World Journal of Urology

How to implement magnetic resonance imaging before prostate biopsy in clinical practice: nomograms for saving biopsies

--Manuscript Draft--

Manuscript Number:	WJUR-D-19-00641R2
Full Title:	How to implement magnetic resonance imaging before prostate biopsy in clinical practice: nomograms for saving biopsies
Article Type:	Original Article
Keywords:	prostate cancer; Multiparemetric resonance imaging; PI-RADS; Nomograms
Corresponding Author:	Luis Mariano Esteban, Ph.D. Universidad de Zaragoza La Almunia De Doña Godina, Zaragoza SPAIN
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Universidad de Zaragoza
Corresponding Author's Secondary Institution:	
First Author:	Ángel Borque-Fernando
First Author Secondary Information:	
Order of Authors:	Ángel Borque-Fernando
	Luis Mariano Esteban, Ph.D.
	Ana Celma
	Sarai Roche
	Jacques Planas
	Lucas Regis
	Inés De Torres
	María Eugenia Semidey
	Enrique Trilla
	Juan Morote
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	Purpose To combine multiparametric MRI (mpMRI) findings and clinical parameters to provide nomograms for diagnosing different scenarios of aggressiveness of prostate cancer (PCa). Methods A cohort of 346 patients with suspicion of PCa because of abnormal finding in digital rectal examination (DRE) and/or high prostate specific antigen (PSA) level received mpMRI prior to prostate biopsy (PBx). A conventional 12-core transrectal PBx with 2 extra cores from suspicious areas in mpMRI was performed by cognitive fusion. Multivariate logistic regression analysis was performed combining age, PSA density (PSAD), DRE, number of previous PBx, and mpMRI findings to predict 3 different scenarios: PCa, significant PCa (ISUP-group >2), or aggressive PCa (ISUP-group >3).

	We validate models by ROC curves, calibration plots, probability density functions (PDF), and clinical utility curves (CUC). Cut-off probabilities were estimated for helping decision-making in clinical practice. Results Our cohort showed 39.6% incidence of PCa, 32.6% of significant PCa, and 23.4% of aggressive PCa. The AUC of predictive models were 0.856,0.883, and 0.911, respectively. The PDF and CUC showed 11% missed diagnoses of significant PCa (35 cases of 326 significant PCa expected in 1000 proposed Bx) when choosing <18% as the cutoff of probability for not performing PBx; the percentage of saved PBx was 47% (474 avoided PBx in 1000 proposed). Conclusion We developed clinical and mpMRI-based nomograms with a high discrimination ability for 3 different scenarios of PCa aggressiveness (https://urostatisticalsolutions.shinyapps.io/MRIfusionPCPrediction/). Specific clinical cutoff points allow us to save a high number of PBx with a minimum of missed diagnoses.
Response to Reviewers:	Dear Editor and Reviewers: Thank you again for the comments and suggestions concerning our manuscript. All comments were carefully reviewed and we made revisions that hope will earn your approval. The revised sections are highlighted in yellow color in the manuscript. Responses to the reviewer' comments: Reviewer #2: The manuscript is well written and the normograms are easily to adopt into clinical daily routine and can help urologists in the decision, if a PI-RADS 3 lesion has to be biopsied. I only have some little annotations: Introduction: There are more limitations of the mpMRI than cost and time. This sentence should be adjusted. We totally agree with this comment, MRI inter-reader reproducibility is moderate at best, what is more troublesome is that its accuracy depends on the specialization from individual radiologists, which is difficult in many hospitals. We added a new comment in the introduction section: "The remaining limitations of mpMRI are cost, its time-consuming nature and the necessary highly specialization of radiologists to increase inter-reader reproducibility." Results: The author describes a subcohort of PCa or ISUP grade more than 3 being significantly older without telling the actual result. How many patients?What is the mean age of that cohort? And the p value is missing. Thank you for your suggestion, we have complemented the information in the results section: "with a subcohort of PCa (n=137, P50: 70.5, P25-75: 65.0-76.1) or ISUP3 (n=81, P50: 73.1, P25-75: 68.0-77.9) being significantly older ($p < 0.001$ in both cases)." Reviewer #3: changes allow a better interpretation of results Thank for you comment

How to implement magnetic resonance imaging before prostate biopsy in clinical practice: nomograms for saving biopsies

Ángel Borque-Fernando[1], Luis Mariano Esteban[2], Ana Celma[3], Sarai Roche[4], Jacques Planas[3], Lucas Regis[3], Inés de Torres[5,6], Maria Eugenia Semidey [5], Enrique Trilla[3,6], Juan Morote[3,6]

[1] Department of Urology, Hospital Universitario Miguel Servet, Zaragoza, IIS Aragón, Spain

[2] Escuela Universitaria Politécnica La Almunia, Universidad de Zaragoza, Spain

[3] Department of Urology, Hospital Vall d´Hebron, Barcelona, Spain

[4] Department of Radiology, Hospital Vall d´Hebron, Barcelona, Spain

[5] Department of Pathology, Hospital Vall d´Hebron, Barcelona, Spain .

[6] Universidad Autónoma de Barcelona, Barcelona, Spain .

Corresponding author:
Luis Mariano Esteban.
Department of Applied Mathematics, Escuela Universitaria Politécnica de La Almunia, Universidad de Zaragoza.
C/ Mayor 5, 50100 La Almunia de Doña Godina; Zaragoza, Spain
Phone number: 0034659479612
lmeste@unizar.es

Introduction

Prostate cancer (PCa) is the most common solid tumor diagnosed among men in developed countries [1,2]. Its diagnosis is based on prostate biopsy (PBx), which is indicated when there is a high suspicion of PCa based on either an increased level of prostate specific antigen (PSA) or abnormal finding in digital rectal examination (DRE). In this context, PBx can identify the tumor in 30%-40% of cases.

To optimize the efficacy of PBx, we use adjustments such as PSA density (PSAD); percentage of free-PSA; age; or novel serum or urine biomarkers such as PCA3, Prostate Health Index, 4Kscore Test, or SelectMDx. However, presently, imaging in PCa has changed with the increasing relevance of magnetic resonance imaging (MRI) [3], its combination with biomarkers [4], and the guidelines recommend its use [5,6,7,8]. A standardized protocol of multiparametric MRI (mpMRI) of PCa, known as the Prostate Imaging–Reporting and Data System (PI-RADS v2.2015) [9,10], has been recently updated. This protocol attemps to homogenize interpretation criteria and reduce interobserver variability [11]. The remaining limitations of mpMRI are cost, ~~and~~ its time-consuming nature and the necessary highly specialization of radiologists to increase inter-reader reproducibility.

One of the main objectives of PI-RADS is identify not all PCas, but those PCas that are clinically significant, which means high volume and/or high-grade PCa. In addition, the characterization of aggressiveness of PCa has been recently updated by the International Society of Uropathologist (ISUP), wherein some Gleason patterns have been redefined and the Gleason score has been reclassified into 5 groups by focusing on separating non-PCa and low-aggressive PCa. We remark three aspects: the redefinition of Gleason pattern 3 and 4 and its consequence in updating of Gleason score 6 which was renamed as ISUP group 1; following, a better discrimination between Gleason 7 (3+4), renamed as ISUP group 2, and Gleason 7 (4+3), renamed ISUP group 3; and finally, the gap between Gleason score 8, new ISUP group 4, and the Gleason score 9-10, new ISUP group 5 [12].

In this changing era of new imaging ability of mpMRI and an updated Gleason grade with less oncological and prognostic relevance of Gleason score 6 than before, we conducted a prospective study to define the best clinical implementation of mpMRI before PBx in actual clinical practice. We focused on building a nomogram and choosing cutoff points for clinical implementation.

Materials and methods

An ambispective data collection for our analysis was performed in 389 patients with suspected PCa on the basis of PSA level above 4 ng/mL or abnormal finding in DRE, between January 2015 and September 2016. The study was conducted in the Urology Unit of the Vall d'Hebron Institute of Oncology in Barcelona, Spain. All patients underwent a 3T mpMRI (MAGNETOM Trio^{TM}, Siemens) prior to PBx according to the center's protocol, and no one had been submitted to a previous mpMRI-guided PBx. A local ethics committee approved this

biomedical research project (VH-294/2017).

Pre-planned mpMRI was not performed in 43 patients because of contraindications or claustrophobia, and these patients were excluded from the study, thus giving a total cohort of 346 patients. A radiologist with more than 300 prostate mpMRI previous experience (S.R.) reported mpMRI outcomes according to the PI-RADS v2 classification [9]. This was performed prospectively after November 2015, and retrospectively before this date by reclassifying MRI findings from PI-RADS v1 into PI-RADS v2.

An experienced urologist (A.C.) with more than 6,000 PBx practised during last 12 years, reviewed the reported mpMRI. Later, she performed a conventional 12-core systematic transrectal ultrasound-guided PBx, with 2 extra cores by cognitive fusion from each suspicious area, up to a maximum of 3 suspicious areas. Prostate cores were analysed by an uropathologist with more than 20 years' experience in this area (I. de T.). The PCa aggressiveness was reported according to the updated International Society of Urological Pathology (ISUP) and World Health Organization (WHO) classification consensus [13].

Descriptive statistics of the total cohort were provided and stratified by pathological findings: all PCa versus Non PCa, ISUP$\geq$2 versus Non PCa-ISUP 1, and ISUP$\geq$3 versus Non PCa-ISUP 1-2. Chi- squared test, *t*- test and Mann-Whitney test were used as appropriate to establish statistically significant differences between the groups.

A multivariate analysis was performed to diagnose the 3 pathological scenarios proposed named as: all PCa, significant PCa (ISUP$\geq$2), or aggressive significant PCa (ISUP$\geq$3). Logistic regression model combined the candidate variables: age, number of previous biopsies, PSA, free-PSA, PSAD, prostate volume (PV), DRE, and PI-RADS v2. Continuous variables were modeled as linear or nonlinear predictors by using restricted cubic splines.

