Protocole utilisé

Chacune des 80 IRM a été relue de manière rétrospective individuellement par 3 radiologues d'expérience différente, un interne R1, un docteur junior R2 et un expert R3 présentant respectivement 1, 3 et 20 ans d'expertise en IRM de prostate, afin d'évaluer l'EEP.

Les paramètres suivants ont été évalués par chacun des observateurs sur l'ensemble des IRM : le score PIRADS 2.1, le grade EPE et la longueur curviligne de contact capsulaire tumoral en mm ainsi que l'évaluation de l'EEP par le score de LIKERT.

Chaque observateur disposait d'un repère pointant sur l'IRM la lésion à évaluer correspondant à la lésion principale décrite dans le compte rendu anatomopathologique, afin que tous puissent évaluer la même lésion.

Un résumé de l'article de Mehralivand et al (36) leur était fourni afin de respecter au mieux les critères du grade EPE.

L'outil d'évaluation utilisé (PACS) ne possédant pas d'outil de mesure à main levée, la longueur du contact curviligne a été estimée en traçant une série de 2 à 3 lignes droites de mesure.

Les caractéristiques cliniques, biologiques et résultats anatomopathologiques n'étaient pas connus par les observateurs.

L'existence d'une extension extra-prostatique était confirmée par l'examen anatomopathologique, qui est le gold standard.

E. Analyses statistiques

Les variables qualitatives ont été décrites en termes de fréquences et de pourcentages. Les variables quantitatives ont été décrites par la moyenne et l'écart type ou par la médiane et l'intervalle interquartile en cas de distribution non Gaussienne. La normalité des distributions a été vérifiée graphiquement et à l'aide du test de Shapiro-Wilk.

La concordance entre les 3 observateurs pour le score EPE a été évaluée par le coefficient alpha de Krippendorff et son intervalle de confiance à 95%. La concordance entre les observateurs 2 à 2 pour le score EPE a été évaluée par le coefficient Kappa de Cohen et son intervalle de confiance à 95%. La concordance pour le score Likert a été évaluée par la même méthode que celle du score EPE. La reproductibilité entre les observateurs 2 à 2 de la LCC a été évaluée à l'aide du coefficient de corrélation intra-classe et son intervalle de confiance à 95%, du graphique de Bland-Altman et du calcul du biais moyen et son intervalle de confiance à 95%.

Pour chaque observateur, les pouvoirs prédictifs des scores EPE et Likert ont été évalués par le calcul de l'aire sous la courbe ROC (AUC) et son intervalle de confiance à 95%. Les pouvoirs prédictifs des deux scores ont été comparés à l'aide du test de DeLong.

Les performances des différents seuils du score EPE (≥ 1 , ≥ 2 et ≥ 3) et les performances du score Likert EEP ≥ 3 pour prédire les extensions ont été évaluées par la sensibilité, la spécificité et les valeurs prédictives positives et négatives et leurs intervalles de confiance à 95%.

Le niveau de significativité a été fixé à 5%. Les analyses statistiques ont été effectuées à l'aide du logiciel SAS (SAS Institute version 9.4).

III. RÉSULTATS

A. Population de l'étude

L'âge moyen au moment de la prostatectomie était de 64,3 +/- 5,3 ans, le taux de PSA médian de 6 [5;7,6] (min :1,1 ; max :45) et l'ISUP médian de 2 [2 ;3].

Parmi les 80 patients, 35 patients (44%) présentaient une EEP et 45 (56%) n'en présentaient pas et le délai moyen entre l'IRM et la chirurgie était de 113±93 jours [min 26; max 743].

Les autres caractéristiques des patients sont résumées dans le *Tableau 1*.

Effectif	80
Age moyen (années)	64,3±5,3 (min 47; max 78)
PSA médian (ng/mL)	6 [5;7,6] (min1,1; max45)
Stade clinique	
cT1	60
cT2	20
cT3/T4	0
Délai moyen entre l'IRM et la chirurgie	113±93 jours [min 26; max 743]
Volume prostatique médian en IRM	41,5 [32,6;48,4] (min15; max145)
Taille médiane de la lésion principale en millimètres	17,5 (min 7 ; max45)
Extension extra-prostatique	
Positive	35 (44%)
Négative	45 (56%)
Marges de la prostatectomie	
Positives	24
Négatives	56
ISUP médian de la prostatectomie	2 [2;3]
1	16 (20%)
2	42 (52%)
3	13 (16%)
4	6 (8%)
5	3 (4%)

Tableau 1. Données démographiques.

B. Performances diagnostiques

Notre étude ne retrouvait pas de différence significative de la précision du grade EPE comparé au score de Likert pour l'évaluation de l'EEP. En effet les aires sous la courbe (AUC) n'étaient pas significativement différentes avec une AUC pour le grade EPE du

R1 à 0,84 contre une AUC pour le score de Likert à 0,81 (*Figure 5*), pour le R2 0,75 vs 0,69 (*Figure 6*) et pour le R3 0,79 vs 0,79 (*Figure 7*). L'ensemble de ces résultats sont répertoriés avec leur intervalle de confiance (IC) dans le *Tableau 2*.

	AUC grade EPE (intervalle de confiance 95%)	AUC score Likert (intervalle de confiance 95%)	<i>p value</i>
R1	0,84 (0,75 ;0,92)	0,81 (0,72 ;0,91)	<i>p=0,31</i>
R2	0,75 (0,65 ;0,86)	0,69 (0,58 ;0,80)	<i>p=0,12</i>
R3	0,79 (0,69 ;0,88)	0,79 (0,69 ;0,89)	<i>p=0,93</i>

Tableau 2. Comparaison de la précision pour l'évaluation de l'EEP du grade EPE et du score Likert pour chaque lecteur.

