# A novel nomogram to identify candidates for active surveillance among patients with ISUP Grade Group 1 or ISUP Grade Group 2 prostate cancer, according to multiparametric magnetic resonance imaging findings

Stefano Luzzago[1,2], Ottavio de Cobelli[1,6], Gabriele Cozzi[1], Giulia Peveri[3], Vincenzo Bagnardi[4], Michele Catellani[1], Ettore Di Trapani[1], Francesco A. Mistretta[1,2], Paola Pricolo[5], Andrea Conti[1,2], Sarah Alessi[5], Giulia Marvaso[8], Matteo Ferro[1], Deliu-Victor Matei[1], Giuseppe Renne[7], Barbara Jereczek-Fossa[8,6], Giuseppe Petralia[5,9], Gennaro Musi[1]

[1]*Department of Urology, IEO European Institute of Oncology, IRCCS, Via Ripamonti 435, Milan, Italy*

[2]*Università degli Studi di Milano, Milan, Italy*

[3]*Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy*

[4]*Department of Statistics and Quantitative Methods, Università degli Studi di Milano-Bicocca, Milan, Italy*

[5]*Department of Radiology, IEO European Institute of Oncology, IRCCS, Via Ripamonti 435, Milan, Italy*

[6]*Università degli Studi di Milano, Department of Oncology and Hematology-Oncology, Milan, Italy*

[7]*Department of Pathology, IEO European Institute of Oncology, IRCCS, Via Ripamonti 435, Milan, Italy*

[8]*Department of Radiation Oncology, IEO European Institute of Oncology, IRCCS, Via Ripamonti 435, Milan, Italy*

[9]*Precision Imaging and Research Unit – Department of Department of Medical Imaging and Radiation Sciences, IEO European Institute of Oncology IRCCS, Via Ripamonti 435, Milan, Italy*

Corresponding author:
Stefano Luzzago, MD
Department of Urology, IEO European Institute of Oncology, IRCCS, Milan, Italy
Via Giuseppe Ripamonti, 435
20141 Milan, Italy
Tel: +39 33354249298
E-mail: stefanoluzzago@gmail.com

Word count (abstract): 420

Word count (manuscript): 2847

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/bju.15048

**Abstract**

*Objectives*

To develop a novel nomogram that combines clinical, biopsy and multiparametric magnetic resonance imaging (mpMRI) findings and to compare its predictive accuracy to, respectively: 1) Prostate Cancer Research International: Active Surveillance (PRIAS) criteria, 2) Johns Hopkins (JH) criteria, 3) EAU low risk classification and 4) EAU low risk or low volume ISUP GG2 classification.

*Materials and Methods*

Overall, we selected 1 837 patients with ISUP GG 1 or ISUP GG2 prostate cancer that were treated with radical prostatectomy (RP) between 2012 and 2018. The outcome of interest was the presence of unfavourable disease (csPCa) at RP, defined as: ISUP GG≥3 and/or pT≥3a and/or pN1. First, logistic regression models including PRIAS, JH, EAU low risk and EAU low risk or low volume ISUP GG2 binary classifications (not eligible vs. eligible) were used. Second, a multivariable logistic regression model including age, PSA-D, ISUP GG and the percentage of positive cores (Model 1) was fitted. Third, PI-RADS score (Model 2), extracapsular score (ECE) (Model 3) and PI-RADS + ECE score (Model 4) were added to Model 1. Only variables associated with higher csPCa rates in Model 4 were retained in the final simplified Model 5. The area under the ROC-curve (AUC), calibration plots and decision-curve analyses were used.

*Results*

Of 1 837 patients, 775 (42.2%) presented csPCa at RP. Overall, 837 (47.5%), 986 (53.7%), 348 (18.9%) and 209 (11.4%) patients were eligible to AS according to, respectively, low risk classification, low risk or low volume ISUP GG2 classification, PRIAS criteria and JH criteria.

The proportion of csPCa among low risk, low risk or low volume ISUP GG2, PRIAS and JH candidates was, respectively 28.5%, 29.3%, 25.6% and 17.2%. Model 4 and Model 5 (in which only PSA-D, ISUP GG, PI-RADS and ECE score were retained) had greater AUC (0.84), compared to the four proposed AS criteria (all $p<0.001$). The adoption of a 25% nomogram threshold increased the proportion of AS eligible patients from 18.9% (PRIAS) and 11.4% (JH) to 44.4%. Moreover, the same 25% nomogram threshold resulted in significantly lower estimated risks of csPCa (11.3%), compared to PRIAS (Δ:-14.3%), JH (Δ:-5.9%), low risk (Δ:-17.2%) and low risk or low volume ISUP GG2 (Δ:-18.0%) classifications.

*Conclusion*

A novel nomogram that combines clinical, biopsy and mpMRI findings is able to increase of approximately 25% and 35% the absolute frequency of patients suitable for AS, compared to, respectively, PRIAS or JH criteria. Moreover, this nomogram significantly reduces the estimated frequency of csPCa that would be recommended for AS compared to, respectively, PRIAS, JH, EAU low risk and EAU low volume ISUP GG2 classifications.

**Keywords:** Active surveillance; International Society of Urological Pathology Grade Groups; Multiparametric-magnetic resonance imaging; Nomogram; Prostate Cancer

DR. STEFANO LUZZAGO (Orcid ID : 0000-0002-9373-1413)

Article type : Original Article

Article Category: Urological Oncology

**Introduction**

Active surveillance (AS) is considered the recommended strategy for patients with low-risk prostate cancer (PCa) (1,2). Moreover, AS has been proposed as a valid alternative to active treatment in patients with low volume International Society of Urological Pathology Grade Group 2 (ISUP GG2) PCa. Several cohorts with long-term follow-up tested the oncological safety of AS, relative to active treatments (3–5). These prospective studies enrolled patients according to clinical (namely: Prostate Specific Antigen [PSA], clinical stage [cT] and PSA-density [PSA-D]) and biopsy features. However, pathological findings of AS candidates treated with radical prostatectomy (RP) showed in this population subgroup a 10-25% risk of high-grade and/or extraprostatic PCa (6,7). Several nomograms failed to show any superiority in selecting patients with low-grade and gland-confined PCa (8–11). Even more, attempts have been made in order to increase the number of AS-suitable patients, without increasing disease misclassification risk (2). Recently, Gandaglia et al. (12) developed a risk calculator which combines clinical and biopsy information. The application of this novel risk score resulted in an absolute increase of ~10% in the number of patients eligible for AS, if compared to PRIAS criteria, without any significant increase in the risk of aggressive disease. Another tool widely exploited to improve selection of AS candidates (13) is multiparametric magnetic resonance imaging (mpMRI), that is now recommended at the time of AS confirmatory biopsy (14,15).

We hypothesized that a novel nomogram that combines clinical, biopsy and mpMRI findings, would allow to include a substantial number of patients in AS, without negative effects on its oncological safety.