Predictive models were assessed by calibration, discrimination, and clinical utility analysis for future implementation in actual clinical practice. The calibration curve, area under the Receiver Operating Characteristic curve (AUC), probability density functions (PDF), and clinical utility curve (CUC) were performed for this purpose, in the same way as that done in previous analyses developed by our group in other PCa scenarios [14,15,16]. In addition, our internal validation was subjected to a bias-correction AUC using 1000 bootstrap samples.

For clinical use, we investigated the best threshold probability for each model. We considered the delayed diagnosis and PBx saved through CUC. In addition, a nomogram and an App were provided as user-friendly tools to apply models in routine clinical practice.

Statistical analyses were computed using R programming language v.3.3.1 (The R statistical foundation, Vienna, Austria). Tests were two- sided and *p-values*< 0.05 were considered as statistically significant.

Results

A total of 39.6% of patients were found to have any grade of PCa, 32.6% had ISUP$\geq$2, and 23.4% had ISUP$\geq$3.

Table 1 summarizes the characteristics of the patients. The median age of our study patients was 67.7 years (P25-75: 63.0-73.7), with a subcohort of PCa (n=137, P50: 70.5, P25-75: 65.0-76.1) or ISUP≥3 (n=81, P50: 73.1, P25-75: 68.0-77.9) being significantly older ($p < 0.001$ in both cases). Total PSA showed a median value of 6.1 ng/mL (P25-75: 4.7-10.2), which was significantly higher in adverse pathological subcohorts. Measurements of free-PSA were not available in 105 patients, because they had total PSA below 4 ng/mL or over 10 ng/mL; there were no differences between the groups. The PV measured at the time of mpMRI showed a median value of 50.3 mL (P25-75: 38-78.2), which was significantly smaller in adverse pathological findings. Its related variable PSAD had a median value of 0.13 ng/mL/cc (P25-75: 0.08-0.20) and showed significant differences between groups of subanalysis.

Half of the patients underwent a repeat PBx due to prior negative PBx and remained suspicious of harbouring cancer during follow-up; this proportion significantly increased when no adverse pathological findings were obtained. One third of the patients had an abnormal DRE before PBx, and this proportion increased when a PCa or an aggressive PCa was detected. A significantly higher proportion of PI-RADS 4-5 was found in adverse pathological subcohorts.

We found statistically significant differences between all PCa/non-PCa, ISUP≥2/non-PCa-ISUP 1, and ISUP≥3/non-PCa-ISUP 1-2 groups in all predictor variables except free-PSA for all groups and DRE for the last comparison.

In addition, 3 multivariate models and nomograms were built to predict any pathological event in each group: all PCa, ISUP≥2, and ISUP≥3, Figure 1. Odds ratios and p-values of variables are shown in Table 2. From the candidate variables, PSAD showed a nonlinear dependence modeled through restricted cubic splines, and age showed a linear dependence in the first and second model, but with a significant nonlinear relationship in the third model. Number of biopsies (first/repeat), DRE, and PI-RADS are also categorical risk factors.

Multivariate models showed a good calibration (Figure 2) because of a high concordance between actual and predicted probabilities. Moreover, the 3 models showed a high discrimination ability with AUC values of 0.856, 0.876, and 0.911, respectively. An internal validation was performed using 1000 bootstrap samples, yielding AUC bias-corrected values of 0.845, 0.866, and 0.901. These results were confirmed by PDF (Figure 3). Figure 3a shows the all PCa/non-PCa predictive model; there is a moderate overlapping in probabilities for both groups, but by choosing a threshold point near 40%, we can separate both of them. Figure 3b and 3c show the PDF for ISUP's adverse predictive models. In both cases, the probabilities provided by the models for non-PCa/ISUP 1 or non-PCa/ISUP 1-2, respectively, are distributed in a narrow range below 20%. However, the probabilities for complementary adverse subcohorts ISUP≥2 or ISUP≥3 appear in a wide range; therefore, we can separate patients with adverse pathology with a threshold point around 20%.

Finally, in Figure 4, we show the CUC. Here, we present the percentage of delayed diagnoses due to the non-perfect discrimination ability of the model and the saved biopsies at any threshold of probabilities.

From a conservative perspective, we can choose the threshold probability points that provide a delayed diagnosis below 10% at every prediction. We should then choose 20%, 18%, and 16% at threshold probabilities to make a clinical decision, with a percentage of saved biopsies of 34%, 46%, and 60%, for all PCa, ISUP≥2, or ISUP≥3 predictions, respectively.

A more detailed subanalysis of the nomogram applicable in the more challenging clinical scenario that is prediction of ISUP≥2 (Figure 1b) is presented at Figure 5. At this point we evaluate the applicability and benefits of using this nomogram in the subcohort of first PBx or at repeated PBx. The same criteria of a delayed diagnosis of ISUP≥2 below 10% counsels us a threshold of 24% in the subcohort of first PBx and 9% in repeated PBx. Not performing PBx at patients with a probability of ISUP 2 under those cutoff drives us to miss less than 10% of real ISUP 2 in both scenarios, but avoiding 43% and 37% of initially proposed PBx respectively (Table 1, suppl material). A more conservative approach with a threshold of 18% for first PBx lets us to avoid a 37% of proposed first PBx with less than a 5% ISUP ≥2. (Figure 1, suppl material)

The clinical benefits of our nomograms can be evaluated on a user-friendly interactive App: https://urostatisticalsolutions.shinyapps.io/MRIfusionPCPrediction/

Discussion

Contemporary diagnosis of PCa has changed dramatically in terms of criteria, resources, and objectives, leading us to an updated definition of our clinical practice. The year 2015 was a key year because of update in ISUP and PI-RADS v2.2015. Efforts should be made to develop updated new tools that implement these changes and help clinicians in decision-making in this challenging situation. Some characteristics of the design of our study make it especially relevant in this context.

We use a recent population sample from 2015 to 2016. We ensure an updated interpretation of mpMRI and pathology by our central radiologist and pathologist. Further, we evaluate a real and reproducible common practice of systematic transrectal ultrasound (TRUS) PBx by the 12-core conventional approach and fusion cognitive addition.

Finally, we define 3 different goals in prediction: from a more conservative scenario, all PCa vs non-PCa, to a less conservative one, ISUP≥3 (primary Gleason pattern ≥ 4) vs non-PCa/ISUP 1-2, and the intermediate and more realistic scenario ISUP≥2 (Gleason score ≥ 7) vs non-PCa/ISUP 1.

A key point of our analysis is that we do not exclude first or repeat PBx, but we include them as a predictive variable, thus enabling our predictive models to be completely implemented in actual clinical practice. Moreover, we used a well-balanced sample in the distribution of variables that allows us to include and precisely weigh them in our multivariate models, in the case of significant association.

Some previous efforts in developing MRI-based nomograms or multivariable models for PCa prediction have been published [17,18,19,20,21]. Four of them were trained with a cohort of

patients before [17,18] or mainly before [19,21] the 2015-ISUP update interpretation [12]. They also used a Likert scale for MRI interpretation [22,23] or PI-RADS v1 [24] but not the updated PI-RADS v2 [9]. Regarding PBx procedures, three different protocols were performed: a minimum 12-core systematic and MRI-ultrasound software fusion-targeted TRUS-PBx [17] specifically targeted to only one suspicious lesion [20], a minimum 30-core transperineal approach using MRI/TRUS fusion or cognitive PBx [18], or a median systematic 24-core adjusted to PV PBx by the transperineal approach and MRI-software fusion for MRI-suspicious lesions [19].These specific approaches are reproducible but not used conventionally in clinical practice in PBx, and they should be taken into account if we consider using those nomograms. The most recent published study used three different approaches for target PBx: in bore, software fusion, cognitive fusion and even no target PBx in 17% of cases [21]. Two studies [17,19] used 2 different nomogram models, with one for first PBx and the other for repeated PBx, and patients with a low suspicion of PCa (using MRI) were excluded [17,19,20]; DRE was not evaluated as a predictive variable [17]. Some nomograms included patients with mpMRI conducted with a 1.5-T magnet and 3-T magnet indistinctly [18]. One study used mpMRI and biparametric MRI [21]. All that previous heterogeneities have been avoided in our project. Our study was designed to present an updated radiological and pathological tool.

Three final variables were selected for building nomograms in a study by Bjurlin et al [17]: age, PSAD, and MRI findings. PSAD was included in a linear manner, but we know that its gray zone ranges between 0.10 and 0.25 ng/mL/cc. The changes over 0.25 ng/mL/cc should increase the risk of adverse pathological findings. However, changes over this threshold confirm the risk, but should not increase it as they should do below the threshold of 0.25 ng/mL/cc. This nonlinear influence is clearly understood in our models. Van Leeuwen et al did not select PSAD but used variables PSA and PV, and in a non-linear manner [18], as Mehralivand did with PV and PSAD [20]; in addition, Radtke et al included the logarithm of PSA but not PV [19]. The rest of the associated variables included in the other models, such as age, DRE, first/repeated PBx, and MRI findings match our selected predictive variables. Previous PBx as a predictive variable is included in van Leeuwen's, Mehralivand's, Alberts' and our models, and it confers greater applicability than designing separate models as reported in Bjurlin and Radtke's studies. In Mehralivand's study, we find 3 serious caveats against its implementation; (1) the lack of a user-friendly tool as a nomogram for obtaining individualized probabilities, (2) the fact that they regularly include PBx practiced according to PI-RADS 3-5 findings (and exceptionally that they include PBx in PI-RADS 1-2 but not by protocol). We know that PI-RADS 1-2 have a very low probability of harboring PCa or significant PCa, but these cannot benefit from individual prediction with this model, and (3) this study avoids practicing targeted PBx from more than one area suggested as suspicious by MRI [20].