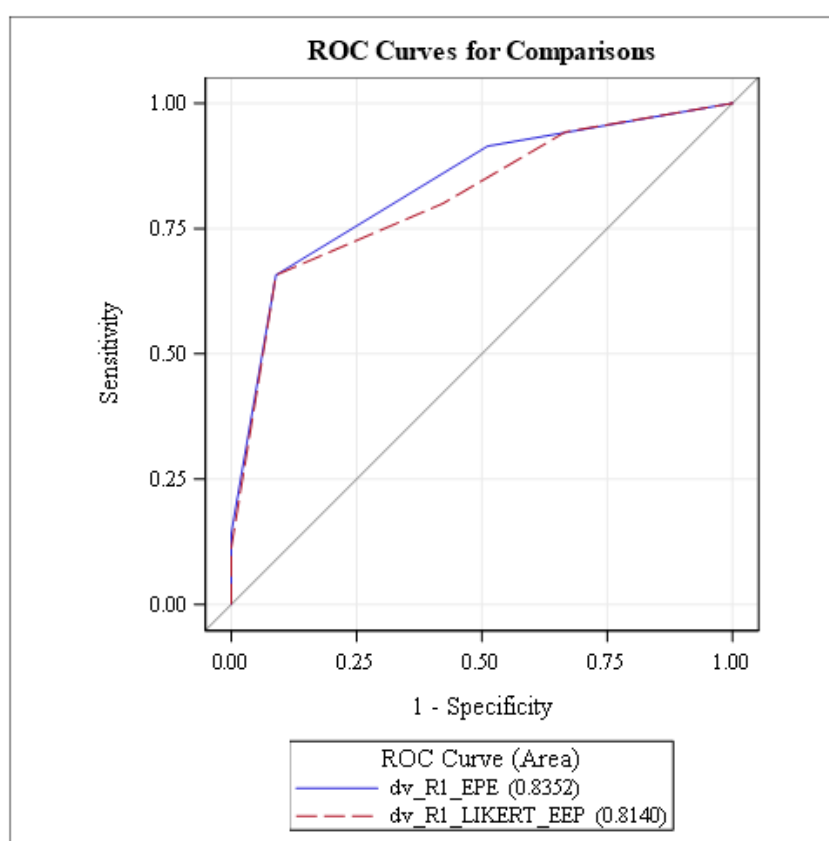


Figure 5. Courbes ROC et AUC du lecteur R1.

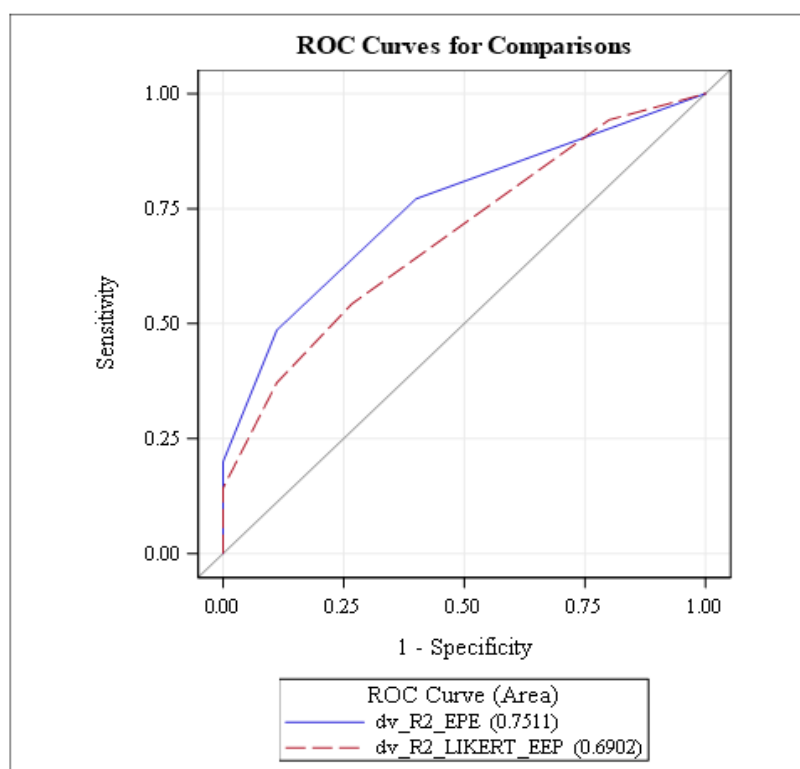


Figure 6. Courbes ROC et AUC du lecteur R2.

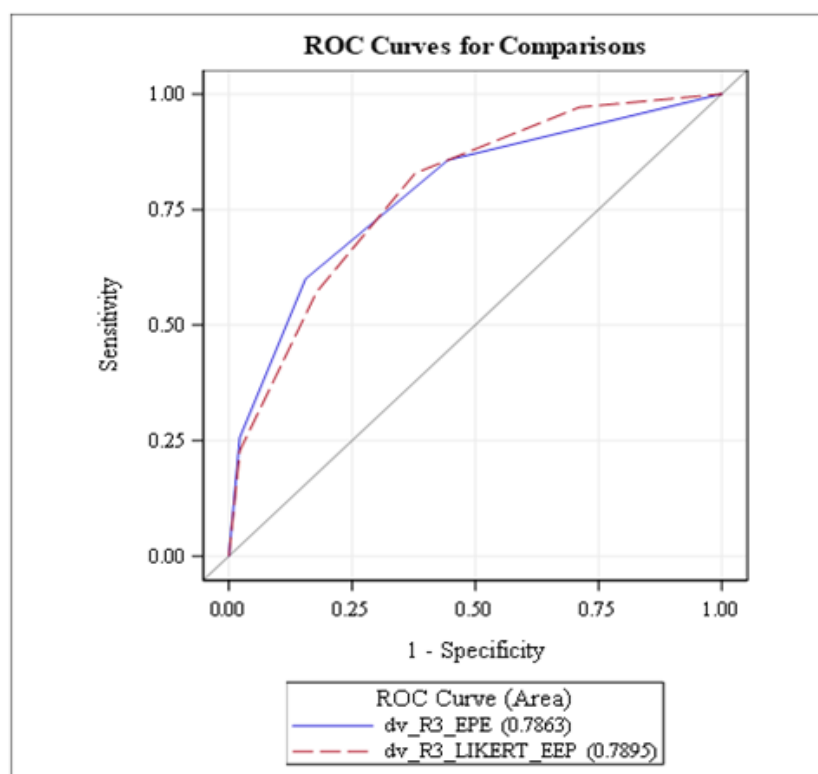


Figure 7. Courbes ROC et AUC du lecteur R3.

Notre étude évaluait pour chaque lecteur la sensibilité (Se), la spécificité (Sp), la valeur prédictive positive (VPP) et la valeur prédictive négative (VPN) du grade EPE ≥ 1 , ≥ 2 et ≥ 3 et du score Likert ≥ 3 (*Tableau 3*).

Pour un grade EPE ≥ 1 , la sensibilité était plus grande pour le lecteur le moins expérimenté (R1), de 91% contre 77 et 86% respectivement pour R2 et R3.

La spécificité était moins bonne pour le lecteur le moins expérimenté, de 48% (R1) contre 60 et 55% pour R2 et R3.