## Materials and methods

*Study population*

This retrospective study was conducted according to the ethical guidelines of the Declaration of Helsinki. From 2012 to 2018, 2 513 patients with ISUP GG1 or ISUP GG2 (16) PCa at biopsy were treated with RP +/- lymph node dissection at a tertiary centre. Patients with missing information (n=387) or that underwent a prostate biopsy with <8 cores taken (n=91) were excluded. Furthermore, patients without mpMRI were not considered (n=198). Overall, 1 837 patients represented the study cohort.

*Multiparametric magnetic resonance imaging technique*

MpMRI was performed on a 1.5-T MR scanner (Avanto, Siemens Medical Solutions, Erlangen, Germany) with a phased-array coil. MpMRI protocol involved: 1) sagittal, coronal and axial T2-weighted images; 2) axial diffusion weighted images; 3) dynamic series of axial T1-weighted images obtained before, during and after injection of gadopentetate dimeglumine (Magnevist, Bayer HealthCare, Berlin, Germany). All images were analysed by one of three dedicated radiologists (G.P, S.A. and P.P.), who were not blinded from clinical findings. A score PI-RADS (PI-RADS v1 till November 2015, and PI-RADS v2 thereafter) was assigned to each suspicious lesion (17,18). The probability of extraprostatic extension was also evaluated using a 5-point scale (ECE score), according to the 2012 ESUR prostate MR guidelines (17).

*Variables definition and outcomes*

All RP specimens were reviewed by a single dedicated pathologist (G.R.). The TNM stage was applied according to the 2002 American Joint Committee on Cancer staging system and the ISUP GG in accordance to ISUP 2014 consensus conference (16). ISUP GG was retrospectively assigned for patients treated before 2015. Clinical stage was assigned after digital rectal examination (DRE). Prostate volume was measured by mpMRI. Biopsies were performed with a 12 cores trans-perineal template, that involved sampling of the entire peripheral zone of the prostate. Additional transition zone cores were taken according to clinician's preference. For patients with positive mpMRI (PI-RADS score≥3), 1-3 cognitive targeted-cores were additionally taken, according to clinician's preference.

The outcome of interest was the presence of csPCa at RP, defined as ISUP GG≥3 and/or extraprostatic disease (pT≥3a) and/or node-positive PCa (pN1) (19).

*Statistical analyses*

We evaluated the added value of mpMRI parameters (i.e. PI-RADS and ECE score) in predicting the risk of csPCa with respect to: 1) Prostate Cancer Research International: Active Surveillance criteria (PRIAS: ISUP GG1, cT≤2a, PSA≤10 ng/ml, PSA-D<0.2 ng/mL/mL and 1 or 2 positive cores) (3); 2) Johns Hopkins criteria (JH: ISUP GG1, cT1c, PSA<10 ng/ml, PSA-D<0.15 ng/mL/mL and 1 or 2 positive cores with ≤50% cancer in any core) (4); 3) EAU Guidelines low risk classification (ISUP GG1, cT≤2a and PSA<10 ng/ml) (2); 4) EAU Guidelines low risk or low volume ISUP GG2 (< 10% pattern 4) classification (2) and 5) a multivariable model including age, PSA-D, ISUP GG and the percentage (%) of positive cores.

Differences in the distribution of continuous and categorical variables were evaluated with, respectively, the t-test and the chi-square test. Logistic regression models including binary classifications as single predictors (namely: PRIAS, JH, low risk, low risk or low volume ISUP GG2,) and a multivariable model including age, PSA-D, ISUP GG at biopsy and % of positive cores (Model 1) were firstly fitted. Subsequently, PI-RADS score (Model 2), ECE score (Model 3), and PI-RADS + ECE score (Model 4) were added to Model 1. Finally, only statistically significant predictors of csPCa in Model 4 were retained in a simplified model (Model 5). The predictive accuracy of the mentioned models was evaluated with respect to discrimination (i.e. the ability of the model to classify a patient with csPCa from one without csPCa) and calibration (i.e. the agreement between the outcome frequencies observed in the data and the predicted probabilities of the model). Discrimination was measured by the area under the ROC curve (AUC). AUC differences were evaluated with the nonparametric approach of DeLong, DeLong, and Clarke-Pearson (20). Calibration was evaluated with the loess plot and comparing the observed proportions of patients with csPCa vs. the predicted risk of csPCa. To reduce overfit bias and for internal validation, the model was subjected to 2 000 bootstrap resamples. Decision-curve analyses (DCA) tested the clinical net benefit associated with the use of the best performing model. To measure the effect of varying decision threshold derived from the best performing model, we evaluated the sensitivity, specificity, number of csPCa cases that would be missed and number of

patients that would be enrolled in AS. A sensitivity analysis was conducted to evaluate whether the association between PI-RADS score and the outcome varied according to PI-RADS versions v1 vs. v2. The interaction terms between the system version and PI-RADS score were added to the best performing model and tested for significance using a likelihood ratio test (LRT).

All statistical analyses were performed using SAS and R software. All reported p-values were two sided.

## Results

*Descriptive analyses*

Of 1 837 patients, 775 (42.2%) presented csPCa at RP (Table 1). Of those, 599 (77.3%) 446 (57.5%) and 81 (10.5%) harboured pT≥3a, ISUP GG≥3 and pN1, respectively. Patients with csPCa at RP were older (65 vs. 63 years; p<0.001) and presented higher PSA (6.8 vs. 5.9 ng/ml) and PSA-D (0.19 vs. 0.14 ng/ml/ml; p<0.001). Moreover csPCa patients exhibited greater % of positive cores (31 vs. 21%; p<0.001) and smaller prostates (38 vs. 43 ml; p<0.001). PI-RADS 5 and ECE score 4-5 lesions accounted for, respectively, 460 (59.4%) vs. 195 (18.4%) and 483 (62.3%) vs. 125 (11.8%) in csPCa vs. non-csPCa patients (p<0.001). Last but not least, csPCa patients more frequently harboured ISUP GG2 PCa at biopsy (52.1 vs. 31.1%; p<0.001). Overall, 837 (47.5%), 986 (53.7%), 348 (18.9%) and 209 (11.4%) patients were eligible to AS according to, respectively, low risk classification, low risk or low volume ISUP GG2 classification, PRIAS criteria and JH criteria. The proportion of csPCa among low risk, low risk or low volume ISUP GG2, PRIAS and JH candidates was, respectively 28.5%, 29.3%, 25.6% and 17.2%.

*Multivariable logistic regression models*

Odds ratios (OR) estimates for all variables included in the considered models are reported in Table 2. In simplified model 5, significant predictors of csPCa were: PSA-D (OR: 1.16; p<0.001), ISUP GG2 (OR: 1.58; p<0.001), PI-RADS score (PI-RADS 3 OR: 1.78, PI-RADS 4 OR: 2.99 and PI-RADS 5 OR: 4.23; all p<0.001) and ECE score (ECE 3 OR: 4.72 and ECE 4-5 OR: 14.35; all p<0.001).

