available at www.sciencedirect.com
journal homepage: www.europeanurology.com

European Association of Urology

Prostate Cancer

Predicting the Need for Biopsy to Detect Clinically Significant Prostate Cancer in Patients with a Magnetic Resonance Imaging–detected Prostate Imaging Reporting and Data System/Likert ≥3 Lesion: Development and Multinational External Validation of the Imperial Rapid Access to Prostate Imaging and Diagnosis Risk Score

Max Peters [a,*], *David Eldred-Evans* [b], *Piet Kurver* [a], *Ugo Giovanni Falagario* [c], *Martin J. Connor* [b], *Taimur T. Shah* [b], *Joost J.C. Verhoeff* [a], *Pekka Taimen* [d], *Hannu J. Aronen* [e], *Juha Knaapila* [f], *Ileana Montoya Perez* [g], *Otto Ettala* [f], *Armando Stabile* [h], *Giorgio Gandaglia* [h], *Nicola Fossati* [h], *Alberto Martini* [h], *Vito Cucchiara* [h], *Alberto Briganti* [h], *Anna Lantz* [i], *Wolfgang Picker* [j], *Erik Skaaheim Haug* [k], *Tobias Nordström* [l], *Mariana Bertoncelli Tanaka* [b], *Deepika Reddy* [b], *Edward Bass* [b], *Peter S.N. van Rossum* [a], *Kathie Wong* [m], *Henry Tam* [b], *Mathias Winkler* [b], *Stephen Gordon* [m], *Hasan Qazi* [n], *Peter J. Boström* [f], *Ivan Jambor* [e], *Hashim U. Ahmed* [b]

[a] *Department of Radiotherapy, University Medical Center Utrecht, Utrecht, The Netherlands;* [b] *Department of Imperial Prostate, Imperial College London, London, UK;* [c] *Department of Urology and Organ Transplantation, University of Foggia, Foggia, Italy;* [d] *University of Turku and Department of Pathology, Turku University Hospital, Turku, Finland;* [e] *Department of Radiology, University of Turku, Turku, Finland;* [f] *Department of Urology, University of Turku and Turku University hospital, Turku, Finland;* [g] *Department of Computing, University of Turku, Turku, Finland;* [h] *Urological Research Institute, IRCCS Ospedale San Raffaele, Milan, Italy;* [i] *Department of Urology, Karolinska University Hospital, Solna, Sweden;* [j] *Department of Radiology, Aleris Cancer Center, Oslo, Norway;* [k] *Section of Urology, Vestfold Hospital Trust, Tønsberg, Norway;* [l] *Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden;* [m] *Department of Urology, Epsom and St. Helier's University Hospital Trust, Surrey, UK;* [n] *Department of Urology, St. George's Hospital NHS Foundation Trust, London, UK*

Article info

Article history:
Accepted July 26, 2022

Keywords:
Magnetic resonance imaging
Rapid Access to Prostate Imaging and Diagnosis pathway

Abstract

Background: Although multiparametric magnetic resonance imaging (MRI) has high sensitivity, its lower specificity leads to a high prevalence of false-positive lesions requiring biopsy.
Objective: To develop and externally validate a scoring system for MRI-detected Prostate Imaging Reporting and Data System (PIRADS)/Likert ≥3 lesions containing clinically significant prostate cancer (csPCa).
Design, setting, and participants: The multicentre Rapid Access to Prostate Imaging and Diagnosis (RAPID) pathway included 1189 patients referred to urology due to elevated age-specific prostate-specific antigen (PSA) and/or abnormal digital rectal examination (DRE); April 27, 2017 to October 25, 2019.

* Corresponding author. Room number Q01.1.110, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands. Tel. +31 (0) 887558800; Fax: +31 (0) 887555850.
E-mail address: m.peters-10@umcutrecht.nl (M. Peters).

https://doi.org/10.1016/j.eururo.2022.07.022

Imperial Rapid Access to Prostate Imaging and Diagnosis risk score
Clinically significant prostate cancer