		R1	R2	R3	Mehralivand et al.
Sensibilité (borne inférieure;borne supérieure)	EPE ≥ 1	91% (77;98)	77% (60;90)	86% (70;95)	75 %
	EPE ≥ 2	66% (48;81)	48% (31;66)	60% (42;76)	61 %
	EPE ≥ 3	14% (5;30)	20% (8;37)	25% (12;43)	30,0 %
Spécificité (borne inférieure;borne supérieure)	EPE ≥ 1	48% (34;64)	60% (44;74)	56% (40;70)	68 %
	EPE ≥ 2	91% (79;98)	89% (76;96)	84% (71;94)	81 %
	EPE ≥ 3	100% (92;100)	100% (92;100)	98% (88;100)	96 %
VPP (borne inférieure;borne supérieure)	EPE ≥ 1	58% (44;71)	60% (44;74)	60% (45;74)	41 %
	EPE ≥ 2	85% (66;96)	77% (55;92)	75% (55;89)	48 %
	EPE ≥ 3	100% (48;100)	100% (59;100)	90% (55;99)	66 %
VPN (borne inférieure;borne supérieure)	EPE ≥ 1	88% (69;97)	77% (60;90)	83% (65;94)	90,0 %
	EPE ≥ 2	77% (64;88)	69% (55;80)	73% (59;84)	88 %
	EPE ≥ 3	60% (48;71)	62% (50;73)	63% (50;74)	82 %

Tableau 3. Performances diagnostiques du grade EPE et de l'étude de Mehralivand et al (2019).

Pour un score Likert ≥ 3 , R1 et R3 avaient une meilleure sensibilité que le lecteur R2, respectivement de 80 et 83% contre 54%. Cela au prix d'une spécificité moins bonne, de 58 et 62%, alors que R2 avait une meilleure spécificité de 73% (*Tableau 4*).

Les Se et Sp moyennes des 3 lecteurs pour le grade EPE ≥ 1 étaient de 85% et 55%.

Les Se et Sp moyennes des 3 lecteurs pour le score Likert ≥ 3 étaient de 72% et 64%.

	R1 Likert ≥ 3	R2 Likert ≥ 3	R3 Likert ≥ 3
Sensibilité (borne inférieure;borne supérieure)	80% (63 ;92)	54% (37 ;71)	83% (66 ;93)
Spécificité (borne inférieure;borne supérieure)	58% (42 ;72)	73% (58 ;85)	62% (47 ;76)
VPP (borne inférieure;borne supérieure)	60% (44 ;74)	61% (42 ;78)	63% (48 ;77)
VPN (borne inférieure;borne supérieure)	79% (61 ;91)	67% (52 ;80)	82% (65 ;93)

Tableau 4. Performances diagnostiques du score Likert ≥ 3 .

C. Reproductibilité inter-observateur

La reproductibilité inter-observateur entre les 3 radiologues pour l'évaluation du grade EPE était modérée avec un coefficient alpha de Krippendorff (α) = 0,69 (0,62 ; 0,77) et faible pour le score Likert, α = 0,54 (0,46 ; 0,62). La RIO du grade EPE a tendance à être meilleure que celle du score Likert, sans différence significative, les IC se chevauchant peu.

La reproductibilité a également été comparée entre chaque observateur, afin d'évaluer l'impact de l'expérience du lecteur sur l'évaluation du grade EPE et du score Likert. Celle-ci est bonne pour le grade EPE entre les 2 radiologues les plus expérimentés (R2 vs R3 : $k=0,63$), modérée entre R1 et R2 ($k=0,56$) et entre R1 et R3 ($k=0,57$).

Elle est modérée pour le score Likert entre R1 et R3 ($k=0,43$) et R2 et R3 ($k=0,41$). Elle est cependant faible entre les deux lecteurs les moins expérimentés (R1 vs R2 : $k=0,36$).

Il n'existe pas de différence significative pour la RIO entre le grade EPE et le score Likert pour chaque groupe de lecteur pris 2 à 2, les intervalles de confiance des kappa se chevauchant.

L'ensemble de ces résultats sont répertoriés dans le *Tableau 5* avec les IC.

Reproductibilité inter-observateur : grade EPE	Kappa (Intervalle de confiance 95%)
R1 vs R2	0,56 (0,44 ;0,67)
R1 vs R3	0,57 (0,44 ;0,70)
R2 vs R3	0,63 (0,50 ;0,78)
Reproductibilité inter-observateur : score LIKERT	Kappa (Intervalle de confiance 95%)
R1 vs R2	0,36 (0,22 ;0,50)
R1 vs R3	0,43 (0,30 ;0,57)
R2 vs R3	0,41 (0,27 ;0,55)

Tableau 5. Reproductibilité 2 à 2 du grade EPE et du score Likert. k de 0 à 0,20 : accord léger ; k de 0,21 à 0,40 : accord faible ; k de 0,41 à 0,60 : accord modéré ; k de 0,61 à 0,80 : accord bon ; k de 0,81 à 1 : accord très bon à presque parfait.

Notre étude a évalué la reproductibilité 2 à 2 des mesures du contact capsulaire, variable quantitative mesurée en mm.

Celle-ci est modérée entre les lecteurs 1 et 2 avec un coefficient de corrélation inter-classe (ICC=0,66) et les lecteurs 2 et 3 (ICC = 0,57) et bonne entre les lecteurs 1 et 3 (ICC = 0,78) (*Tableau 6*).

	ICC (borne inférieure ; borne supérieure)	Biais moyen (borne inférieure ; borne supérieure)
R1 vs R2	0,66 (0,51 ;0,77)	0,85 (-0,75 ;2,45)
R1 vs R3	0,78 (0,67 ;0,85)	0,90 (-0,30 ;2,10)
R2 vs R3	0,57 (0,39 ;0,70)	0,05 (-1,66 ;1,76)

Tableau 6. Reproductibilité 2 à 2 des mesures du contact capsulaire. ICC compris entre 0 et 1. <0,5 : pas de concordance ; entre 0,50 et 0,75 : concordance modérée ; entre 0,75 et 0,9 : concordance bonne ; >0,9 : concordance excellente.

IV. DISCUSSION

L'objectif principal de cette étude était d'évaluer les performances diagnostiques du grade EPE en IRM, pour l'évaluation de l'EEP, grâce à la relecture des imageries par 3 lecteurs d'expérience différente en imagerie prostatique (R1 : 1 an d'expérience ; R2 : 3 ans d'expérience ; R3 : 20 ans d'expérience). Les résultats obtenus ont été comparés à l'étude princeps de Mehralivand et al et au score Likert largement utilisé en pratique clinique pour l'évaluation de l'EEP.