In Figure 1, ROC curves are shown for low risk, low risk or low volume ISUP GG2, PRIAS and JH candidates, as well as for each of the five multivariable models. AUC for low risk, low risk or low volume ISUP GG2, PRIAS and JH patients were, respectively, 0.63, 0.64, 0.56 and 0.56. All proposed models had a significant greater AUC, compared to the four previous mentioned classifications (all p<0.001). Moreover, all proposed models including at least one mpMRI feature (Models 2-5) had a significantly higher AUC, compared to Model 1 (all p<0.001). Among those, Model 4 and Model 5 had the greatest AUCs (0.84), compared to Model 1 (AUC: 0.70; p<0.001), Model 2 (AUC 0.80; p<0.001) and Model 3 (AUC: 0.82, p=0.04). Model 5 was therefore selected as the best performing model. As shown in Figure 2, the model exhibited a good agreement between observed

and predicted outcomes. A nomogram to calculate the individual probability of csPCa, based on Model 5, is shown in Figure 3. The DCA (Supplementary Figure 1) showed that the adoption of a risk score based on Model 5 improved clinical risk prediction, as compared to the strategy of treating all patients and to low risk, low risk or low volume ISUP GG 2, PRIAS and JH classifications, starting approximately at a threshold probability of 5%.

*Threshold to predict unfavourable disease*

The proportion of patients eligible for AS according to different nomogram thresholds, and the proportion of csPCa patients, are shown in Figure 4.
The adoption of a 25% nomogram threshold increased the proportion of AS eligible patients from 18.9% (PRIAS) and 11.4% (JH) to 44.4%. Moreover, the adoption of the same 25% threshold resulted in a modest reduction of the percentage of AS eligible patients, relative to low risk (Δ:-3.1%) and low risk or low volume ISUP GG 2 classifications (Δ:-9.3%). However, the same 25% nomogram threshold resulted in significantly lower estimated risks of csPCa (11.3%), compared to PRIAS (Δ:-14.3%), JH (Δ:-5.9%), low risk (Δ:-17.2%) and low risk or low volume ISUP GG2 (Δ:-18.0%) classifications. Last but not least, the sensitivity and the specificity of this nomogram derived 25% threshold were, respectively, 85% and 73%.

*Sensitivity analysis*

The prognostic value of PI-RADS score did not significantly differ according to the PI-RADS version used (Model 5 p-value for the likelihood ratio test: 0.56) (Supplementary table 1).

**Discussion**

Active Surveillance protocols aim to identify patients who truly harbour csPCa (3,5). Nonetheless, pathological series of AS candidates treated with RP showed a risk of csPCa up to 25% (6,7). Stricter AS inclusion criteria appear not to be the solution. Komisarenko et al. (21) reported no significant differences in adverse pathology rates in men who underwent RP, according to different AS protocols. Moreover, existing nomograms showed only modest accuracy in predicting the outcomes of patients followed with AS (8,9). Therefore, an important unmet need is how to increase the number of patients suitable for AS without increasing the risk of csPCa (2). We thus developed a novel nomogram combining clinical, biopsy and mpMRI findings.

First, we reported significant rates of csPCa among AS candidates according to four different established inclusion criteria. Specifically, the rates of csPCa in PRIAS vs. JH vs. low risk vs. low risk or low volume ISUP GG2 patients were, respectively, 25.6% vs. 17.2% vs. 28.5% vs. 29.3%. In previous series, the four mentioned inclusion criteria (2–4) showed good reliability, even though more than 20% of enrolled patients shifted from AS to definitive treatment within the first two years. Our findings, although consistent with other reports, probably represent the worst-case scenario. Moreover, our cohort was recruited in the period when AS was widely used, which might have led to uncontrolled selection bias. Therefore, the aforementioned criteria still need to be considered reliable in the enrolment of AS patients.

Second, of all clinical and biopsy variables tested in multivariable logistic regression models predicting csPCa at RP, only PSA-D and ISUP GG retained independent predictor status. Several ongoing AS protocols (3,4) consider PSA-D cut-offs between 0.15 and 0.2 ng/ml/ml to assess patient's eligibility. Moreover, PSA-D remains a significant predictor of csPCa among AS candidates, even in the recent mpMRI era (22). Additionally, previous series (5) showed significantly worse outcomes for ISUP GG 2 vs. ISUP GG 1 AS patients during follow-up.

Third, all models (Models 2-5) that included at least one mpMRI feature, namely PI-RADS and ECE score, had significantly higher AUCs compared to PRIAS, JH, low risk, low risk or low volume ISUP GG2 classifications and Model 1. Recently, Sanguedolce et al. (23) retrospectively studied 135 patients selected for AS and submitted to baseline mpMRI. Authors demonstrated that PI-RADS score (HR 3.2) was significantly associated with failure free survival over time. The ESUR 2012 MRI guidelines (17) proposed the

adoption of a 5-point scale (ECE score) to assess the probability of PCa extracapsular extension. Although these recommendations have not been yet confirmed by PI-RADS v2 (18), ECE score represented the strongest predictor of csPCa in multivariable logistic regression models. Yoo et al. (24) reported a 11 greater risk of adverse histology in patients with low-risk PCa and suspicious upstaging on mpMRI. Based on these considerations, we suggest the evaluate ECE score among AS candidates, before considering their eligibility.

Fourth, the adoption of a mpMRI-based nomogram (Model 5) significantly increased the accuracy of PRIAS, JH, low risk and low risk or low volume ISUP GG2 classifications in determining optimal AS candidates. Available pathological nomograms (8,9) failed to achieve superiority to AS criteria in selecting patients with low-grade organ-confined PCa. Moreover, to the best of our knowledge, some mpMRI risk-tools (25–27) were only used to reduce the rate of misclassified disease, without evaluating the possibility of increasing the number of AS eligible patients. Recently, Gandaglia et al. (12) combined PSA, cT, ISUP GG, number of positive cores and PSA-D to develop a novel risk calculator to increase the number of AS eligible patients. With an AUC value of 0.75, Gandaglia et al. (12) nomogram allowed to include ~10% more patients in AS compared to PRIAS criteria. We therefore combined clinical, biopsy and mpMRI parameters to create a novel nomogram to further increase the number of AS eligible patients, without increasing the risk of csPCa. With an AUC of 0.84 and according to a risk-threshold value of 25%, our nomogram resulted in an absolute increase of approximately 25% and 35% of the absolute frequency of AS suitable patients, compared to PRIAS and JH criteria, respectively. Moreover, the same 25% risk-threshold value guarantees a minimum reduction in the number of AS eligible patients, relative to low risk (Δ:-3.1%) and low risk or low volume ISUP GG2 (Δ:-9.3%) categories. Last but not least, the mentioned nomogram risk-threshold value of 25% was associated with significantly lower estimated frequency of csPCa that would be recommended for AS (11.3%), relative to the same four established AS inclusion criteria: PRIAS (Δ:-14.3%), JH (Δ:-5.9%), low risk (Δ:-17.2%) and low risk or low volume ISUP GG2 (Δ:-18.0%). These csPCa estimated rates are consistent with previous reports (6,7) and with Gandaglia et al. (12) findings, which reported a minimum csPCa rate of 12-15%, despite the use of several AS inclusion criteria. However, it should be stated that our results were obtained by testing the predictive model in the population where it was developed. To reduce overfit bias

and for internal validation, the model was subjected to 2000 bootstrap resamples. Nonetheless, results could be less impressive once this instrument is applied to a different population (i.e. validated externally).