Intervention: Visual-registration or image-fusion targeted and systematic transperineal biopsies for an MRI score of ≥4 or 3 + PSA density ≥0.12 ng/ml/ml.
Outcome measurements and statistical analysis: Fourteen variables were used in multivariable logistic regression for Gleason ≥3 + 4 (primary) and Gleason ≥4 + 3, and PROMIS definition 1 (any ≥4 + 3 or ≥6 mm any grade; secondary). Nomograms were created and a decision curve analysis (DCA) was performed. Models with varying complexity were externally validated in 2374 patients from six international cohorts.
Results and limitations: The five-item Imperial RAPID risk score used age, PSA density, prior negative biopsy, prostate volume, and highest MRI score (corrected c-index for Gleason ≥3 + 4 of 0.82 and 0.80–0.86 externally). Incorporating family history, DRE, and Black ethnicity within the eight-item Imperial RAPID risk score provided similar outcomes. The DCA showed similar superiority of all models, with net benefit differences increasing in higher threshold probabilities. At 20%, 30%, and 40% of predicted Gleason ≥3 + 4 prostate cancer, the RAPID risk score was able to reduce, respectively, 11%, 21%, and 31% of biopsies against 1.8%, 6.2%, and 14% of missed csPCa (or 9.6%, 17%, and 26% of foregone biopsies, respectively).
Conclusions: The Imperial RAPID risk score provides a standardised tool for the prediction of csPCa in patients with an MRI-detected PIRADS/Likert ≥3 lesion and can support the decision for prostate biopsy.
Patient summary: In this multinational study, we developed a scoring system incorporating clinical and magnetic resonance imaging characteristics to predict which patients have prostate cancer requiring treatment and which patients can safely forego an invasive prostate biopsy. This model was validated in several other countries.

1. Introduction

In patients referred with an elevated prostate-specific antigen (PSA), an area scoring of 3, 4, or 5 on magnetic resonance imaging (MRI) scoring systems is identifiable in the majority who undergo multiparametric MRI (mpMRI) [1]. The level of suspicion can range from equivocal to highly suspicious [2]. Current guidelines recommend MRI-targeted biopsy of these areas regardless of the suspicion score [3,4]. Performing prostate biopsies, without risk stratification, would result in approximately half of MRI lesions being diagnosed either with benign histology or as clinically insignificant prostate cancer (ciPCa) [5].

A risk-stratified MRI-directed pathway might allow a more advantageous balance of risks from biopsy-related complications and overdiagnosis of ciPCa against the risk of missing clinically significant prostate cancer (csPCa) [6].

Several risk models incorporating MRI findings have recently been developed [7–11]. As different MRI scoring systems such as Prostate Imaging Reporting and Data System version 2 (PIRADSv2) and Likert are used across various healthcare settings, applying a scoring system based on only one system is limited [2,3]. Previous models no longer reflect current practice, having been developed by incorporating biopsies in patients with an MRI score of 1 or 2 [1,5,12,13]. Finally, models often lack external validation [14].

A personalised risk score provided from a robustly developed and externally validated risk model might allow patients, with both an mpMRI with score of ≥3 and a low risk of csPCa, to defer immediate biopsy. We aimed to design and externally validate a contemporary and more representative risk prediction model for patients with mpMRI-detected lesion(s) with a score of ≥3 than what is present in the current literature.

2. Patients and methods

2.1. Study population

The development cohort used the Rapid Access to Prostate Imaging and Diagnosis (RAPID) registry (RAPID-Online), funded by NHS England, in order to demonstrate the deliverability of a high-quality MRI-first pathway following the results of PROMIS [12,15]. All consecutive patients referred on the MRI-directed pathway were included. The RAPID pathway involved straight-to-test mpMRI in patients referred by their primary care physician to a urology department due to elevated age-specific PSA (≥2.5 ng/ml at 45–49 yr, ≥3 ng/ml at 50–69 yr, or PSA ≥5 ng/ml at ≥70 yr) and/or abnormal digital rectal examination (DRE). Visual-registration or image-fusion targeted and systematic transperineal biopsies were performed for an MRI score of ≥4 or 3 with PSA density ≥0.12 ng/ml/ml. We have previously reported that approximately 43% of patients avoided immediate biopsy this way [16]. At analysis, three UK centres participated in RAPID-Online with a further eight centres currently. Initial analyses have been described [17,18]. For this analysis, patients who received a biopsy needed prebiopsy mpMRI with a visible MRI lesion assessed using the PIRADS (version 2.0) or Likert scoring system [18]. To avoid a clustering bias, the highest scoring lesion was selected for patients with multiple lesions [19]. MRI lesions with PIRADS or Likert ≥3 were included. We excluded patients with previous prostate cancer (PCa) or no MRI-directed targeted biopsy within ≤3 mo following MRI.

2.2. Candidate predictor variables

A multidisciplinary panel of urologists, radiologists, and researchers selected candidate predictors that were measurable, available, and reliable based on existing evidence. Four classes were selected: demographic (age, Black ethnicity, and family history of PCa), clinical (5-alpha reductase inhibitors, prior negative biopsy, and DRE), laboratory (PSA and PSA density), and radiological (prostate volume, PIRADSv2 score, Likert score, combined MRI score, number of MRI lesions, and index lesion size; Supplementary material).