Our calibration curves showed an excellent correlation between predicted probabilities of adverse pathology and actual findings in all 3 estimation scenarios. Our internal validation

showed AUC for an overall PCa prediction of 0.856 over 0.82-0.76 for first or repeated PBx in Bjurlin's nomograms [17], and 0.839-0.791 in Alberts' models [21]. For ISUP≥2 the AUC was 0.876 versus 0.91-0.86 in Bjurlin's models [17], 0.876 in training van Leeuwen cohort [18], 0.83-0.81 for first or previous PBx in Radtke training cohort [19], 0.84 in training Mehralivand cohort [20], and 0.843-0.850 in Alberts' study [21]. To the best of our knowledge, there are no previous comparative nomograms published to compare with our AUC of 0.911 for predicting ISUP≥3 as an aggressive significant PCa.

Finally, we analyzed the clinical utility of nomograms by PDF and CUC [14,15,16]. With these tools we chose a cutoff of 18% probability of significant PCa (ISUP≥2 vs Non-PCa/ISUP 1); therefore, by not performing PBx for patients with a probability of ISUP≥2 under 18%, we could avoid 47% of PBx procedures (474 PBx avoided over 1000 indicated) but missed 11% of significant PCa (327 significant PCas would be expected in 1000 PBx indicated; 11% of these significant PCas would be missed, 35 cases in 1000 PBx) (Figure 1a, suppl material). We explored the management of nomogram focused on ISUP≥2 prediction (Figure 1b) and two different scenario, first or repeated PBx, with different prevalence of disease, 41% and 25% respectively (Table 1). A threshold of 24% in the subcohort of first PBx avoided 436 indications in 1000 proposed PBx missing 37 ISUP≥2; a more conservative cutoff of 18% saved 367 patients in 1000 PBx, missing 18 ISUP≥2 (Figure 1b-c, suppl material). In the subcohort of repeated PBx the threshold of 9% saved 372 PBx in 1000 patients, with 21 ISUP≥2 missed (Figure 1d, suppl material). A more conservative or aggressive criteria in saving PBx and missing diagnoses can be chosen by different cutoff points in PDF and CUC, and especially by implementing a dynamic and friendly useful App for clinical use, (https://urostatisticalsolutions.shinyapps.io/MRIfusionPCPrediction/) [25]. An analysis of cutoff points was performed by van Leeuwen et al evaluating a 10% nomogram-derived risk as the threshold for PBx, resulting in avoidance of 28% of biopsies whilst missing significant PCa in only 2.6% of the population. Mehralivand investigated 3 cutoff points of probabilities: 10% 15%, and 20%; a 10% threshold can avoid 17% of PBx in PI-RADS 3-5 patients, but miss 3% of significant PCa [20]. A similar conservative approach with our model reveals how a cut-off point of 7% saves 25.1% of PBx but misses 2.7% of significant PCa.

Integrated nomograms with PCA3 and PHI for PCa prediction showed AUC of 0.725 and 0.80 in their first evaluations [26,27]. The 4Kscore Test is a multivariate model for significant PCa prediction with an AUC of 0.82 [28]. The North American Prostate Cancer Prevention Trials-based Cancer Risk Calculator and the European Randomized Study of Screening for Prostate Cancer-derived Prostate Risk Indicator showed an AUC for PCa of 0.702 and 0.79 and significant PCa of 0.698 and 0.86, respectively [29,30].

We can consider some potential limitations of our study. First, the central and well-trained radiological and pathological review could diminish data reproducibility in less-experienced hands, but this expertise is desirable and should be improved by the ISUP-2015 and PI-RADS

v2.2015 attemps of standarization included in our models. Second, the cost-benefit of introducing an additional 3T MRI pre-PBx step has not been evaluated against saved PBx and missed PCa/aggressive PCa; unfortunately, we have no data to evaluate costs of missed diagnosis as it occurs generally. The gold standard for pathology is radical prostatectomy rather than PBx findings. However, other criteria of significant PCa (length or percentage of core involvement in PBx) and agressive PCa (ISUP-2015) can be evaluated from PBx findings. These limitations were also noted in previous studies; however, for aggressive PCa, the presence of ISUP≥3 or even ISUP≥2 should be enough to take an active decision in clinical practice. Most of the urologists would indicate PBx in PIRADS 4-5 without further consideration, especially in first PBx. In this category our nomograms probably offer no more than a confirmatory calculus, but it is in the cases of PIRADS 1-3 where the rest of the variables included, especially non-linear PSAD, can offer them its main utility. Our study was developed in a cohort of cognitive fusion PBx, and its accuracy with software fusion PBx should be evaluated before using this approach. Finally, the lack of external validation is a limitation for the clinical implementation of our predictive models. We encourage external validations of our models in advance.

Our present findings improve previous results, but the most relevant advantages of our proposal are the inclusion of updated pathological and radiological criteria, the contemporary cohort, the common and reproducible PBx approach and MRI target cores, the 3 scenarios of prediction proposed, and an easy and intuitive way to choose cutoff points for routine clinical practice in every patient by PDF and CUC as our App shows.

Conclusion

We developed mpMRI-based nomograms with a high discrimination ability for 3 different scenarios of PCa aggressiveness (https://urostatisticalsolutions.shinyapps.io/MRIfusionPCPrediction/). Specific clinical cutoff points allow us to save a high number of PBx with a minimum of missed diagnoses.

Author contributions

A Borque-Fernando: Project development, Data analysis, Manuscript writing

LM Esteban: Project development, Data analysis, Manuscript writing

A Celma: Project development, Data collection, Manuscript writing

S Roche: Project development, Data collection, Manuscript writing

J Planas: Project development, Data collection, Manuscript writing

L Regis: Project development, Data collection, Manuscript writing

I de Torres: Project development, Data collection, Manuscript writing

ME Semidey: Project development, Data collection, Manuscript writing

E Trilla: Project development, Data collection, Manuscript writing

J Morote: Project development, Data collection, Manuscript writing

Funding

No funding was received for this work.

Compliance with ethical standards

Conflict of interest

The authors declare that there is no conflict of interest from any of the co-authors.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent

Patients prospectively collected signed informed consent form.

References

[1] Siegel R, Miller KD, Ahmedin J (2017) Cáncer Statistics. Ca Cáncer J 67: 7–30. https://doi.org/10.3322/caac.21387.

[2] Galceran J, Ameijide A, Carulla M, Mateos A, Quirós JR, et al (2017) Cancer incidence in Spain, 2015. Clin Transl Oncol 19: 799–825. https://doi.org/10.1007/s12094-016-1607-9.

[3] Li M, Huang Z, Yu H, Wang Y, Zhang Y, Song B (2019) Comparison of PET/MRI with multiparametric MRI in diagnosis of primary prostate cancer: A meta-analysis. European Journal of Radiology 113: 225-231. https://doi.org/10.1016/j.ejrad.2019.02.028

[4] Cuocolo R, Stanzione A, Rusconi G, Petretta M, Ponsiglione A, Fusco F, Longo N, Persico F, Cocozza S, Brunetti A, Imbriaco M (2018) PSA-density does not improve bi-parametric prostate MR detection of prostate cancer in a biopsy naïve patient population. *European journal of radiology 104:* 64-70. https://doi.org/10.1016/j.ejrad.2018.05.004

[5] Sanda MG, Cadeddu J, Kirkby E, Chen RC, Crispino T, Fontanarosa J, Freedland SJ, Greene K, Klotz LH, Makarov DV, Nelson JB, Rodrigues G, Sandler HM, Taplin ME, Treadwell JR (2018) Clinically Localized Prostate Cancer: AUA/ASTRO/SUO Guideline. Part II: Recommended Approaches and Details of Specific Care Options. J Urol 199: 990–997. https://doi.org/10.1016/j.juro.2018.01.002.

[6] Carroll PR, Parsons JK, Andriole G, Bahnson RR, Carlsson S, Castle EP, Catalona WJ, Dahl DM, Davis JW, Epstein JI, Etzioni RB, Farrington T, Hemstreet GP, Jarrard D, Kibel AS, Kim S, Lowrance W, Maroni P, Mohler J, Morgan TM, Moses KA, Nadler RB, Poch M, Scales C, Shaneyfelt TM, Smaldone MC, Sonn G, Sprenkle P, Wake R, Yuh B (2019) NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®) Prostate Cancer Early Detection Version 1.2019. https://www.nccn.org/professionals/physician_gls/pdf/prostate_detection.pdf.

[7] Mohler JL, Antonarakis ES, Armstrong AJ, D'Amico AV, Davis BJ, Dorff T, Eastham JA, Enke CA, Farrington T, Ellis R, Higano CS, Horwitz EM, Hurwitz M, Ippolito JE, Kane CJ, Kuettel ME, Lang JM, Netto G, Penson DF, Plimack ER, Pow-Sang JM, Pugh TJ, Richey S, Roach III M, Rosenfeld S, Schaeffer E, Shabsigh A, Small EJ, Spratt DE, Srinivas S, Tward J (2019) NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®) Prostate Cancer Version 1.2019. https://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf.

[8] Mottet N, van den Bergh R, Briers E, Cornford P, De Santis M, Fanti S, Gillesen S, Grummel J, Henry AM, Lam TB, Mason MD, van der Kwast T, van der Poel HG, Rouvière O, Tilki D, Wiegel T, van der B T, Cumberbatch M, Fossati N, Gross T, Lardas M, Liew M, Moris L, Schoots IG, Willemse PM (2019) EAU-EANM-ESTRO-ESUR-SIOG, Guidelines on Prostate Cancer. 2019. https://uroweb.org/guideline/prostate-cancer/#3.

[9] PI-RADS TM Prostate Imaging – Reporting and Data System (2015). https://www.acr.org/-/media/ACR/Files/RADS/Pi-RADS/PIRADS-V2.pdf?la=en. Accessed May 14, 2019.

[10] Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, Margolis D, Schnall MD, Shtern F, Tempany CM, Thoeny HC, Verma S, Catto J (2016) PI-RADS Prostate Imaging – Reporting and Data System: 2015, Version 2. Eur Urol 69: 16–40. https://doi.org/10.1016/j.eururo.2015.08.052.