Fifth, sensitivity analyses did not show different results according to the PI-RADS version used. This last check, further confirms the reliability of our risk tool and of PI-RADS scoring system (17,18) in predicting csPCa at RP.

Despite its strengths, our study presents several limitations. First, the current data are retrospective and influenced by inherent selection bias. In consequence, a substantial proportion of men with highly mpMRI suspicious lesions (PI-RADS score 4-5) and a lower proportion of patients with non-suspicious mpMRI lesions (PI-RADS score≤3) were treated with RP. Second, the accuracy of our model was not compared with other AS selection criteria that are widely used in clinical practice (2). However, Komisarenko et al. (21) previously reported no significant differences in the rates of csPCa at RP according to different inclusion criteria. Third, we did not assess the inter-reader agreement among mpMRI readers. Despite Moldovan et al. demonstrated that the negative predictive value varies significantly between readers (28), other studies reported moderate-to-high concordance rates between expert radiologists (29). Fourth, cognitive-targeted cores were taken according to clinician's preference and not with a standard targeted-biopsy protocol. Moreover, we were unable to perform mpMRI-US fusion targeted-biopsies since this technology was unavailable during the study period. Despite cognitive targeted-biopsies could be associated with incorrect sampling of suspicious lesions, previous studies failed to demonstrate a significant advantage of any MRI-guided biopsy technique over the others for csPCa detection (30). Fifth, as previously stated, lack of external validation currently limits predictions concerning the possible impact of our nomogram in clinical practice.

**Conclusion**

A novel nomogram that combines clinical, biopsy and mpMRI findings is able to increase of approximately 25% and 35% the absolute frequency of patients suitable for AS, compared to, respectively, PRIAS or JH criteria. Moreover, this nomogram significantly reduces the estimated frequency of csPCa that would be recommended for AS compared to, respectively, PRIAS, JH, EAU low risk and EAU low volume ISUP GG2 classifications.

**Acknowledgments**: This work was partially supported by the Italian Ministry of Health with Ricerca Corrente and 5x1000 funds

**Conflicts of interest:** None

**Funding:** None

**References**


1. Moschini M, Carroll PR, Eggener SE, Epstein JI, Graefen M, Montironi R, et al. Low-risk Prostate Cancer: Identification, Management, and Outcomes. European Urology. 2017.
2. Briganti A, Fossati N, Catto JWF, Cornford P, Montorsi F, Mottet N, et al. Active Surveillance for Low-risk Prostate Cancer: The European Association of Urology Position in 2018. Eur Urol. 2018;
3. Bokhorst LP, Valdagni R, Rannikko A, Kakehi Y, Pickles T, Bangma CH, et al. A Decade of Active Surveillance in the PRIAS Study: An Update and Evaluation of the Criteria Used to Recommend a Switch to Active Treatment. Eur Urol. 2016;70(6):954–60.
4. Tosoian JJ, Trock BJ, Landis P, Feng Z, Epstein JI, Partin AW, et al. Active surveillance program for prostate cancer: An update of the Johns Hopkins experience. International Braz J Urol. 2011.
5. Klotz L, Vesprini D, Sethukavalan P, Jethava V, Zhang L, Jain S, et al. Long-term follow-up of a large active surveillance cohort of patients with prostate cancer. J Clin Oncol. 2015;33(3):272–7.
6. Iremashvili V, Pelaez L, Manoharan M, Jorda M, Rosenberg DL, Soloway MS. Pathologic prostate cancer characteristics in patients eligible for active surveillance: A head-to-head comparison of contemporary protocols. Eur Urol. 2012;
7. Conti SL, Dall'Era M, Fradet V, Cowan JE, Simko J, Carroll PR. Pathological Outcomes of Candidates for Active Surveillance of Prostate Cancer. J Urol. 2009;
8. Iremashvili V, Manoharan M, Parekh DJ, Punnen S. Can nomograms improve our ability to select candidates for active surveillance for prostate cancer? Prostate Cancer Prostatic Dis. 2016;
9. Wang SY, Cowan JE, Clint Cary K, Chan JM, Carroll PR, Cooperberg MR. Limited ability of existing nomograms to predict outcomes in men undergoing active surveillance for prostate cancer. BJU Int. 2014;
10. Kattan MW, Eastham JA, Wheeler TM, Maru N, Scardino PT, Erbersdobler A, et al. Counseling men with prostate cancer: A nomogram for predicting the presence of small, moderately differentiated, confined tumors. J Urol. 2003;
11. Truong M, Slezak JA, Lin CP, Iremashvili V, Sado M, Razmaria AA, et al. Development and multi-institutional validation of an upgrading risk tool for Gleason 6 prostate cancer. Cancer. 2013;

12. Gandaglia G, van den Bergh RCN, Tilki D, Fossati N, Ost P, Surcel CI, et al. How can we expand active surveillance criteria in patients with low- and intermediate-risk prostate cancer without increasing the risk of misclassification? Development of a novel risk calculator. BJU Int. 2018;
13. Schoots IG, Petrides N, Giganti F, Bokhorst LP, Rannikko A, Klotz L, et al. Magnetic resonance imaging in active surveillance of prostate cancer: A systematic review. Eur Urol [Internet]. 2015;67(4):627–36. Available from: http://dx.doi.org/10.1016/j.eururo.2014.10.050
14. 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. 2017;71(4):618–29.
15. NCCN. NCCN guidelines : Prostate cancer. Surg. 2016;
16. Epstein JI, Zelefsky MJ, Sjoberg DD, Nelson JB, Egevad L, Magi-Galluzzi C, et al. A Contemporary Prostate Cancer Grading System: A Validated Alternative to the Gleason Score. Eur Urol. 2016;
17. Barentsz JO, Richenberg J, Clements R, Choyke P, Verma S, Villeirs G, et al. ESUR prostate MR guidelines 2012. Eur Radiol. 2012 Apr;22(4):746–57.
18. 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. 2016 Jan;69(1):16–40.
19. Gandaglia G, Ploussard G, Isbarn H, Suardi N, De Visschere PJL, Futterer JJ, et al. What is the optimal definition of misclassification in patients with very low-risk prostate cancer eligible for active surveillance? Results from a multi-institutional series. Urol Oncol Semin Orig Investig. 2015;
20. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;
21. Komisarenko M, Timilshina N, Richard PO, Alibhai SMH, Hamilton R, Kulkarni G, et al. Stricter Active Surveillance Criteria for Prostate Cancer do Not Result in Significantly Better Outcomes: A Comparison of Contemporary Protocols. J Urol. 2016;
22. Alberts AR, Roobol MJ, Drost FJH, van Leenders GJ, Bokhorst LP, Bangma CH, et al. Risk-stratification based on magnetic resonance imaging and prostate-specific antigen density may reduce unnecessary follow-up biopsy procedures in men on active

surveillance for low-risk prostate cancer. BJU Int. 2017;