2.3. Data collection and data quality

All RAPID sites prospectively collected data using a customised Research Electronic Data Capture (REDCap) tool hosted at Imperial College London. Training was conducted prior to the entry of each site into the registry, to ensure adherence to a standardised protocol. Standardised data collection forms were used with detailed definitions and instructions for each predictor variable. Quality assurance checks were performed throughout the collection period. Any issues with data quality led to a re-review of the primary health record by the coordinating centre (source data verification).

2.4. Procedures

2.4.1. MRI acquisition and reporting

The mpMRI scans were performed using a 1.5 T scanner with a pelvic phased-array coil in accordance with the European Society of Uroradiology guidelines [20]. The sequences included T1-weighted, T2-weighted, and dynamic contrast-enhanced images, and multiple *b* values (for apparent diffusion coefficient [ADC] maps and a high *b* value of 1500). MRI scores were reviewed by a second reader in a weekly multidisciplinary tumour board. Any discordance was resolved by consensus. Radiologists had over 3 yr of experience reporting >100 MRI scans per annum (Supplementary material) [21].

2.4.2. Biopsy protocol

Patients underwent a transperineal prostate biopsy according to a standard operating procedure at all participating sites [17]. Targeting was performed with visual-registration or image-fusion software using a biopsy system that employs elastic image coregistration (Biopsee; MedCom GmbH, Darmstadt) [22]. Each MRI lesion was targeted under real-time transrectal ultrasound guidance (using Hitachi Preirus device) and potted individually. The minimum was three biopsy cores per lesion. Additional nontargeted systematic sampling was performed [17]. Biopsies were evaluated in accordance with the International Society of Urological Pathology standards [23] by specialist uropathologists with ≥10 yr of experience. Pathologists were not blinded to other clinical characteristics.

2.5. Validation cohorts

The model was externally validated using six international cohorts from four different countries including 154, 351, 324, 385, 570, and 590 patients (total 2374). The prevalence of Gleason ≥3 + 4 ranged from 27% to 65%. These cohorts have been described previously, and represent data from different settings encompassing single or multicentre clinical trials and consecutive case series. The datasets incorporate a range of MRI scanners. All MRI scans in the external validation cohorts were contemporary and performed between 2013 and 2019 (for details, see the Supplementary material).

2.6. Outcomes

The primary endpoint was any-length Gleason ≥3 + 4. This is a conservative definition for csPCa, and its widespread availability allowed external validation. Alternative models were constructed using the secondary endpoints any amount of Gleason ≥4 + 3 and PROMIS definition 1 (≥4 + 3 or any grade ≥6 mm).

2.7. Statistical analysis

Multivariable variable selection was done with a backward stepwise approach minimising the Akaike's information criterion (AIC) [24]. Continuous variables were assessed for nonlinearity and transformed using the natural logarithm when appropriate.

Multicollinearity was defined as a correlation coefficient *r* of >0.5. Multiple imputation by chained equations was used for missing data, generating 20 imputed datasets. The results were pooled with Rubin's [25] rules.

We created three models, decreasing clinical complexity in two steps from the model with the lowest AIC value. First, DRE was removed (model 2), and second, Black ethnicity (Gleason ≥3 + 4) or lesion size (Gleason≥4 + 3) and family history of PCa were removed (model 1). We assessed whether this would reduce calibration, the C-statistic, and net benefit. We further compared the model with PSA density against a model with PSA and prostate volume separately and with all three variables (despite potential collinearity) included, to see which combination provided the most parsimonious model.

Internal validation was performed using 2000 bootstrap resamples from each imputed dataset, in which the modelling process was repeated. This provided correction of the final models' β coefficients and C-statistic. We assessed the models' performance by evaluating discrimination (Harrell's c-index, calibration plots) and plotting decision curves to assess the clinical utility of the model in predicting csPCa compared with the approaches of a biopsy-in-all or a biopsy-in-none strategy [26]. We further calculated the amount of missed csPCa against the number of biopsies prevented, given different model-predicted probabilities of csPCa. Nomograms and an online risk calculator were created to visually assess an individual patient's risk.

For external validation, the predicted probabilities in the six external cohorts on a per-patient basis were calculated using the complete models and compared with the observed probabilities in that population. The Delong test was used to compare C-statistics [27]. External validation was conducted in a masked manner with no data pooling across cohorts; the RAPID investigators had no access to the data from the cohorts outside the UK. To assess temporal variation, the database was divided into a model-derivation set (years 2017–2018), which was subsequently validated in data from 2019.