[11] Schaudinn A, Gawlitza J, Mucha S, Linder N, Franz T, Horn LC, Kahn T, Busse H (2019) Comparison of PI-RADS v1 and v2 for multiparametric MRI detection of prostate cancer with whole-mount

histological workup as reference standard. European Journal of Radiology. https://doi.org/10.1016/j.ejrad.2019.04.012

[12] Epstein JI, Zelefsky MJ, Sjoberg DD, Nelson JB, Egevad L, Magi-Galluzzi C, Vickers AJ, Parwani AV, Reuter VE, Fine SW, Eastham JA, Wiklund P, Han M, Reddy CA, Ciezki JP, Nyberg T, Klein EA (2016) A Contemporary Prostate Cancer Grading System: A Validated Alternative to the Gleason Score. Eur Urol 69: 428–435. https://doi.org/10.1016/j.eururo.2015.06.046.

[13] Humphrey PA, Moch H, Cubilla AL, Ulbright TM, Reuter VE (2016) The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs Part B: Prostate and Bladder Tumours. Eur Urol 70: 106–119. https://doi.org/10.1016/j.eururo.2016.02.028.

[14] Rubio-Briones J, Borque A, Esteban LM, Casanova J, Fernandez-Serra A, Rubio L, Casanova-Salas I, Sanz G, Domínguez-Escrig J, Collado A, Gómez-Ferrer A, Iborra I, Ramírez-Backhaus M, Martínez F, Calatrava A, Lopez-Guerrero JA (2015) Optimizing the clinical utility of PCA3 to diagnose prostate cancer in initial prostate biopsy. BMC Cancer. 15: 633. https://doi.org/10.1186/s12885-015-1623-0.

[15] Borque-Fernando A, Esteban-Escaño LM, Rubio-Briones J, García-Ruiz R, Tejero-Sánchez A, Muñoz-Rivero MT, Alfaro-Torres J, Marquina-Ibáñez IM, Hakim-Alonso S, Mejía-Urbáez E, Gil-Fabra J, Gil-Martínez P, Álvarez-Alegret R, Sanz G, Gil-Sanz MJ (2017) A preliminary study of the ability of the 4Kscore test , the Prostate Cancer Prevention Trial-Risk Calculator and the European Research Screening Prostate-Risk Calculator for predicting high-grade prostate cancer. Actas Urol Esp 40: 155–163.

[16] Borque A, Rubio-Briones J, Esteban LM, Sanz G, Domínguez-Escrig J, Ramírez-Backhaus M, Calatrava A, Solsona E (2014) Implementing the use of nomograms by choosing threshold points in predictive models: 2012 updated Partin Tables vs a European predicitive nomogram for organ-confined disease in prostate cancer. BJU Int 113: 878–86. https://doi.org/10.1111/bju.12532.

[17] Bjurlin MA, Rosenkrantz AB, Sarkar S, Lepor H, Huang WC, Huang R, Venkataraman R, Taneja SS (2018) Prediction of Prostate Cancer Risk Among Men Undergoing Combined MRI-targeted and Systematic Biopsy Using Novel Pre-Biopsy Nomograms That Incorporate MRI Findings. Urology 112: 112–120. https://doi.org/10.1016/j.jchromb.2016.09.003.

[18] van Leeuwen PJ, Hayen A, Thompson JE, Moses D, Shnier R, Böhm M, Abuodha M, Haynes AM, Ting F, Barentsz J, Roobol M, Vass J, Rasiah K, Delprado W, Stricker PD (2017) A multiparametric magnetic resonance imaging-based risk model to determine the risk of significant prostate cancer prior to biopsy. BJU Int 120: 774–781. https://doi.org/10.1111/bju.13814.

[19] Radtke JP, Bonekamp D, Kesch C, Freitag MT, Alt CD, Celik K, Distler F, Roth W, Wieczorek K, Duensing S, Roethke MC, Teber D, Schlemmer HP, Hohenfellner M, Hadaschik BA, Wiesenfarth M, Stock C (2017) Combined clinical parameters and multiparametric MRI for advanced risk modeling of prostate cancer - patient-tailored risk stratification can reduce unnecessary biopsies. Eur Urol 72: 888–896. https://doi.org/10.1016/S1569-9056(17)30557-2.

[20] Mehralivand S, Shih JH, Rais-Bahrami S, Oto A, Bednarova S, Nix JW, Thomas JV, Gordetsky JB, Gaur S, Harmon SA, Siddiqui MM, Merino MJ, Parnes HL, Wood BJ, Pinto PA, Choyke PL, Turkbey B (2018) A Magnetic Resonance Imaging–Based Prediction Model for Prostate Biopsy Risk Stratification. JAMA Oncol 4: 678–685. https://doi.org/10.1001/jamaoncol.2017.5667.

[21] Alberts AR, Roobol MJ, Verbeek FJM, Schoots IG, Chiu PK, Osses F, Tijsterman JD, Beerlage HP, Mannaerts CK, Albers P, Arsov C, Schimmo L, Novara G, Vickers A (2019) Prediction of High-grade

Prostate Cancer Following Multiparametric Magnetic Resonance Imaging : Improving the Rotterdam European Randomized Study of Screening for Prostate Cancer Risk Calculators. Eur Urol 75: 310–318. https://doi.org/10.1016/j.eururo.2018.07.031.

[22] Rosenkrantz AB, Kim S, Lim RP, Hindman N, Deng FM, Babb JS, Taneja SS (2013) Prostate cancer localization using multiparametric MR imaging: Comparison of Prostate Imaging Reporting and Data System (PI-RADS) and Likert scales. Radiology 269: 482–92. https://doi.org/10.1148/radiol.13122233/-/DC1.

[23] Wysock JS, Rosenkrantz AB, Huang WC, Stifelman MD, Lepor H, Deng FM, Melamed J, Taneja SS (2014) A prospective, blinded comparison of magnetic resonance (MR) imaging-ultrasound fusion and visual estimation in the performance of MR-targeted prostate biopsy: The profus trial. Eur Urol 66: 343–351. https://doi.org/10.1016/j.eururo.2013.10.048.

[24] Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, Rouviere O, Logager V, Fütterer JJ (2012) ESUR prostate MR guidelines 2012. https://doi.org/10.1007/s00330-011-2377-y.

[25] Borque-Fernando A, Esteban LM, Morote-Robles J, Sanz G (2019) WHO Predictive model. https://urostatisticalsolutions.shinyapps.io/FusionBiopsy/. Accessed May 14, 2019.

[26] Chun FK, de la Taille A, van Poppel H, Marberger M, Stenzl A, Mulders PFA, Huland H, Abbou CC, Stillebroer AB, van Gils MP, Schalken JA, Fradet Y, Marks LS, Ellis W, Partin AW, Haese A (2009) Prostate Cancer Gene 3 (PCA3): Development and Internal Validation of a Novel Biopsy Nomogram. Eur Urol 56: 659–668. https://doi.org/10.1016/j.eururo.2009.03.029.

[27] Lughezzani G, Lazzeri M, Larcher A, Lista G, Scattoni V, Cestari A, Buffi NM, Bini V, Guazzoni G (2012) Development and internal validation of a prostate health index based nomogram for predicting prostate cancer at extended biopsy. J Urol 188: 1144–1150. https://doi.org/10.1016/j.juro.2012.06.025.

[28] Parekh DJ, Punnen S, Sjoberg DD, Asroff SW, Bailen JL, Cochran JS, Concepcion R, David RD, Deck KB, Dumbadze I, Gambla M, Grable MS, Henderson RJ, Karsh L, Krisch EB, Langford TD, Lin DW, McGee SM, Munoz JJ, Pieczonka CM, Rieger-Christ K, Saltzstein DR, Scott JW, Shore ND, Sieber PR, Waldmann TM, Wolk FN, Zappala SM (2015) A multi-institutional prospective trial in the USA confirms that the 4Kscore accurately identifies men with high-grade prostate cancer. Eur Urol 68: 464–70. https://doi.org/10.1016/j.eururo.2014.10.021.

[29] Thompson IM, Ankerst DP, Chi C, Goodman PJ, Tangen CM, Lucia MS, Feng Z, Parnes HL, Coltman CA (2006) Assessing prostate cancer risk: Results from the Prostate Cancer Prevention Trial. J Natl Cancer Inst 98: 529–534. https://doi.org/10.1093/jnci/djj131.

[30] Roobol MJ, Van Vugt HA, Loeb S, Zhu X, Bul M, Bangma CH, Van Leenders AG, Steyerberg EW, Schröder FH (201) Prediction of prostate cancer risk: The role of prostate volume and digital rectal examination in the ERSPC risk calculators. Eur Urol 61: 577–583. https://doi.org/10.1016/j.eururo.2011.11.012.

LEGENDS figures

- Figure 1: Nomograms of a) all prostate cancer prediction; b) ISUP group 2-5, Gleason score $\geq$ 7 prediction; and c) ISUP group 3-5, primary Gleason pattern $\geq$ 4 prediction.

- Figure 2: Calibration of predictive models. These graphics show the correlation between predicted probabilities (X-axis) and actual incidence of the event (Y-axis).

- Figure 3: Probability density functions of predictive models. These graphics shows the distribution of assigned probabilities of predicted event, among actual events (red area) and hypothetical events (blue area).