A. Performances diagnostiques du grade EPE et du score Likert

Les performances diagnostiques du grade EPE, évaluées par l'AUC, étaient de 0,84 pour le lecteur R1 contre 0,75 pour R2 et 0,79 pour R3, valeurs qui semblent en accord avec l'étude de Mehralivand et al (AUC = 0,77) et avec d'autres études récentes retrouvant des AUC variant de 0,75 à 0,81 (37–39). Dans cette étude, l'expérience du lecteur ne semblait pas améliorer les performances diagnostiques.

En revanche les résultats pour le score Likert semblaient plus variables avec des AUC pour les lecteurs R1, R2 et R3 respectivement de 0,81, 0,69 et 0,79, résultats qui semblent en accord avec la littérature récente pour R1 et R3. En effet, Reisæter et al. (37) retrouve en 2020 des AUC de 0,76 et 0,77 pour 2 lecteurs de même expérience (>10 ans), Park et al (38) en 2020 des AUC de 0,78 et 0,79 pour des lecteurs d'expérience différente (de 3 et > 15 ans). Les résultats de R2 semblaient quant à eux discordant avec une performance diagnostique moindre.

Les AUC du grade EPE étaient supérieures aux AUC du score de Likert pour R1 et R2, suggérant une meilleure performance diagnostique et étaient égales (AUC = 0,79) pour le lecteur R3 plus expérimenté. Toutefois, il n'existait pas de différence significative ($p>0,05$) pour chaque lecteur, ce qui est en accord avec la littérature (37,38,40).

Notre étude retrouvait, pour le grade EPE ≥ 1 , une meilleure sensibilité au prix d'une moins bonne spécificité comparativement à l'étude de Mehralivand et al. En effet, les sensibilités étaient de 91%, 77% et 85% respectivement pour R1, R2 et R3 contre 75% dans l'étude de Mehralivand et les spécificités de 48%, 60% et 55% contre 68%.

Nos résultats vont cependant dans le sens de la méta-analyse récente de Li et al (41) portant sur 4 études et 1294 patients qui retrouve pour le grade EPE une sensibilité de 82% et une spécificité de 63%.

Une importante méta-analyse de 2015 (21) s'intéressant aux performances d'une lecture simple de l'IRM, sans utilisation de grade ou de score, retrouvait une sensibilité de 57% et une spécificité de 91% pour la détection de l'extension extra-prostatique. Dans la littérature, le grade EPE semble apporter une meilleure sensibilité, au prix d'une moins bonne spécificité, pour des performances diagnostiques globalement identiques, tout en limitant la variabilité liée aux centres et à l'expérience des lecteurs (38,39).

Deux études récentes, de Park et al (2020) et Asfuroglu et al (2022) (38,40), ont évalué les performances diagnostiques et la RIO entre les 4 scores IRM évaluant la présence d'une EEP les plus étudiés dans la littérature, que sont le score ESUR, le score Likert, le grade EPE et la LCC ≥ 10 mm. Les performances diagnostiques et les RIO étaient bonnes pour tous ces scores dans ces 2 études.

B. Reproductibilité inter-observateur des scores EPE et Likert

Dans notre étude, la RIO entre les 3 radiologues était modérée pour le grade EPE et faible pour le score Likert. Elle tendait donc à être meilleure pour le grade EPE, sans différence significative.

La RIO avait également été évalué entre chaque radiologue, 2 à 2, pour le grade EPE puis pour le score Likert, afin de juger l'impact de l'expérience des lecteurs sur celle-ci. Là encore il n'existait pas de différence significative pour chaque groupe de lecteur 2 à 2, même si l'on remarquait une tendance à une meilleure RIO pour le grade EPE pour chaque groupe de radiologue avec une RIO bonne entre les 2 radiologues les plus expérimentés ($k=0,63$) et modérée entre les autres (voir *Tableau 5*).

Dans la littérature, 4 études récentes ont évaluées la RIO du score EPE (37–40). Deux d'entre elles retrouvaient une bonne RIO chez des lecteurs d'expérience différente, la troisième une bonne RIO chez des lecteurs d'expérience semblable. Par ailleurs, celle de Reisæter et al. (37) retrouve une RIO modérée chez deux radiologues expérimentés.

C. Reproductibilité inter-observateur de la mesure de la longueur de contact capsulaire

L'objectif secondaire de notre étude était d'évaluer la RIO de la mesure de la LCC, qui en pratique clinique nous paraissait parfois varier de manière importante entre les lecteurs.

Notre étude retrouve une RIO modérée (pour R1 vs R2 et R2 vs R3) à bonne (R1 vs R3) en fonction des lecteurs.

Dans la littérature et notamment dans une récente méta-analyse de Kim et al. (24) la RIO de la mesure de LCC paraît meilleure que celle de notre étude avec des RIO variant de bonnes à très bonnes. Cette différence pourrait être expliquée par les niveaux d'expérience différents des lecteurs ou par notre méthode de mesure. En effet, l'outil d'évaluation utilisé ne possédant pas d'outil de mesure à main levée, la longueur du contact curviligne avait été estimée en traçant une série de 2 à 3 lignes droites, technique pouvant induire un biais. La mesure de la LCC était réalisée de façon hétérogène dans les études de cette méta analyse, de manière curviligne ou linéaire (*Figure 8*).

Une étude rétrospective récente de Eurboonyanun et al (42) portant sur 95 patients s'est intéressée à la méthode de mesure de la LCC (curviligne VS linéaire) et a conclu à une meilleure évaluation de l'EEP avec la mesure linéaire de LCC (sensibilité de 76% et spécificité de 72% pour un cut off de LCC ≥ 15 mm) comparé à la mesure curviligne. Il serait intéressant d'évaluer si cette donnée se confirme sur une plus grande cohorte de patients.

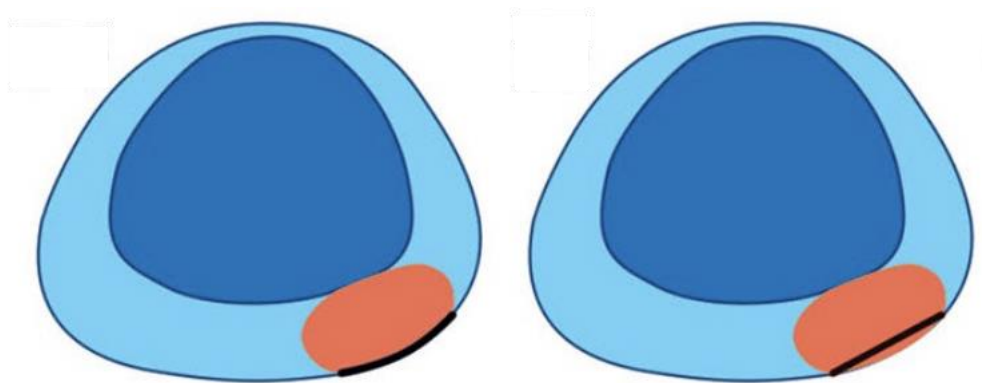


Figure 8. Schéma tiré de l'article de Eurboonyanun et al (42) qui montre la différence entre la mesure curviligne de la longueur du contact capsulaire à gauche vs la mesure linéaire de la longueur du contact capsulaire à droite.