23. Sanguedolce F, Petralia G, Sokhi H, Tagliabue E, Anyamene N, Hellawell G, et al. Baseline Multiparametric MRI for Selection of Prostate Cancer Patients Suitable for Active Surveillance: Which Features Matter? Clinical Genitourinary Cancer. 2017;
24. Yoo S, Hong JH, Byun S-S, Lee JY, Chung BH, Kim C-S. Is suspicious upstaging on multiparametric magnetic resonance imaging useful in improving the reliability of Prostate Cancer Research International Active Surveillance (PRIAS) criteria? Use of the K-CaP registry. Urol Oncol Semin Orig Investig. 2017;1–7.
25. Stamatakis L, Siddiqui MM, Nix JW, Logan J, Rais-Bahrami S, Walton-Diaz A, et al. Accuracy of multiparametric magnetic resonance imaging in confirming eligibility for active surveillance for men with prostate cancer. Cancer. 2013;
26. Siddiqui MM, Truong H, Rais-Bahrami S, Stamatakis L, Logan J, Walton-Diaz A, et al. Clinical implications of a multiparametric magnetic resonance imaging based nomogram applied to prostate cancer active surveillance. J Urol. 2015;
27. Shukla-Dave A, Hricak H, Akin O, Yu C, Zakian KL, Udo K, et al. Preoperative nomograms incorporating magnetic resonance imaging and spectroscopy for prediction of insignificant prostate cancer. BJU Int. 2012;
28. Moldovan PC, Van den Broeck T, Sylvester R, Marconi L, Bellmunt J, van den Bergh RCN, et al. What Is the Negative Predictive Value of Multiparametric Magnetic Resonance Imaging in Excluding Prostate Cancer at Biopsy? A Systematic Review and Meta-analysis from the European Association of Urology Prostate Cancer Guidelines Panel. Eur Urol. 2017;
29. Luzzago S, Petralia G, Musi G, Catellani M, Alessi S, Di Trapani E, et al. Multiparametric Magnetic Resonance Imaging Second Opinion May Reduce the Number of Unnecessary Prostate Biopsies: Time to Improve Radiologists' Training Program? Clinical Genitourinary Cancer. 2018;
30. Wegelin O, van Melick HHE, Hooft L, Bosch JLHR, Reitsma HB, Barentsz JO, et al. Comparing Three Different Techniques for Magnetic Resonance Imaging-targeted Prostate Biopsies: A Systematic Review of In-bore versus Magnetic Resonance Imaging-transrectal Ultrasound fusion versus Cognitive Registration. Is There a Preferred Technique? European Urology. 2017.

## Figure legends

**Figure 1.** ROC curves for comparisons between considered models

PRIAS: Prostate Cancer Research International: Active Surveillance
JH: Johns Hopkins

**Figure 2.** Actual and bias-corrected calibration plot of the best performing model, including PSA-D, ISUP Grade Group at biopsy, PI-RADS and ECE score.

PSA-D: Prostate Specific Antigen Density
ISUP: International Society of Urological Pathology
PI-RADS: Prostate Imaging Reporting and Data System
ECE: Extracapsular Extension

**Figure 3.** Nomogram for risk of unfavourable disease (csPCa: ISUP group ≥ 3 and/or pT≥3a and/or pN1) among patients with ISUP Grade Group 1 or ISUP Grade Group 2 prostate cancer treated with radical prostatectomy between 2012 and 2018. The value for each factor (PSA-D, ISUP Grade Group, PI-RADS and ECE score) corresponds to points vertically above on the top scale. The three-point values are added together to determine the total points, which then correspond to estimated risk of csPCa shown on the scales below.

PSA-D: Prostate Specific Antigen Density
ISUP: International Society of Urological Pathology
PI-RADS: Prostate Imaging Reporting and Data System
ECE: Extracapsular Extension

**Figure 4.** Proportion of patients potentially eligible for active surveillance and proportion of pathologically unfavourable disease according to low risk, low risk or low volume ISUP grade group 2, PRIAS or JH criteria and risk calculator thresholds.

ISUP: International Society of Urological Pathology
PRIAS: Prostate Cancer Research International: Active Surveillance
JH: Johns Hopkins

**Supplementary Figure 1.** Decision Curve Analysis showing the net benefit associated with the use of the best performing model (Model 5) in the detection of unfavourable disease (csPCa: ISUP group ≥ 3 and/or pT≥3a and/or pN1).

Violet line: best performing model
Blue line: low risk
Red line: low risk or low volume ISUP grade group 2
Brown line: Johns Hopkins criteria
Green line: Prostate Cancer Research International: Active Surveillance
Grey line: criteria assuming that all patients have unfavourable disease and undergo active surveillance.
Black line: criteria assuming that no patients have unfavourable disease and undergo active surveillance.

PRIAS: Prostate Cancer Research International: Active Surveillance
JH: Johns Hopkins

**Table 1:** Descriptive characteristics of 1 837 patients with ISUP Grade Group 1 or ISUP Grade Group 2 prostate cancer treated with radical prostatectomy between 2012 and 2018, stratified accordingly to the presence or absence of unfavourable disease (csPCa) at final pathology (ISUP Grade Group ≥ 3 and/or pT≥3a and/or pN1).
Data are shown as medians for continuous variables or as counts and percentages (%) for categorical variables.