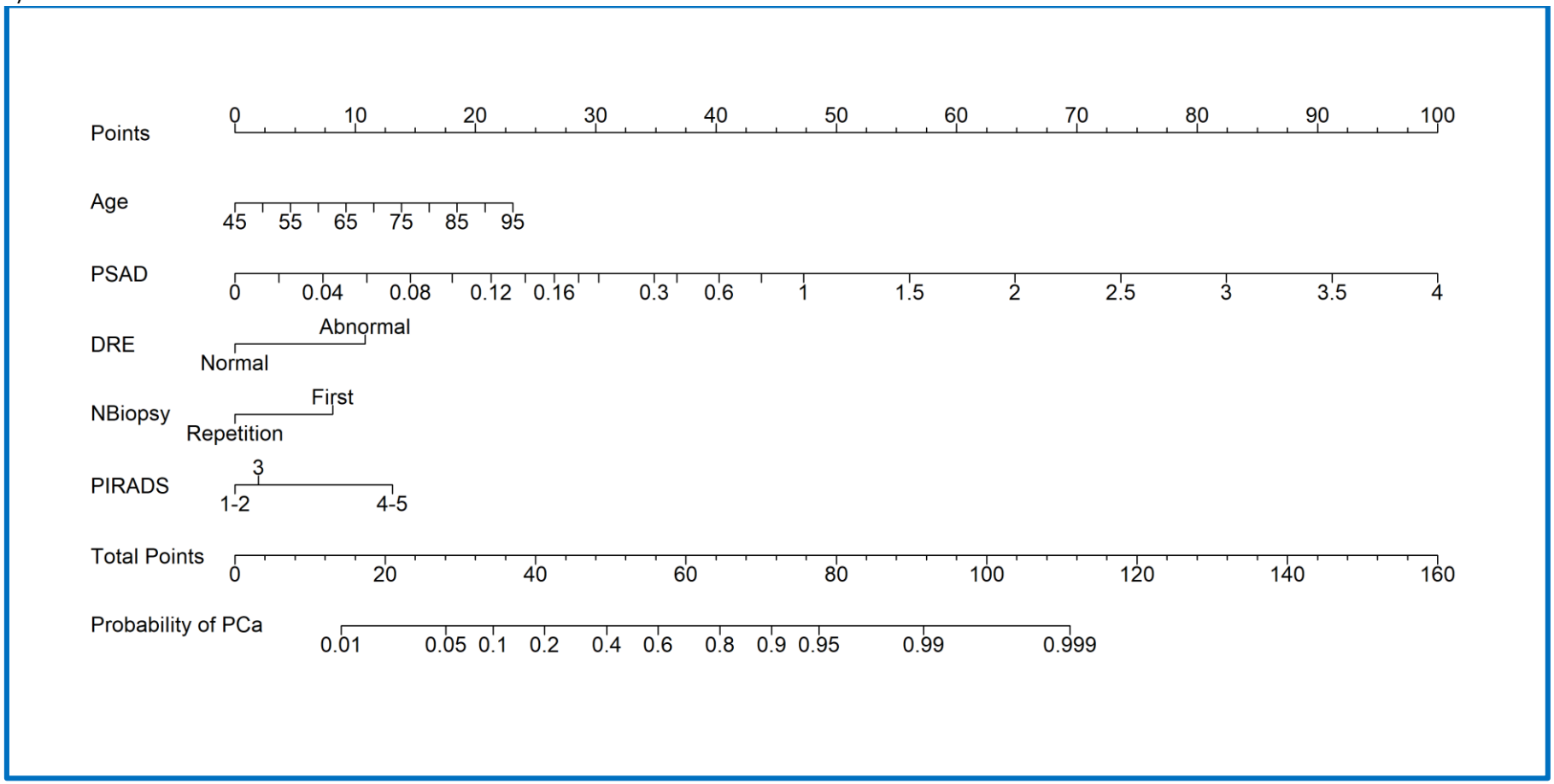
- Figure 4: Clinical utility curves of different predictive models. These graphics show how depending on the threshold of probability for decision-making, we can avoid a percentage of biopsies (red line) and the corresponding risk of missing undetected diagnoses (blue line); a) all PCa prediction, b) ISUP $\geq$ 2 prediction, c) ISUP $\geq$ 3 prediction.

- Figure 5: Probability density functions and clinical utility curves of ISUP$\geq$2 prediction for first PBx and repeated PBx.

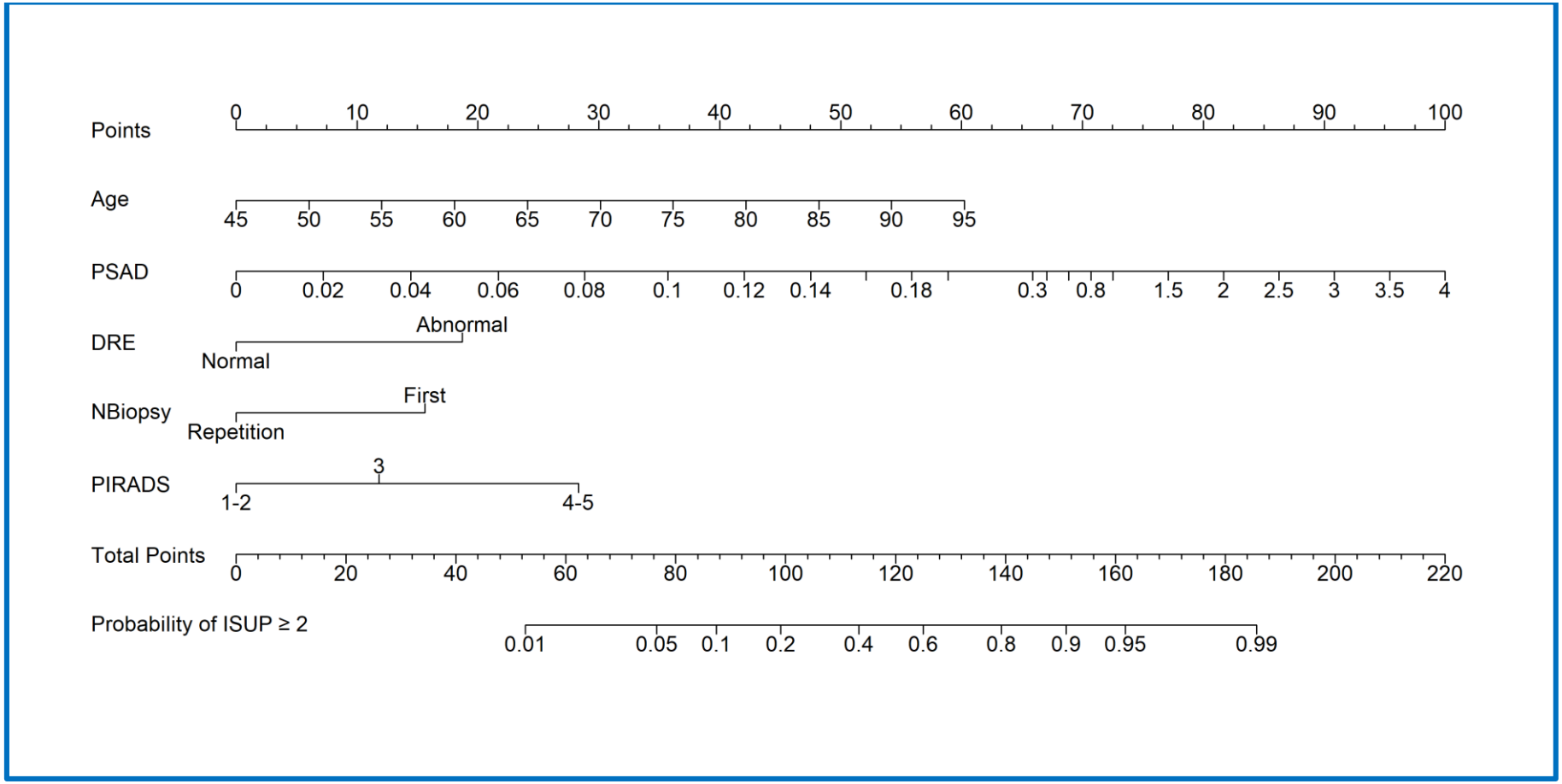
Figure 1 Click here to access/download;Figure;FINAL_Figure1_revised.docx

Figure 1: Nomograms of a) all prostate cancer prediction; b) ~~significant prostate cancer~~ ISUP group 2-5, Gleason score ≥ 7 prediction; and c) ~~aggressive significant prostate cancer~~ ISUP group 3-5, primary Gleason pattern ≥ 4 prediction

a)



b)


Points
0
10
20
30
40
50
60
70
80
90
100
Age
45
50
55
60
65
70
75
80
85
90
95
PSAD
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.18
0.3
0.8
1.5
2
2.5
3
3.5
4
DRE
Abnormal
Normal
NBiopsy
First
Repetition
PIRADS
3
1-2
4-5
Total Points
0
20
40
60
80
100
120
140
160
180
200
220
Probability of ISUP ≥ 2
0.01
0.05
0.1
0.2
0.4
0.6
0.8
0.9
0.95
0.99

c)