D. Limites et atouts de l'étude

Notre étude possède plusieurs limites. Premièrement, il s'agit d'une étude rétrospective, mono-centrique, comme la plupart des études évaluant le grade EPE depuis sa proposition en 2019.

Deuxièmement, notre étude ne distingue pas les performances diagnostiques du grade EPE pour l'évaluation de l'EEP focale de l'évaluation de l'EEP étendue.

Hors il a été montré dans la littérature que l'EEP focale (vs étendue) sur la pièce de prostatectomie ne montrent que rarement des récives post prostatectomies (43,44),

Il aurait donc été intéressant d'évaluer les performances du grade EPE dans la détection des EPE non focales comme Park et al (38). D'autant plus que Mehralivand et al intègre la LCC reconnue comme sensible dans la détection de ces EEP focales (22,23). Troisièmement, le grade EPE a été comparé au score Likert, utilisé en pratique clinique, il aurait été intéressant de réaliser des comparaisons avec les autres scores IRM et cliniques. Et enfin nos mesures n'étaient pas parfaitement curviligne du fait du

système de post-traitement des images n'intégrant pas d'outils de mesure à main levée ce qui constitue un biais dans nos objectifs d'étude.

La force de notre étude est d'être la première à évaluer le grade EPE sur une IRM à 1,5 T avec une antenne de surface et de comparer les performances diagnostiques et la RIO chez 3 lecteurs d'expérience différente.

E. Perspectives

Le score IRM idéal pour l'évaluation de l'EEP dans le CaP devrait comporter à la fois des caractéristiques IRM quantitatives telles que la LCC, la mesure de l'ADC, ou le volume tumoral ainsi que des caractéristiques qualitatives telles que celles décrites dans le PIRADS, chaque caractéristique possédant une définition précise. Une utilisation simple en pratique clinique ainsi qu'un bon accord inter-lecteur seraient des points essentiels.

Le grade EPE a le mérite de standardiser la graduation du risque d'EEP de manière simple en utilisant seulement 3 caractéristiques IRM, ce qui le rend facile à enseigner et apprendre.

La présence d'une invasion extra-capsulaire franche ou l'atteinte d'un organe adjacent ont l'avantage d'être facilement reconnus par les radiologues.

Toutefois ce grade reste subjectif. Cela étant principalement lié à la définition de l'irrégularité et du bombement capsulaire qui reste floue et donc dépendante de l'expérience du radiologue.

Le cut-off de LCC de 15 mm utilisé comme marqueur d'EEP par Mehralivand et al n'est pas expliqué, même si quelques études ont retrouvé cette mesure (27) et

notamment une récente de 2021 de Eurboonyanun et al (42). Ce cut-off reste cependant débattu dans la littérature. Il pourrait aussi être différent suivant la localisation antérieure ou postérieure de la tumeur, ceci restant à prouver (45).

Son intégration est intéressante dans ce score, la LLC ayant montré une bonne Se dans la détection de l'EEP, qu'elle soit étendue ou focale. De plus, il existe une bonne reproductibilité inter-observateur de sa mesure.

Dans la littérature, d'autres critères IRM quantitatifs prédictifs d'EEP non pris en compte dans le grade EPE ont été analysés, notamment le volume tumoral de la lésion index. Deux études ont évalué ses performances en tant que prédicteur d'EEP. Lim et al (46) obtiennent une Se de 78,4% et une spécificité de 73,5% dans l'évaluation de l'EEP, pour les lésions de plus de 2,1 mm³. Rud et al (47) retrouvent une Se de 81% en zone périphérique, 86% en zone de transition et une spécificité respectivement de 60% et 50% pour les lésions de plus de 0,9 mm³.

La mesure du coefficient apparent de diffusion (ADC) est aussi un marqueur utile pour prédire l'agressivité du CaP (48). Il a pour les lésions de la zone périphérique une relation inverse avec le score de Gleason (49,50). Elle a démontré pour les tumeurs avec un ADC bas ($\leq$ ou égal à 89×10^{-5} mm²/s) environ 50% de probabilité d'EEP (47,51,52). Une méta-analyse de 2019 de Bai et al (53) incluant 18 études, retrouve une sensibilité et une spécificité de 80,5% et 69,1% dans la détection de l'EEP. Toutefois, ces études n'utilisaient pas les mêmes cut-offs dans la mesure de l'ADC, rendant son utilisation difficile en pratique clinique.

Une étude rétrospective plus récente (2021) de Koichi et al (54) s'est intéressée spécifiquement aux tumeurs présentant en IRM un contact capsulaire, sans irrégularité, bombement, effacement ou franchissement capsulaire. Sur les 93 patients

de l'étude, 58 possédaient ces caractéristiques IRM. Pour ces tumeurs, l'association LCC et mesure de l'ADC retrouvaient de bonnes performances diagnostiques dans l'évaluation de l'EEP (AUC = 0,82) avec une Se de 73,7% et une Sp de 84,6%. Cette combinaison paraît intéressante pour l'évaluation de l'EEP. Elle associe en effet seulement 2 caractéristiques IRM, quantitatives, souvent moins sujettes à de mauvaise RIO.

Ces caractéristiques pourraient être intéressantes à intégrer dans un prochain score d'évaluation de l'EEP.

V. CONCLUSION

Le grade EPE est un bon outil clinique, permettant de standardiser et de graduer le risque d'extension extra prostatique en IRM de manière simple, tout en limitant l'impact de l'expérience du radiologue.

Notre étude montre que le grade EPE a des performances diagnostiques au moins équivalentes à l'étude princeps de Mehralivand et al et au score Likert, non dépendantes de l'expérience du radiologue.

La reproductibilité inter-observateur de ce grade est modérée et à tendance à être meilleure que la reproductibilité inter-observateur du score Likert.

Ce score est malgré tout perfectible, certaines caractéristiques restant mal définies ou subjectives.