| | | Overall N=1 837 | non csPCa N=1 062 (57.8%) | csPCa N=775 (42.2%) | *p* |
|---|---|---|---|---|---|
| **Age, median (IQR)** | | 64 (59-69) | 63 (58-68) | 65 (59-69) | *<0.001* |
| **PSA, ng/mL median (IQR)** | | 6.3 (4.8-8.7) | 5.9 (4.5-7.9) | 6.8 (5.1-10.0) | *<0.001* |
| **PSA-D, ng/mL/mL cc median (IQR)** | | 0.15 (0.11-0.23) | 0.14 (0.10-0.19) | 0.19 (0.13-0.27) | *<0.001* |
| **Number of cores taken, median (IQR)** | | 14 (12-15) | 14 (12-15) | 13 (12-15) | *0.25* |
| **Number of positive cores, median (IQR)** | | 3 (2-6) | 3 (2-5) | 4 (2-6) | *<0.001* |
| **% of positive cores, median (IQR)** | | 25 (14-42) | 21 (13-36) | 31 (17-50) | *<0.001* |
| **Maximum % of core involvement, median (IQR)** | | 30 (10-60) | 25 (10-50) | 40 (20-70) | *<0.001* |
| **cT, N (%)** | T1 | 1184 (64.5) | 773 (72.8) | 411 (53.0) | *<0.001* |
| | T2a | 415 (22.6) | 198 (18.6) | 217 (28.0) | |
| | T2b | 135 (7.3) | 54 (5.1) | 81 (10.5) | |
| | T2c | 35 (1.9) | 22 (2.1) | 13 (1.7) | |
| | T3 | 68 (3.7) | 15 (1.4) | 53 (6.8) | |
| **ISUP grade group at biopsy, N (%)** | 1 | 1103 (60.0) | 732 (68.9) | 371 (47.9) | *<0.001* |
| | 2 | 734 (40.0) | 330 (31.1) | 404 (52.1) | |
| **PI-RADS, N (%)** | 1 | 11 (0.6) | 10 (0.9) | 1 (0.1) | *<0.001* |
| | 2 | 104 (5.7) | 99 (9.3) | 5 (0.6) | |
| | 3 | 375 (20.4) | 326 (30.7) | 49 (6.3) | |
| | 4 | 692 (37.7) | 432 (40.7) | 260 (33.5) | |
| | 5 | 655 (35.7) | 195 (18.4) | 460 (59.4) | |
| **ECE, N (%)** | 1 | 326 (17.7) | 314 (29.6) | 12 (1.5) | *<0.001* |
| | 2 | 518 (28.2) | 426 (40.1) | 92 (11.9) | |
| | 3 | 385 (21.0) | 197 (18.5) | 188 (24.3) | |
| | 4 | 425 (23.1) | 99 (9.3) | 326 (42.1) | |
| | 5 | 183 (10.0) | 26 (2.4) | 157 (20.3) | |
| **mpMRI prostate volume (ml), median (IQR)** | | 41 (32-57) | 43 (33-61) | 38 (30-50) | *<0.001* |
| **pT, N (%)** | T0 | 6 (0.3) | 6 (0.6) | 0 (0.0) | *<0.001* |
| | T2 | 1232 (67.1) | 1056 (99.4) | 176 (22.7) | |
| | T3a | 476 (25.9) | 0 (0.0) | 476 (61.4) | |
| | T3b | 123 (6.7) | 0 (0.0) | 123 (15.9) | |
| **pN, N (%)** | N0 | 902 (49.1) | 457 (43.0) | 445 (57.4) | *<0.001* |
| | N1 | 81 (4.4) | 0 (0.0) | 81 (10.5) | |
| | NX | 854 (46.5) | 605 (57.0) | 249 (32.1) | |
| **ISUP grade group at pathology, N (%)** | 0 | 6 (0.3) | 6 (0.6) | 0 (0.0) | *<0.001* |
| | 1 | 502 (27.3) | 450 (42.4) | 52 (6.7) | |

| | | | | | |
|---|---|---|---|---|---|
| | 2 | 883 (48.1) | 606 (57.1) | 277 (35.7) | |
| | 3 | 347 (18.9) | 0 (0.0) | 347 (44.8) | |
| | 4 | 62 (3.4) | 0 (0.0) | 62 (8.0) | |
| | 5 | 37 (2.0) | 0 (0.0) | 37 (4.8) | |
| **Low risk, N (%)** | Not eligible for AS | 964 (52.5) | 438 (41.2) | 526 (67.9) | *<0.001* |
| | Eligible for AS | 873 (47.5) | 624 (58.8) | 249 (32.1) | |
| **Low risk or low volume ISUP grade group 2, N (%)** | Not eligible for AS | 851 (46.3) | 365 (34.4) | 486 (62.7) | *<0.001* |
| | Eligible for AS | 986 (53.7) | 697 (65.6) | 289 (37.3) | |
| **PRIAS, N (%)** | Not eligible for AS | 1489 (81.1) | 803 (75.6) | 686 (88.5) | *<0.001* |
| | Eligible for AS | 348 (18.9) | 259 (24.4) | 89 (11.5) | |
| **JH, N (%)** | Not eligible for AS | 1628 (88.6) | 889 (83.7) | 739 (95.4) | *<0.001* |
| | Eligible for AS | 209 (11.4) | 173 (16.3) | 36 (4.6) | |

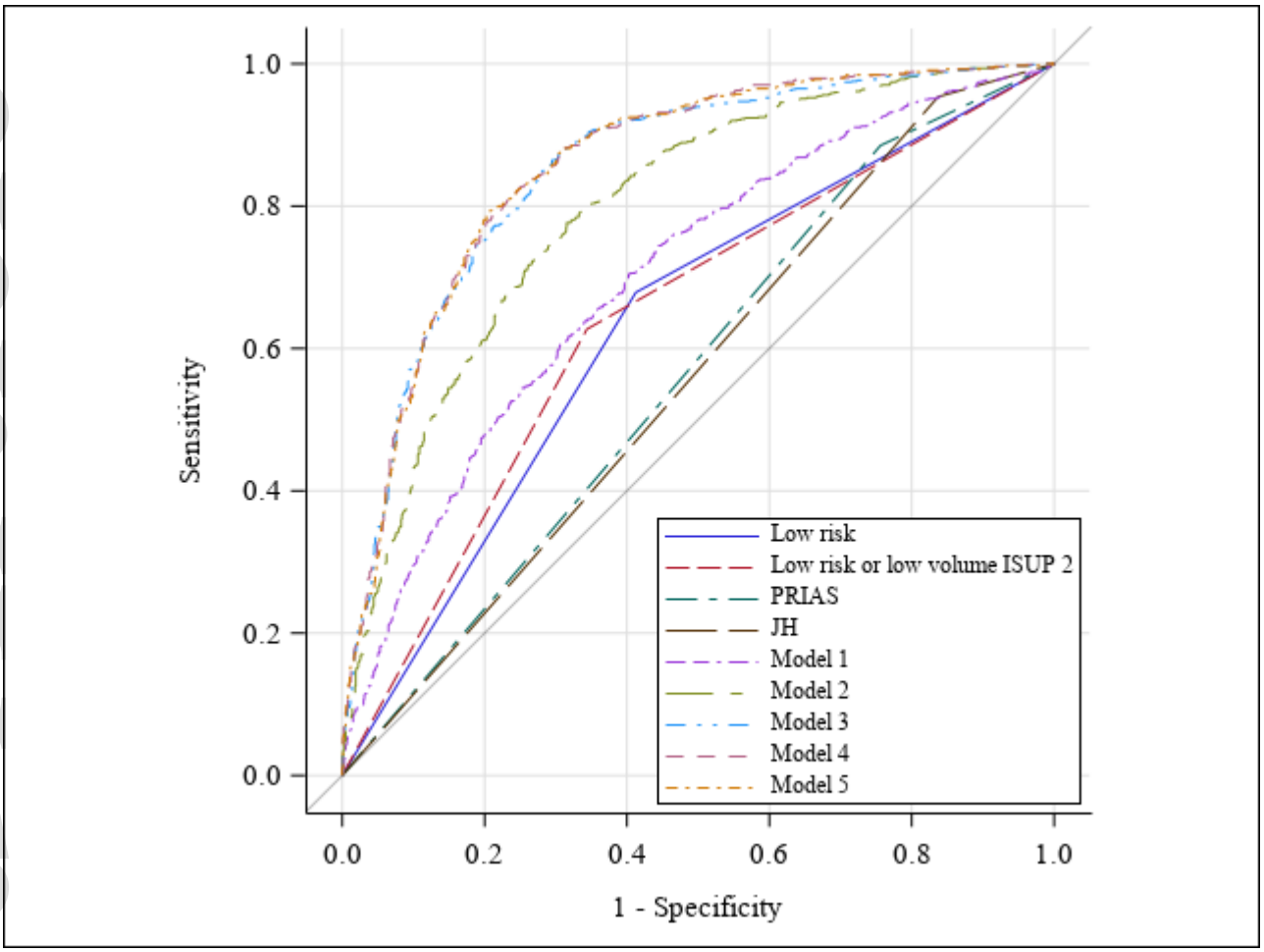
IQR: interquartile range
PSA: Prostate Specific Antigen
PSA-D: PSA density
ISUP: International Society of Urological Pathology
PI-RADS: Prostate Imaging - Reporting and Data System
ECE: Extracapsular extension
mpMRI: multiparametric magnetic resonance imaging
cT: clinical T stage
pT: pathological T stage
pN: pathological N stage
PRIAS: Prostate Cancer Research International Active Surveillance
JH: Johns Hopkins
AS: Active Surveillance