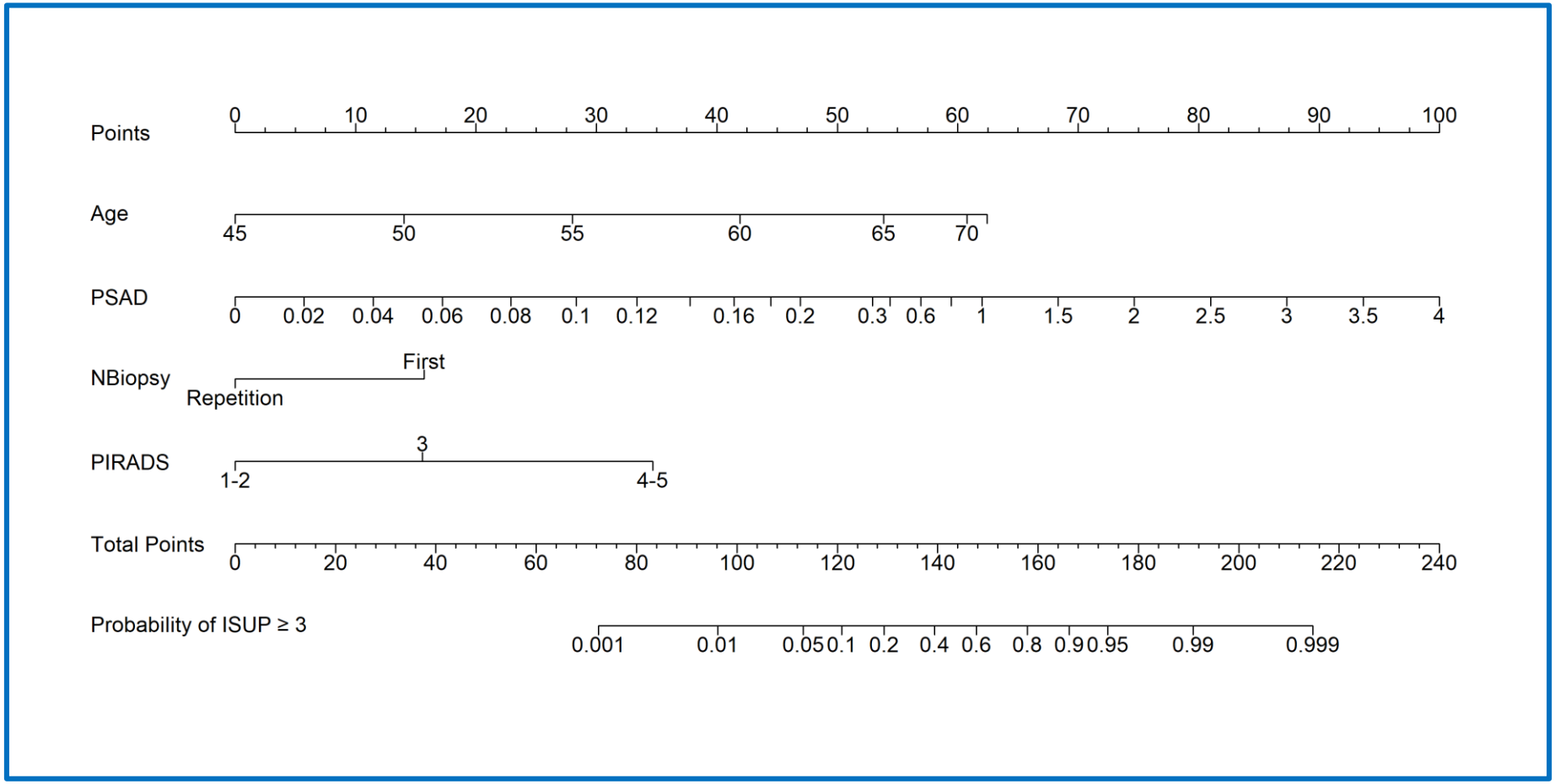

Points
0 10 20 30 40 50 60 70 80 90 100
Age
45 50 55 60 65 70
PSAD
0 0.02 0.04 0.06 0.08 0.1 0.12 0.16 0.2 0.3 0.6 1 1.5 2 2.5 3 3.5 4
NBiopsy
First
Repetition
PIRADS
3
1-2
4-5
Total Points
0 20 40 60 80 100 120 140 160 180 200 220 240
Probability of ISUP ≥ 3
0.001 0.01 0.05 0.1 0.2 0.4 0.6 0.8 0.9 0.95 0.99 0.999