We additionally externally validated the model by Mehralivand et al [8] on our RAPID-Online data and the external cohorts.

Prediction modelling was adopted from the study of Steyerberg [28]. Data were analysed with R studio version 4.1.2 (R Foundation, Vienna, Austria) using packages rms, mice, psfmi, rmda, polycor, pROC, and ggplot2, among others, and STATA (version 15).

An extended description of the analysis is provided in the Supplementary material. The study follows the TRIPOD statement for reporting multivariable prediction model development and validation (checklist in the Supplementary material) [29].

2.8. Ethical considerations

Permission was granted by the institutional review board of each participating site either as consented research or as a continuous quality improvement project depending on local ethics committee requirements.

3. Results

The RAPID development cohort included 1189 patients with an MRI-detected PIRADS/Likert ≥3 lesion who had MRI targeted and systematic biopsies (April 27, 2017 to October 25, 2019). Gleason ≥3 + 4 was detected in 681 (57%) and Gleason ≥4 + 3 in 378 (32%; Fig. 1 and Table 1). Only 64 (2.5%) patients were excluded due to not undergoing biopsies within 3 mo.

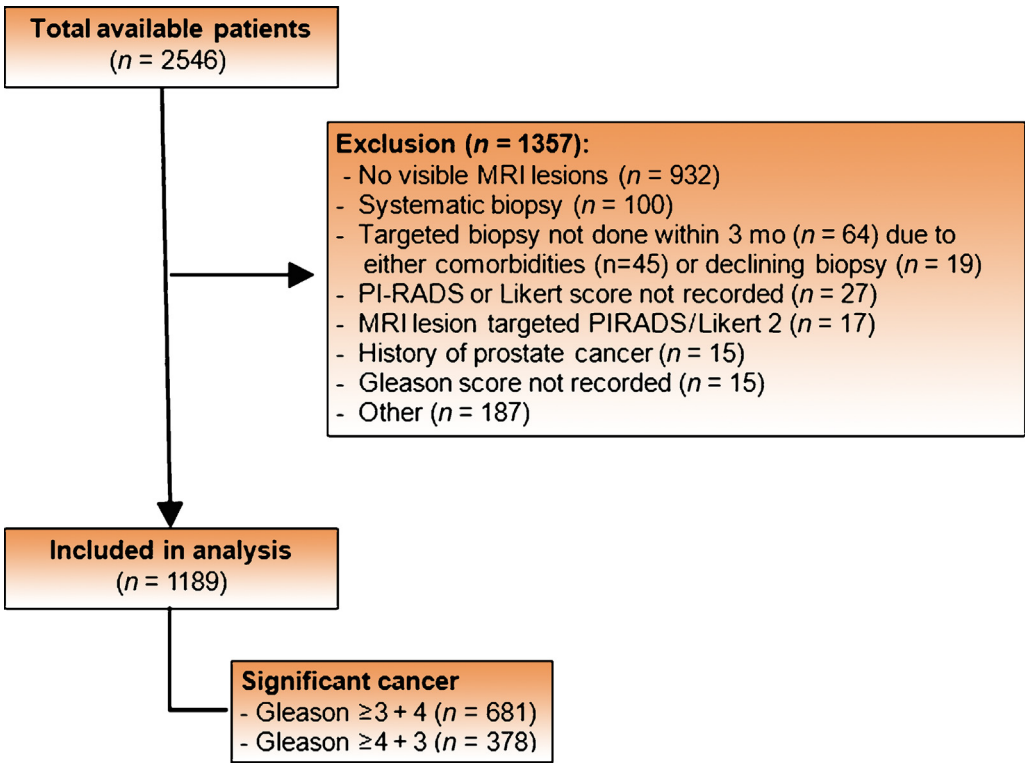

Fig. 1 – Flowchart describing the study population. MRI = magnetic resonance imaging; PIRADS = Prostate Imaging Reporting and Data System.