**Table 2.** Multivariate logistic models evaluating the association between selected covariables and the risk of unfavourable disease (csPCa: ISUP Grade Group ≥ 3 and/or pT≥3a and/or pN1). Odds ratios (OR) with 95% confidence interval (CI) are provided.

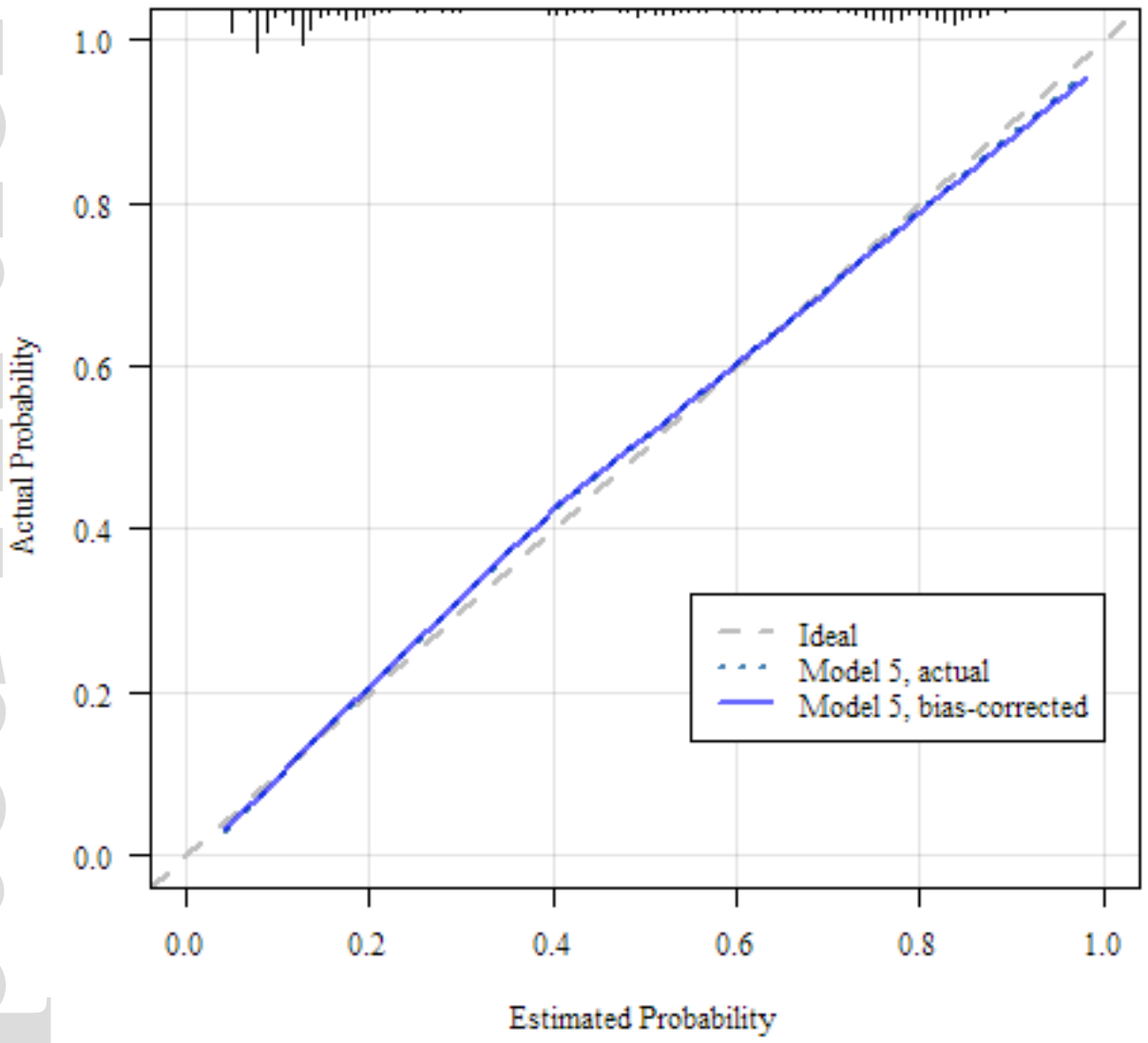
| | | MODEL 1 | | MODEL 2 | | MODEL 3 | | MODEL 4 | | MODEL 5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **Level** | **OR (95% CI)** | ***p*** | **OR (95% CI)** | ***p*** | **OR (95% CI)** | ***p*** | **OR (95% CI)** | ***p*** | **OR (95% CI)** | ***p*** |
| **Age** | + 5 years | 1.16 (1.08-1.24) | *<0.001* | 1.08 (1.00-1.16) | *0.06* | 1.05 (0.96-1.14) | *0.26* | 1.04 (0.95-1.13) | *0.40* | - | - |
| **% of positive cores** | +10% | 1.13 (1.08-1.19) | *<0.001* | 1.07 (1.01-1.13) | *0.02* | 1.06 (1.00-1.12) | *0.07* | 1.04 (0.99-1.11) | *0.14* | - | - |
| **PSA-D** | + 0.05 (ng/mL)/mL | 1.17 (1.12-1.21) | *<0.001* | 1.09 (1.05-1.14) | *<0.001* | 1.09 (1.05-1.14) | *<0.001* | 1.08 (1.04-1.12) | *<0.001* | 1.16 (1.08-1.26) | *<0.001* |
| **ISUP grade group** | 2 vs. 1 | 1.97 (1.6-2.41) | *<0.001* | 1.72 (1.37-2.15) | *<0.001* | 1.55 (1.21-1.98) | *<0.001* | 1.52 (1.19-1.94) | *<0.001* | 1.58 (1.24-2.01) | *<0.001* |
| **PI-RADS** | 3 vs. ≤ 2 | - | - | 2.46 (1.02-5.93) | *<0.001* | - | - | 1.75 (0.72-4.25) | *<0.001* | 1.78 (0.73-4.32) | *<0.001* |
| | 4 vs. ≤ 2 | - | | 8.57 (3.70-19.88) | | - | | 2.87 (1.21-6.82) | | 2.99 (1.26-7.08) | |
| | 5 vs. ≤ 2 | - | | 29.48 (12.66-68.68) | | - | | 4.03 (1.65-9.81) | | 4.23 (1.74-10.28) | |
| **ECE** | 3 vs. ≤ 2 | - | - | - | - | 6.09 (4.55-8.16) | *<0.001* | 4.67 (3.43-6.37) | *<0.001* | 4.72 (3.47-6.44) | *<0.001* |
| | 4-5 vs. ≤ 2 | - | | - | | 22.14 (16.54-29.63) | | 14.05 (9.94-19.86) | | 14.35 (10.16-20.27) | |