Figure 2

Click here to access/download;Figure;Figure2_revised.tiff

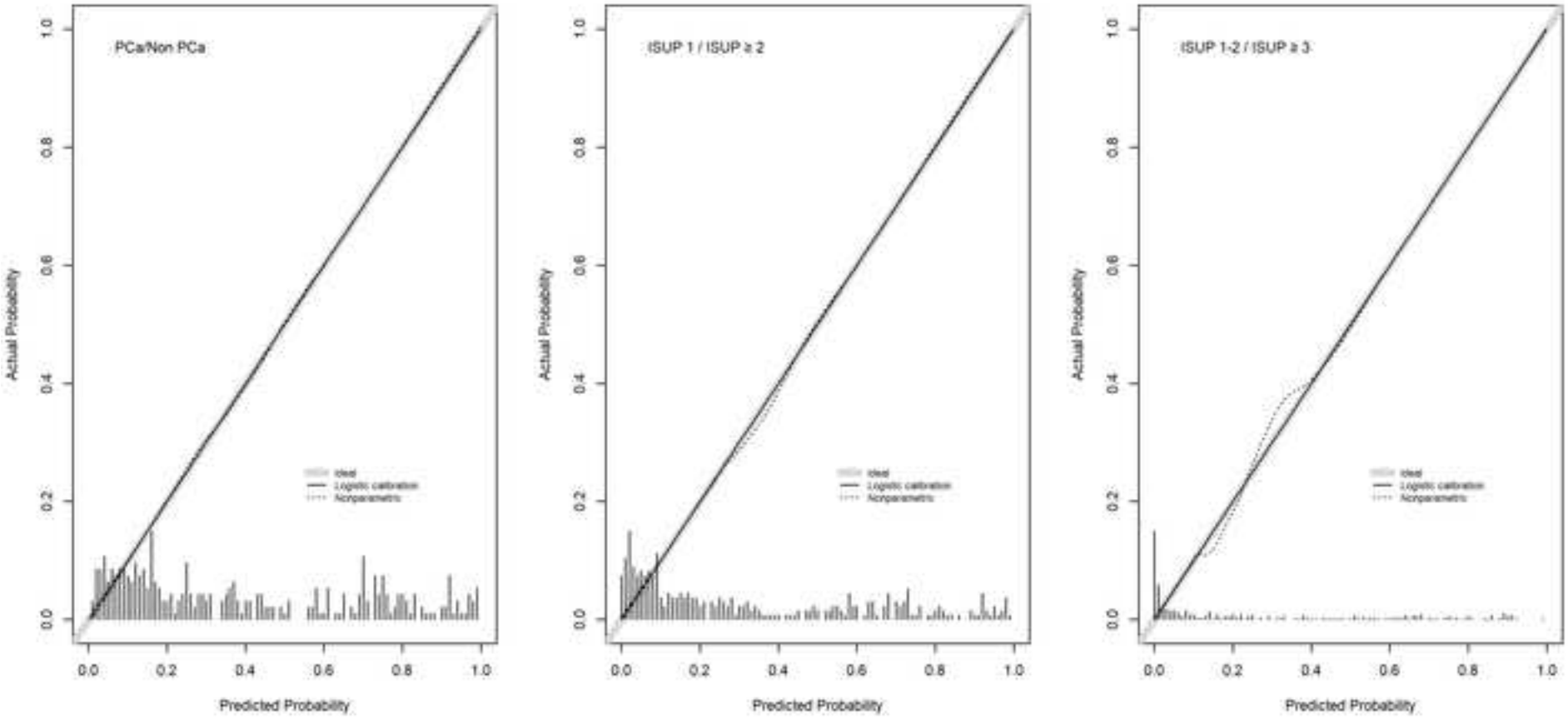


Figure 3

Click here to access/download;Figure;Figure3_revised.tiff

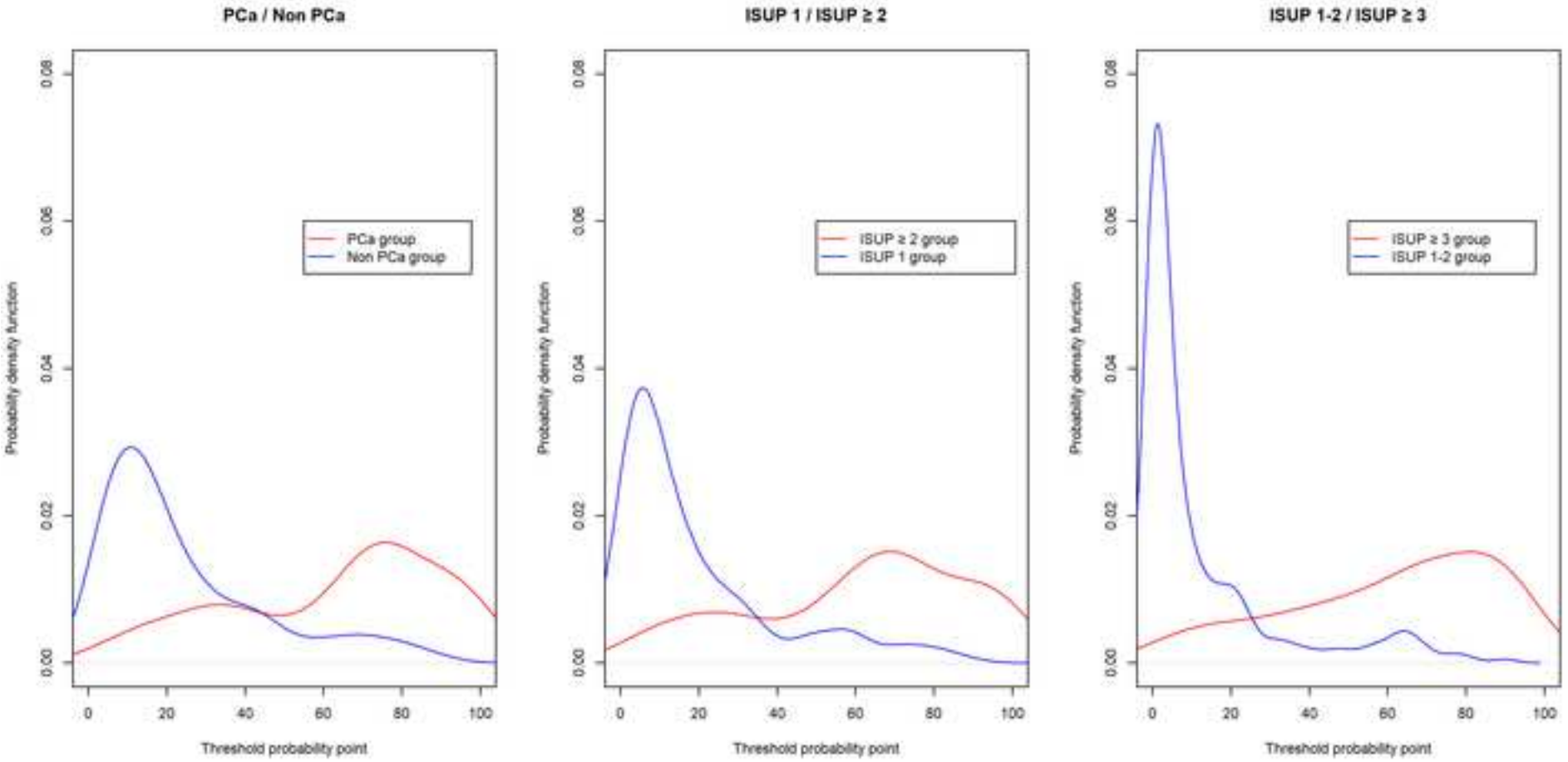
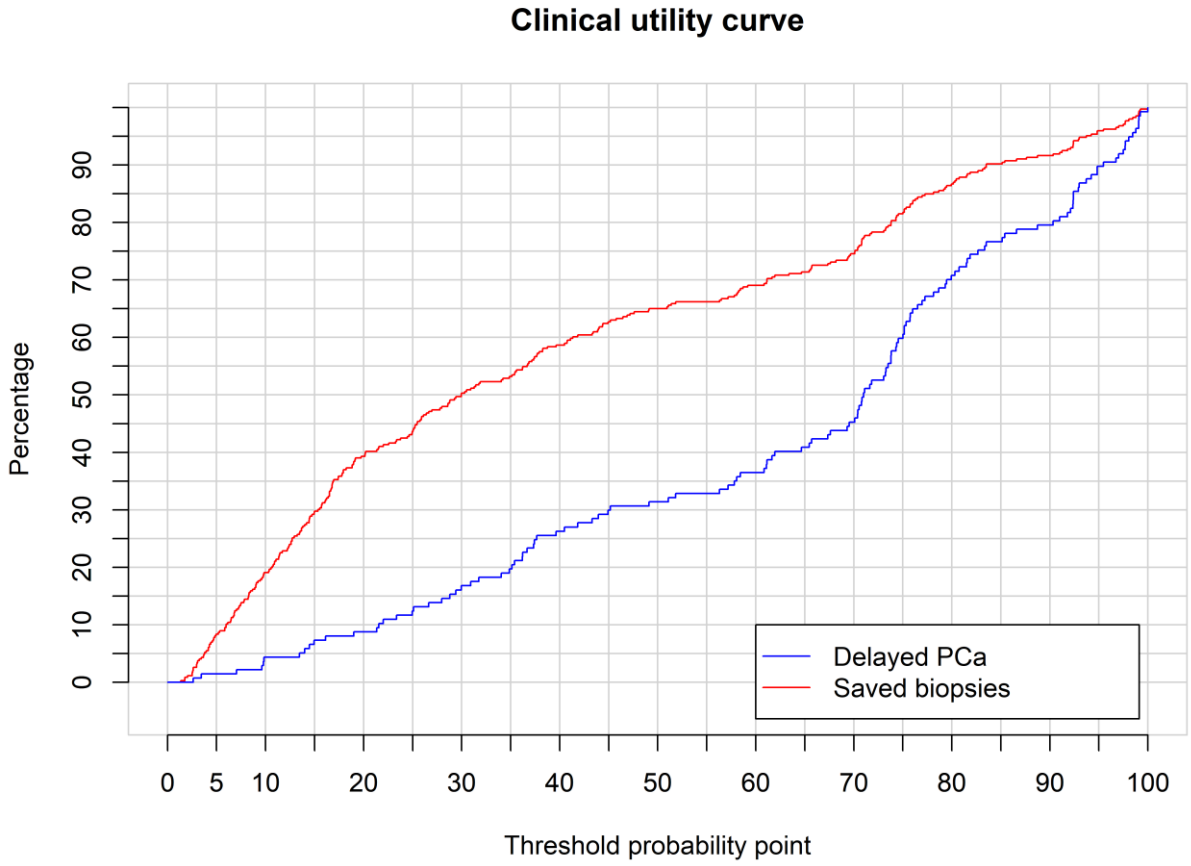


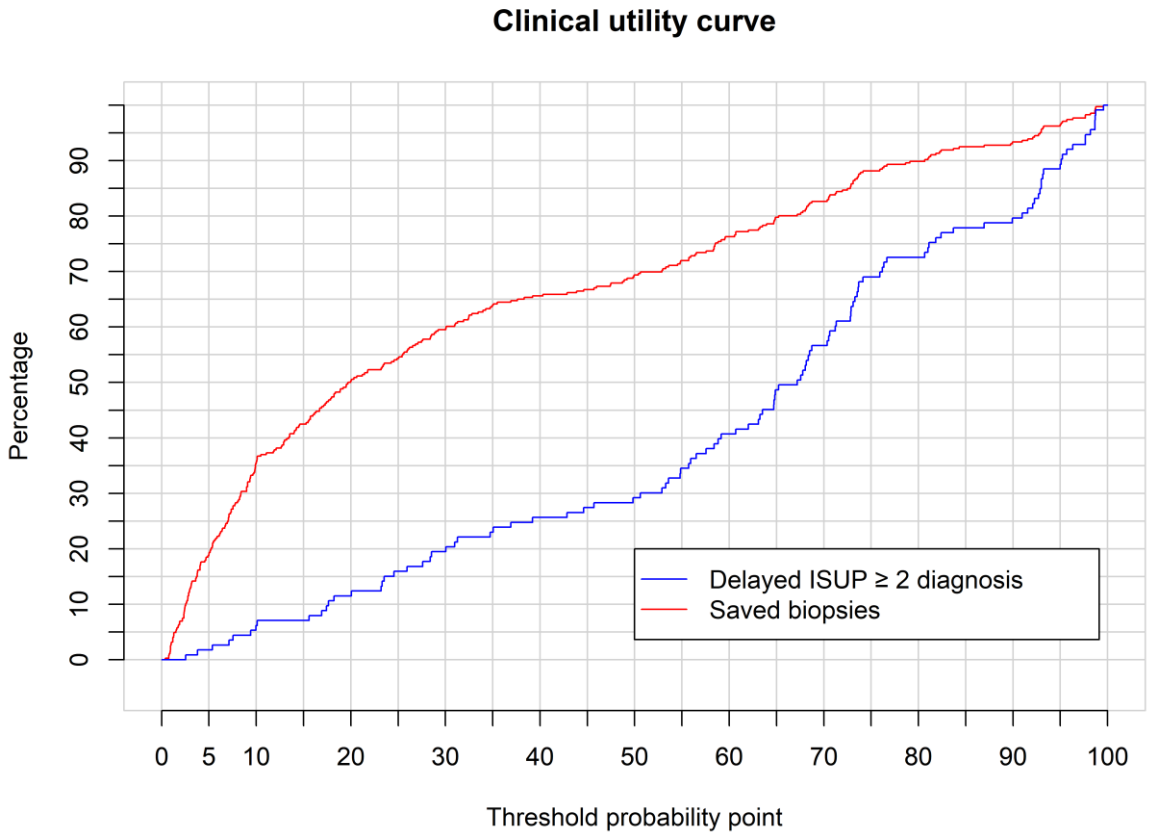
Figure 4

Figure 4: Clinical utility curves of different predictive models. These graphics show how depending on the threshold of probability for decision-making, we can avoid a percentage of biopsies (red line) and the corresponding risk of missing undetected diagnoses (blue line); a) any PCa prediction, b) ISUP ≥ 2 prediction, c) ISUP ≥ 3 prediction.

a)



b)


Clinical utility curve
Percentage
0
10
20
30
40
50
60
70
80
90
0
5
10
20
30
40
50
60
70
80
90
100
Threshold probability point
Delayed ISUP ≥ 2 diagnosis
Saved biopsies

c)

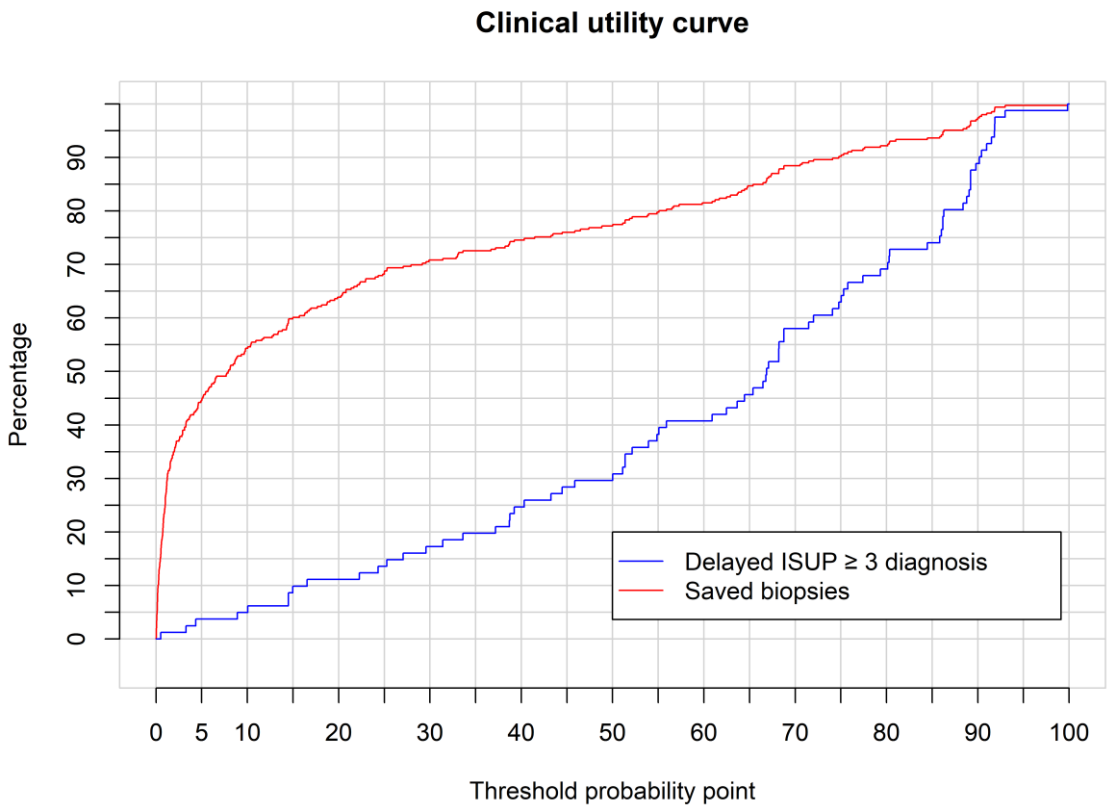

Clinical utility curve
Percentage
0
10
20
30
40
50
60
70
80
90
0
5
10
20
30
40
50
60
70
80
90
100
Threshold probability point
Delayed ISUP ≥ 3 diagnosis
Saved biopsies

Figure 5: Probability density functions and clinical utility curves of ISUP ≥ 2 prediction for first PBx and repeated PBx.