VI. BIBLIOGRAPHIE

1. Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. *Int J Cancer*. 5 avr 2021;
2. Le cancer de la prostate - Les cancers les plus fréquents [Internet]. [cité 6 oct 2022]. Disponible sur: <https://www.e-cancer.fr/Professionnels-de-sante/Les-chiffres-du-cancer-en-France/Epidemiologie-des-cancers/Les-cancers-les-plus-frequents/Cancer-de-la-prostate>
3. Hull GW, Rabbani F, Abbas F, Wheeler TM, Kattan MW, Scardino PT. Cancer control with radical prostatectomy alone in 1,000 consecutive patients. *J Urol*. févr 2002;167(2 Pt 1):528-34.
4. Mikel Hubanks J, Boorjian SA, Frank I, Gettman MT, Houston Thompson R, Rangel LJ, et al. The presence of extracapsular extension is associated with an increased risk of death from prostate cancer after radical prostatectomy for patients with seminal vesicle invasion and negative lymph nodes. *Urol Oncol*. janv 2014;32(1):26.e1-7.
5. Tollefson MK, Karnes RJ, Rangel LJ, Bergstralh EJ, Boorjian SA. The impact of clinical stage on prostate cancer survival following radical prostatectomy. *J Urol*. mai 2013;189(5):1707-12.
6. Wheeler TM, Dillioglulugil O, Kattan MW, Arakawa A, Soh S, Suyama K, et al. Clinical and pathological significance of the level and extent of capsular invasion in clinical stage T1-2 prostate cancer. *Hum Pathol*. août 1998;29(8):856-62.
7. Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur Urol*. avr 2017;71(4):618-29.

8. Roethke MC, Lichy MP, Kniess M, Werner MK, Claussen CD, Stenzl A, et al. Accuracy of preoperative endorectal MRI in predicting extracapsular extension and influence on neurovascular bundle sparing in radical prostatectomy. *World J Urol.* oct 2013;31(5):1111-6.
9. Partin AW, Borland RN, Epstein JI, Brendler CB. Influence of wide excision of the neurovascular bundle(s) on prognosis in men with clinically localized prostate cancer with established capsular penetration. *J Urol.* juill 1993;150(1):142-6; discussion 146-148.
10. Cheng L, Slezak J, Bergstralh EJ, Myers RP, Zincke H, Bostwick DG. Preoperative prediction of surgical margin status in patients with prostate cancer treated by radical prostatectomy. *J Clin Oncol Off J Am Soc Clin Oncol.* août 2000;18(15):2862-8.
11. Park BH, Jeon HG, Jeong BC, Seo SI, Lee HM, Choi HY, et al. Influence of magnetic resonance imaging in the decision to preserve or resect neurovascular bundles at robotic assisted laparoscopic radical prostatectomy. *J Urol.* juill 2014;192(1):82-8.
12. Loeb S, Smith ND, Roehl KA, Catalona WJ. Intermediate-term potency, continence, and survival outcomes of radical prostatectomy for clinically high-risk or locally advanced prostate cancer. *Urology.* juin 2007;69(6):1170-5.
13. Kwon T, Lee C, Jung J, Kim CS. Neurovascular bundle size measured on 3.0-T magnetic resonance imaging is associated with the recovery of erectile function after robot-assisted radical prostatectomy. *Urol Oncol.* sept 2017;35(9):542.e11-542.e17.
14. Avulova S, Zhao Z, Lee D, Huang LC, Koyama T, Hoffman KE, et al. The Effect of Nerve Sparing Status on Sexual and Urinary Function: 3-Year Results from the CEASAR Study. *J Urol.* mai 2018;199(5):1202-9.

15. Rozet F, Hennequin C, Beauval JB, Beuzeboc P, Cormier L, Fromont-Hankard G, et al. RETRACTED: Recommandations françaises du Comité de Cancérologie de l'AFU – Actualisation 2018–2020 : cancer de la prostate French ccAFU guidelines – Update 2018–2020: Prostate cancer. *Progres En Urol J Assoc Francaise Urol Soc Francaise Urol*. nov 2018;28(12S):S79-130.
16. Kozikowski M, Malewski W, Michalak W, Dobruch J. Clinical utility of MRI in the decision-making process before radical prostatectomy: Systematic review and meta-analysis. *PloS One*. 2019;14(1):e0210194.
17. Gupta RT, Faridi KF, Singh AA, Passoni NM, Garcia-Reyes K, Madden JF, et al. Comparing 3-T multiparametric MRI and the Partin tables to predict organ-confined prostate cancer after radical prostatectomy. *Urol Oncol*. nov 2014;32(8):1292-9.
18. Augustin H, Fritz GA, Ehammer T, Auprich M, Pummer K. Accuracy of 3-Tesla magnetic resonance imaging for the staging of prostate cancer in comparison to the Partin tables. *Acta Radiol Stockh Swed* 1987. juin 2009;50(5):562-9.
19. Morlacco A, Sharma V, Viers BR, Rangel LJ, Carlson RE, Froemming AT, et al. The Incremental Role of Magnetic Resonance Imaging for Prostate Cancer Staging before Radical Prostatectomy. *Eur Urol*. mai 2017;71(5):701-4.
20. Feng TS, Sharif-Afshar AR, Wu J, Li Q, Luthringer D, Saouaf R, et al. Multiparametric MRI Improves Accuracy of Clinical Nomograms for Predicting Extracapsular Extension of Prostate Cancer. *Urology*. août 2015;86(2):332-7.
21. de Rooij M, Hamoen EHJ, Witjes JA, Barentsz JO, Rovers MM. Accuracy of Magnetic Resonance Imaging for Local Staging of Prostate Cancer: A Diagnostic Meta-analysis. *Eur Urol*. août 2016;70(2):233-45.
22. Rosenkrantz AB, Shanbhogue AK, Wang A, Kong MX, Babb JS, Taneja SS. Length of capsular contact for diagnosing extraprostatic extension on prostate MRI:

Assessment at an optimal threshold. J Magn Reson Imaging JMRI. avr 2016;43(4):990-7.

23. Baco E, Rud E, Vlatkovic L, Svindland A, Eggesbø HB, Hung AJ, et al. Predictive value of magnetic resonance imaging determined tumor contact length for extracapsular extension of prostate cancer. J Urol. févr 2015;193(2):466-72.

24. Kim TH, Woo S, Han S, Suh CH, Ghafoor S, Hricak H, et al. The Diagnostic Performance of the Length of Tumor Capsular Contact on MRI for Detecting Prostate Cancer Extraprostatic Extension: A Systematic Review and Meta-Analysis. Korean J Radiol. juin 2020;21(6):684-94.