| | | Low risk | | Low risk or low volume ISUP grade group 2 | | PRIAS | | JH | |
|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **Level** | **OR (95% CI)** | ***p*** | **OR (95% CI)** | ***p*** | **OR (95% CI)** | ***p*** | **OR (95% CI)** | ***p*** |
| **Eligible for AS** | No vs. Yes | 3.01 (2.48-3.65) | *<0.001* | 3.21 (2.65-3.90) | *<0.001* | 2.49 (1.91-3.23) | *<0.001* | 3.99 (2.75-5.8) | *<0.001* |

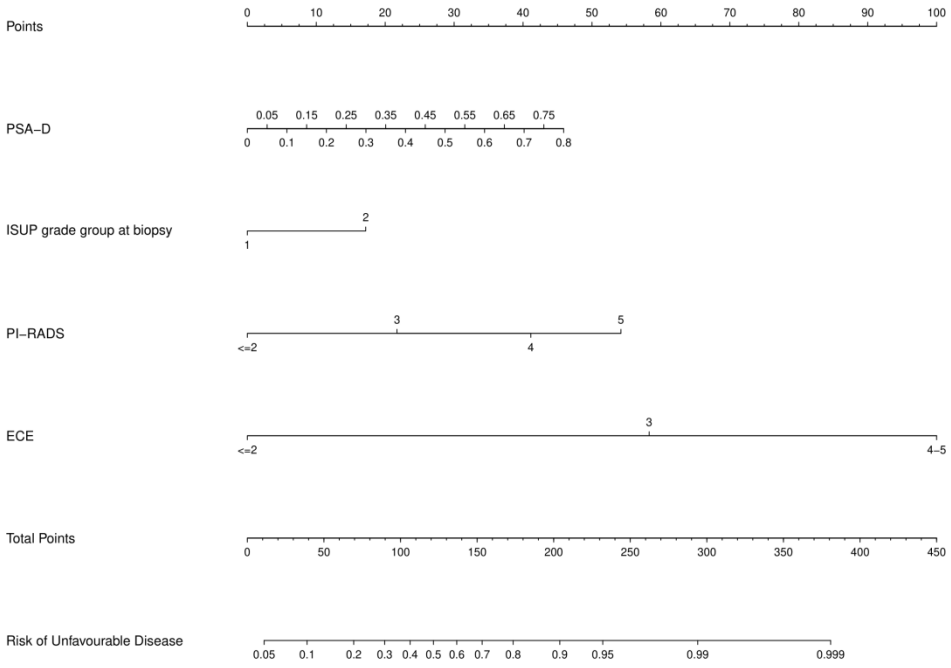
PSA: Prostate Specific Antigen
PSA-D: PSA density
ISUP: International Society of Urological Pathology
PI-RADS: Prostate Imaging Reporting and Data System
ECE: Extracapsular Extension
PRIAS: Prostate Cancer Research International: Active Surveillance
JH: Johns Hopkins
AS: Active Surveillance



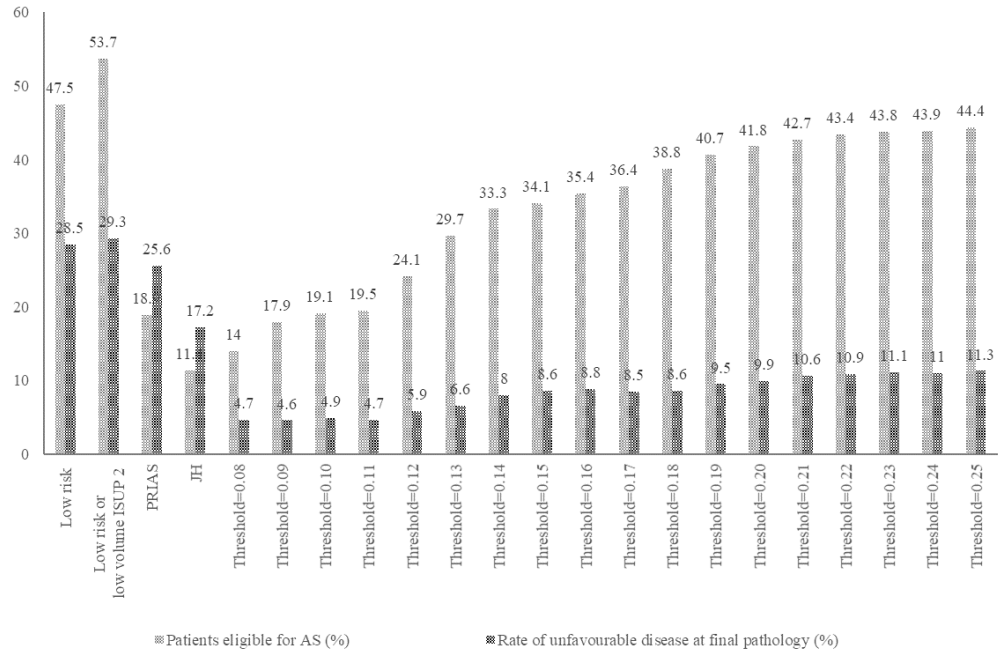

bju_15048_f1.png




bju_15048_f2.png




bju_15048_f3.tiff




bju_15048_f4.png