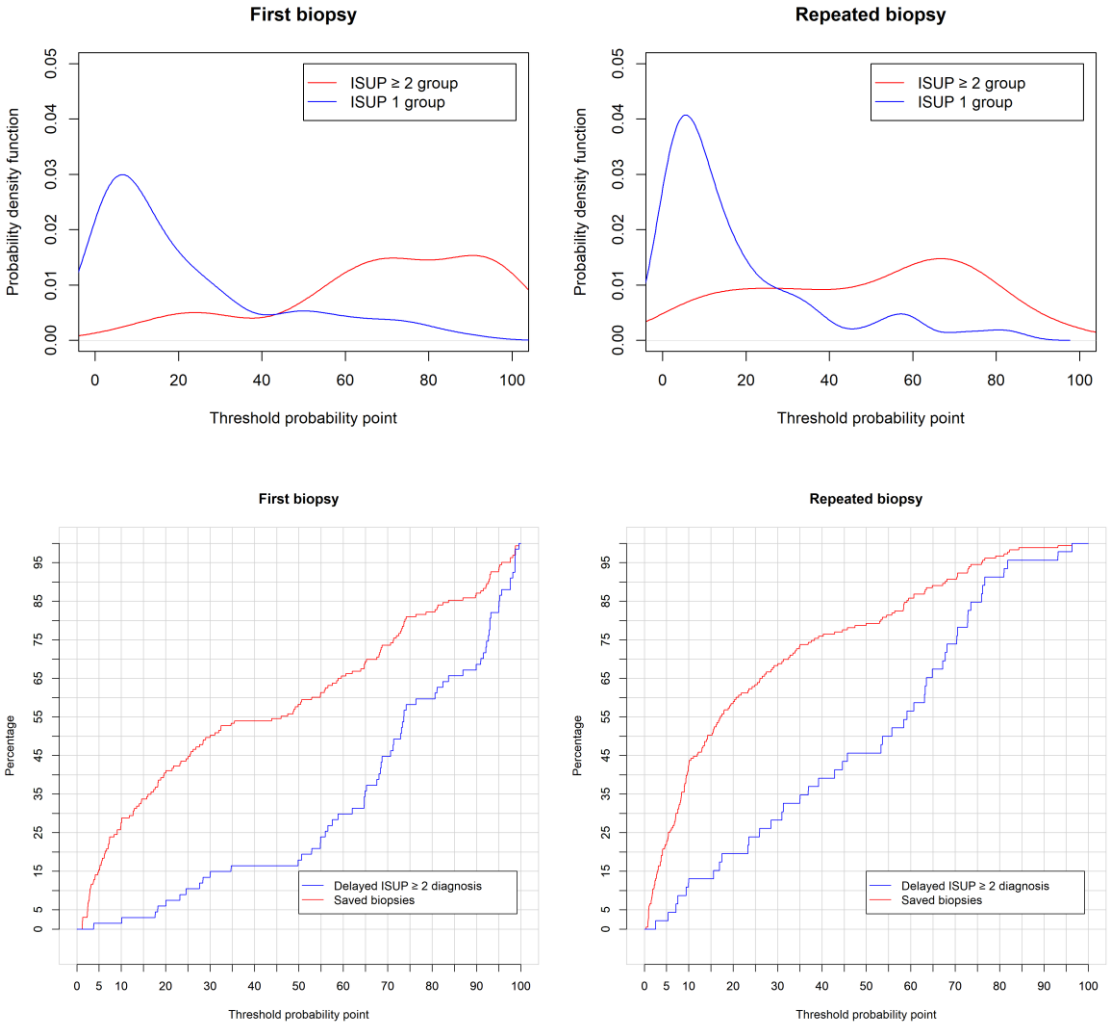


Table 1

Table 1: Clinical and demographics characteristics: Total sample stratified by predicted prostate cancer groups.

VARIABLE	All PCa	Non PCa		ISUP ≥ 2	Non PCa-ISUP 1		ISUP ≥ 3	Non PCa-ISUP 1-2		Total
CONTINUOUS	**p50(p25-p75)**	**p50(p25-p75)**	**p-value**	**p50(p25-p75)**	**p50(p25-p75)**	**p-value**	**p50(p25-p75)**	**p50(p25-p75)**	**p-value**	**p50(p25-p75)**
Age (years)	70.5(65-76.1)	66.2(60.7-71.5)	<.0001	72.4(65.5-77.1)	66.2(60.3-71.2)	<.0001	73.1(68-77.9)	66.2(60.7-71.2)	<.0001	67.7 (63.0-73.7)
PSA (ng/mL)	8(5-12.7)	5.7(4.4-8.3)	<.0001	8.8(5.2-13.8)	5.7(4.5-8.4)	<.0001	9.6(6.2-17.3)	5.7(4.4-8.4)	<.0001	6.1(4.7-10.2)
fPSA*	1.15(0.78-1.65)	1.18(0.8-1.59)	0.67	1.07(0.7-1.59)	1.19(0.81-1.62)	0.58	1.23(0.83-1.74)	1.16(0.79-1.54)	0.62	1.17(0.79-1.63)
Prostate volume (mL.) (MRI)	41(34-57)	57(45-81)	<.0001	40(34-57)	56(43-80)	<.0001	40(33-57)	55(41-79)	<.0001	50.5(38-72.8)
PSA density (ng/mL/cc)	0.17(0.13-0.30)	0.10(0.07-0.16)	<.0001	0.19(0.14-0.33)	0.11(0.07-0.16)	<.0001	0.22(0.16-0.37)	0.12(0.07-0.17)	<.0001	0.13(0.08-0.20)
CATEGORICAL	**N(%)**	**N(%)**		**N(%)**	**N(%)**		**N(%)**	**N(%)**		
	137	**209**		**113**	**233**		**81**	**265**		**346**
Number of biopsy			0.004			0.002			0.004	
• First	78 (57%)	85 (41%)		67 (59%)	96 (41%)		50 (62%)	113 (43%)		163 (47%)
• Repetition	59 (43%)	124(59%)		46 (41%)	137 (59%)		31 (38%)	152 (57%)		183 (53%)
DRE										
• Normal	80 (58%)	157 (75%)	0.002	64 (57%)	173 (74%)	0.001	49 (60%)	188 (71%)	0.102	237 (68%)
• Abnormal	57 (42%)	52 (25%)		49 (43%)	60 (26%)		32 (40%)	77 (29%)		109 (32%)
PI-RADS v.2.2015										
• 1-2	12 (9%)	66 (32%)	<.0001	6 (5%)	72 (31%)	<.0001	1 (1%)	77 (29%)	<.0001	78 (23%)
• 3	18 (13%)	74 (35%)		14 (12%)	78 (33%)		5 (6%)	87 (33%)		92 (27%)
• 4-5	107 (78%)	69 (33%)		93 (82%)	83 (36%)		75 (93%)	101 (38%)		176 (50%)

PCa: prostate cancer at biopsy; ISUP: International Society of Uropathologists grade group; p25-p50-p75: percentile 25, percentile 50 (median), percentile 75, respectively; PSA: prostate specific antigen; fPSA: free PSA *(not available in 105 patients, if there was total PSA below 4 ng/mL or over 10 ng/mL); DRE: digital rectal examination; PI-RADS v.2.2015

Table 2: Multivariate models of different prostate cancer predictions

	All PCa / Non PCa		ISUP ≥ 2 / Non PCa-ISUP 1		ISUP ≥ 3 / Non PCa-ISUP 1-2	
VARIABLE	**Odds ratio (C.I. 95%)**	**p-value**	**Odds ratio (C.I. 95%)**	**p-value**	**Odds ratio (C.I. 95%)**	**p-value**
Age	1.75 (1.19-2.56)	<.0001	2.33 (1.54-3.53)	<.0001	3.78 (1.93-7.43)	0.0010 (linear)
	-------		--------			0.0084 (no linear)
PSA density	5.96 (3.28-10.85)	<.0001 (linear)	7.22 (3.68-14.17)	<.0001 (linear)	9.34 (4.06-21.48)	<.0001 (linear)
		0.0003 (no linear)		<.0001(no linear)		<.0001(no linear)
DRE: • Abnormal/Normal (reference)	3.61 (1.86-7.01)	0.0001	3.64 (1.79-7.41)	0.0004	2.99 (1.36-6.59)	0.0065
Number of biopsy: • First/Repeated (reference)	2.62 (1.47-4.69)	0.0012	2.94 (1.57-5.52)	0.0008	--------	-------
PI-RADS v2.2015: • 1-2/4-5 (reference)	0.21 (0.10-0.46)	<.0001	0.14 (0.05-0.37)	<.0001	0.05(0.01-0.36)	0.0034
• 3/4-5 (reference)	0.27 (0.13-0.53)	0.0002	0.32 (0.15-0.67)	0.0025	0.15 (0.06-0.43)	0.0004

Supplementary Figure 1