25. Kongnyuy M, Sidana A, George AK, Muthigi A, Iyer A, Ho R, et al. Tumor contact with prostate capsule on magnetic resonance imaging: A potential biomarker for staging and prognosis. Urol Oncol. janv 2017;35(1):30.e1-30.e8.

26. Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. Eur Urol. janv 2016;69(1):16-40.

27. Woo S, Kim SY, Cho JY, Kim SH. Length of capsular contact on prostate MRI as a predictor of extracapsular extension: which is the most optimal sequence? Acta Radiol Stockh Swed 1987. avr 2017;58(4):489-97.

28. Bakir B, Onay A, Vural M, Armutlu A, Yıldız SÖ, Esen T. Can Extraprostatic Extension Be Predicted by Tumor-Capsule Contact Length in Prostate Cancer? Relationship With International Society of Urological Pathology Grade Groups. AJR Am J Roentgenol. mars 2020;214(3):588-96.

29. Wibmer A, Vargas HA, Donahue TF, Zheng J, Moskowitz C, Eastham J, et al. Diagnosis of Extracapsular Extension of Prostate Cancer on Prostate MRI: Impact of Second-Opinion Readings by Subspecialized Genitourinary Oncologic Radiologists.

AJR Am J Roentgenol. juill 2015;205(1):W73-78.

30. Mullerad M, Hricak H, Wang L, Chen HN, Kattan MW, Scardino PT. Prostate cancer: detection of extracapsular extension by genitourinary and general body radiologists at MR imaging. Radiology. juill 2004;232(1):140-6.
31. Freifeld Y, Diaz de Leon A, Xi Y, Pedrosa I, Roehrborn CG, Lotan Y, et al. Diagnostic Performance of Prospectively Assigned Likert Scale Scores to Determine Extraprostatic Extension and Seminal Vesicle Invasion With Multiparametric MRI of the Prostate. AJR Am J Roentgenol. mars 2019;212(3):576-81.
32. Costa DN, Passoni NM, Leyendecker JR, de Leon AD, Lotan Y, Roehrborn CG, et al. Diagnostic Utility of a Likert Scale Versus Qualitative Descriptors and Length of Capsular Contact for Determining Extraprostatic Tumor Extension at Multiparametric Prostate MRI. AJR Am J Roentgenol. mai 2018;210(5):1066-72.
33. Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, et al. ESUR prostate MR guidelines 2012. Eur Radiol. avr 2012;22(4):746-57.
34. Li W, Dong A, Hong G, Shang W, Shen X. Diagnostic performance of ESUR scoring system for extraprostatic prostate cancer extension: A meta-analysis. Eur J Radiol. oct 2021;143:109896.
35. PIRADS-V2-1.pdf [Internet]. [cité 2 oct 2022]. Disponible sur: <https://www.acr.org/-/media/ACR/Files/RADS/Pi-RADS/PIRADS-V2-1.pdf?la=en>
36. Mehralivand S, Shih JH, Harmon S, Smith C, Bloom J, Czarniecki M, et al. A Grading System for the Assessment of Risk of Extraprostatic Extension of Prostate Cancer at Multiparametric MRI. Radiology. mars 2019;290(3):709-19.
37. Reisæter LAR, Halvorsen OJ, Beisland C, Honoré A, Gravdal K, Losnegård A, et al. Assessing Extraprostatic Extension with Multiparametric MRI of the Prostate: Mehralivand Extraprostatic Extension Grade or Extraprostatic Extension Likert Scale?

Radiol Imaging Cancer. janv 2020;2(1):e190071.

38. Park KJ, Kim MH, Kim JK. Extraprostatic Tumor Extension: Comparison of Preoperative Multiparametric MRI Criteria and Histopathologic Correlation after Radical Prostatectomy. Radiology. juill 2020;296(1):87-95.

39. Xu L, Zhang G, Zhang X, Bai X, Yan W, Xiao Y, et al. External Validation of the Extraprostatic Extension Grade on MRI and Its Incremental Value to Clinical Models for Assessing Extraprostatic Cancer. Front Oncol. 2021;11:655093.

40. Asfuroğlu U, Asfuroğlu BB, Özer H, Gönül İl, Tokgöz N, İnan MA, et al. Which one is better for predicting extraprostatic extension on multiparametric MRI: ESUR score, Likert scale, tumor contact length, or EPE grade? Eur J Radiol. avr 2022;149:110228.

41. Li W, Shang W, Lu F, Sun Y, Tian J, Wu Y, et al. Diagnostic Performance of Extraprostatic Extension Grading System for Detection of Extraprostatic Extension in Prostate Cancer: A Diagnostic Systematic Review and Meta-Analysis. Front Oncol. 2021;11:792120.

42. Eurboonyanun K, Pisuchpen N, O'Shea A, Lahoud RM, Atre ID, Harisinghani M. The absolute tumor-capsule contact length in the diagnosis of extraprostatic extension of prostate cancer. Abdom Radiol N Y. août 2021;46(8):4014-24.

43. Bloch BN, Genega EM, Costa DN, Pedrosa I, Smith MP, Kressel HY, et al. Prediction of prostate cancer extracapsular extension with high spatial resolution dynamic contrast-enhanced 3-T MRI. Eur Radiol. oct 2012;22(10):2201-10.

44. Lee S, Kim KB, Jo JK, Ho JN, Oh JJ, Jeong SJ, et al. Prognostic Value of Focal Positive Surgical Margins After Radical Prostatectomy. Clin Genitourin Cancer. août 2016;14(4):e313-319.

45. Matsumoto K, Akita H, Narita K, Hashiguchi A, Takamatsu K, Takeda T, et al.

Prediction of extraprostatic extension by MRI tumor contact length: difference between anterior and posterior prostate cancer. *Prostate Cancer Prostatic Dis.* déc 2019;22(4):539-45.

46. Lim C, Flood TA, Hakim SW, Shabana WM, Quon JS, El-Khodary M, et al. Evaluation of apparent diffusion coefficient and MR volumetry as independent associative factors for extra-prostatic extension (EPE) in prostatic carcinoma. *J Magn Reson Imaging JMRI.* mars 2016;43(3):726-36.

47. Rud E, Diep L, Baco E. A prospective study evaluating indirect MRI-signs for the prediction of extraprostatic disease in patients with prostate cancer: tumor volume, tumor contact length and tumor apparent diffusion coefficient. *World J Urol.* avr 2018;36(4):629-37.

48. Wu X, Reinikainen P, Vanhanen A, Kapanen M, Vierikko T, Ryymin P, et al. Correlation between apparent diffusion coefficient value on diffusion-weighted MR imaging and Gleason score in prostate cancer. *Diagn Interv Imaging.* janv 2017;98(1):63-71.

49. Hambrock T, Somford DM, Huisman HJ, van Oort IM, Witjes JA, Hulsbergen-van de Kaa CA, et al. Relationship between apparent diffusion coefficients at 3.0-T MR imaging and Gleason grade in peripheral zone prostate cancer. *Radiology.* mai 2011;259(2):453-61.

50. Itou Y, Nakanishi K, Narumi Y, Nishizawa Y, Tsukuma H. Clinical utility of apparent diffusion coefficient (ADC) values in patients with prostate cancer: can ADC values contribute to assess the aggressiveness of prostate cancer? *J Magn Reson Imaging JMRI.* janv 2011;33(1):167-72.

51. Granja MF, Pedraza CM, Flórez DC, Romero JA, Palau MA, Aguirre DA. Predicting extracapsular involvement in prostate cancer through the tumor contact

length and the apparent diffusion coefficient. Radiologia. août 2017;59(4):313-20.

52. Woo S, Cho JY, Kim SY, Kim SH. Extracapsular extension in prostate cancer: added value of diffusion-weighted MRI in patients with equivocal findings on T2-weighted imaging. AJR Am J Roentgenol. févr 2015;204(2):W168-175.

53. Bai K, Sun Y, Li W, Zhang L. Apparent diffusion coefficient in extraprostatic extension of prostate cancer: a systematic review and diagnostic meta-analysis. Cancer Manag Res. 11 avr 2019;11:3125-37.

54. Ito K, Chiba E, Oyama-Manabe N, Washino S, Manabe O, Miyagawa T, et al. Combining the Tumor Contact Length and Apparent Diffusion Coefficient Better Predicts Extraprostatic Extension of Prostate Cancer with Capsular Abutment: A 3 Tesla MR Imaging Study. Magn Reson Med Sci MRMS Off J Jpn Soc Magn Reson Med. 1 juill 2022;21(3):477-84.

VII. ANNEXES

Annexe 1

Classification de D'Amico

Risque d'évolutivité faible	PSA < 10 ng/mL et score de Gleason ≤ 6 et stade clinique T1c ou T2a.
Risque d'évolutivité faible	PSA compris entre 10 et 20 ng/mL ou score de Gleason de 7 ou stade T2b.
Risque d'évolutivité faible	PSA > 20 ng/mL ou score de Gleason ≥ 8 ou stade clinique T2c.

Annexe 2 : Score de CAPRA

Variable	Level	CAPRA-S points
Pre-surgical PSA level (ng/ml)	0.00 to 6.00	0
	6.01 to 10.00	1
	10.01 to 20.00	2
	> 20.00	3
Pathologic Gleason score	≤ 3 + 3 = 6	0
	3 + 4 = 7	1
	4 + 3 = 7	2
	≥ 4 + 4 = 8	3
Surgical margin status	Negative	0
	Positive	2
Extracapsular extension	No	0
	Yes	1
Seminal vesicle invasion	No	0
	Yes	2
Lymph node invasion	No	0
	Yes	1

AUTEUR(E) : Nom : GOSSE

Prénom : Rémi

Date de soutenance : 17 octobre 2022

Titre de la thèse : Reproductibilité inter observateur et performance de l'évaluation par IRM de l'extension extra-prostatique du cancer de la prostate : étude rétrospective sur une série de 80 patients.

Thèse - Médecine - Lille « 2022 »

Cadre de classement : Radiodiagnostic et imagerie médicale, imagerie génito-urinaire

DES + FST/option : Radiodiagnostic et imagerie médicale, imagerie génito-urinaire

Mots-clés : cancer de prostate, extension extra-prostatique, IRM, grade EPE.

Résumé : Introduction : Le cancer de prostate est le cancer le plus fréquent chez l'homme. La mise en évidence d'une extension extra-prostatique (EEP) en préopératoire est un élément pronostique majeur, modifiant la prise en charge thérapeutique. Plusieurs nomogrammes cliniques et scores IRM ont été proposés afin d'évaluer cette EEP, sans recommandations officielles. Parmi eux le score Likert et le grade EPE proposé par Mehralivand et al en 2019 permettant de standardiser la détection d'une EEP sur l'IRM. Objectif : L'objectif principal était d'évaluer les performances diagnostiques du grade EPE par la relecture de 80 IRM par 3 lecteurs d'expérience différente, de les comparer à l'étude princeps et au score Likert et d'évaluer sa reproductibilité inter-observateur (RIO). L'objectif secondaire est d'évaluer la RIO de la mesure de la longueur de contact capsulaire (LCC), critère du grade EPE. Matériel et méthode : une étude rétrospective, monocentrique a été réalisée au CHU de Lille, intégrant 80 patients ayant eu une IRM avant prostatectomie, relue par 3 radiologues d'expérience différente (1, 3 et > 20 ans) n'ayant pas connaissance du statut anatomopathologique. Les performances du grade EPE et du score Likert ont été mesurées pour chaque lecteur par l'aire sous la courbe (AUC), les sensibilité (Se) et spécificité (Sp) et les RIO par le coefficient alpha de Krippendorff. Résultats : Notre étude montre des performances diagnostiques du grade EPE aussi bonnes que celles retrouvées dans l'étude de Mehralivand et celles du score Likert. En effet, il n'existe pas de différence significative ($p > 0,05$) des AUC du grade EPE et du score Likert pour chaque lecteur (AUC respectivement de 0,84, 0,75 et 0,79 pour le grade EPE et de 0,81, 0,69 et 0,78 pour le score Likert). La RIO de ces scores est modérée, avec une tendance à être meilleure pour le grade EPE et entre les lecteurs plus expérimentés, sans différence significative. La RIO de la mesure de LCC est quant à elle modérée à bonne, fonction des lecteurs. Conclusion : Le grade EPE est un bon outil clinique, permettant de standardiser et de graduer le risque d'EEP en IRM de manière simple, tout en limitant l'impact de l'expérience du radiologue.

Composition du Jury :

Président : Monsieur le Professeur Philippe PUECH

**Assesseurs : Monsieur le Professeur Xavier LEROY
Monsieur le Docteur Jonathan OLIVIER
Monsieur le Docteur Louis COTTEN**

Directeur de thèse : Monsieur le Professeur Philippe PUECH