Table 1 – Summary characteristics

Covariate	Primary outcome		
	No PCa (*N* = 508)	Gleason ≥3 + 4 cancer (*N* = 681)	*p* value
Age (yr)	64.3 (7.6) [42.3–83.5]	67.8 (7.7) [42.1–86.7]	<0.001
PSA (ng/ml)	6.2 (4.7–8.9) [0.4–71]	8.4 (5.9–14.0) [1.5–1000]	<0.001
PSA density	0.13 (0.09–0.19) [0.009–1.56]	0.22 (0.15–0.36) [0.02–26.3]	<0.001
Black ethnicity	69 (14%)	76 (11%)	0.24
Family history of PCa	59 (12%)	89 (13%)	0.51
Prior prostate biopsy	42 (8.3%)	19 (2.8%)	<0.001
5-ARIs	15 (3.0%)	17 (2.5%)	0.76
Abnormal DRE	116 (28%)	300 (53.0%)	<0.001
Prostate volume	49 (35–67) [17–245]	38 (30–53) [12–150]	<0.001
Lesion diameter	12 (9.0–18.2) [2–50]	15 (10–23) [3–58]	<0.001
Number of lesions			0.01
1	373 (73%)	462 (68%)	
2	126 (25%)	187 (28%)	
3	8 (1.6%)	28 (4%)	
4	1 (0.2%)	4 (0.6%)	
Lesion score			<0.001
3	184 (36%)	67 (9.8%)	
4	234 (46%)	234 (34%)	
5	90 (18%)	380.0 (56%)	

5-ARI = 5-alpha reductase inhibitors; DRE = digital rectal examination; PCa = prostate cancer; PSA = prostate-specific antigen.
Summary statistics are presented as mean (±standard deviation) for continuous normally distributed data, median and interquartile range for skewed continuous variables, and *n* (%) for categorical data. Ranges are added in square brackets at the end. Differences were testes with a Student *t* test, Mann-Whitney U test, and chi-square test for these data types, respectively.

3.1. Primary outcome: Gleason ≥3 + 4

3.1.1. Model development and performance

Eleven variables were used. PSA was excluded due to collinearity with PSA density. After backward elimination, eight variables remained in the full model (age, Black ethnicity, PSA density, family history of PCa, prior negative biopsy, prostate volume, positive DRE, and highest MRI score), with a C-statistic of 0.832 (AIC 1201.9; model 3).

Table 2 – Multivariable model coefficients for Gleason ≥3 + 4

Variable	β (+SE), p value (model 1)	β (+SE), p value (model 2)	β (+SE), p value (model 3)
Intercept	−1.851187 (0.655685), 0.005	−1.812339 (0.676587), 0.007	−1.929103 (0.685040), 0.005
Log(PSA density)	1.103418 (0.132386), <0.0001	1.134648 (0.133952), <0.0001	1.100873 (0.134728), <0.0001
Prior negative biopsy	−1.020887 (0.334249), 0.002	−1.027477 (0.337287), 0.002	−0.890331 (0.337592), 0.008
MRI volume, by cc increase	−0.008079 (0.003220), 0.01	−0.007505 (0.003227), 0.02	−0.007888 (0.003265), 0.02
MRI score 4 (ref. = 3)	0.933048 (0.184440), <0.0001	0.958390 (0.185986), <0.0001	0.951417 (0.187133), <0.0001
MRI score 5 (ref. = 3)	1.886100 (0.201489), <0.0001	1.915564 (0.202535), <0.0001	1.819472 (0.204162), <0.0001
Age, per year increase	0.052568 (0.009823), <0.0001	0.052062 (0.010005), <0.0001	0.050507 (0.010116), <0.0001
Black ethnicity		−0.386964 (0.216905), 0.07	−0.381812 (0.216763), 0.08
Positive family history		0.394579 (0.214867), 0.07	0.394296 (0.216736), 0.07
DRE			0.570272 (0.152647), 0.0002

β = coefficient from model (the corresponding odds ratio can be calculated by e^β, where e = Euler's number [2.71828]); DRE = digital rectal examination; MRI = magnetic resonance imaging; PSA = prostate-specific antigen; ref. = reference; SE = standard error.
Coefficients in the table are the corrected coefficients after internal validation (correction factor 0.9544 and additional intercept correction of 0.00748) and are used in all subsequent steps of the analyses.

Removal of DRE resulted in AIC increase (1215.4) with similar C-statistic (0.826; model 2). The final model consisted of five independent predictors (age, PSA density, prior negative biopsy, prostate volume, and highest MRI score; model 1). The C-statistic was 0.823 (AIC 1218.4; Table 2). The models with PSA and prostate volume separately performed worse (AIC 1233 vs 1218.4 of model 1) as well as the model with all three variables included (AIC 1220.3 vs 1218.4).

3.1.2. *Decision curve analysis and trade-offs*

A decision curve analysis showed no difference between the three models regarding the net benefit over a range of 0–40% threshold probabilities compared with recommending biopsy for all (Fig. 2A). All three models were equally able to reduce biopsies for a range of predicted probabilities, while minimising missed csPCa (Fig. 3A and Supplementary Table 3). The difference in net benefit increased with higher threshold probabilities, from 0.04 at 20% to 0.07 at 40%. At 20%, 30%, and 40% of predicted Gleason ≥3 + 4 PCa, the RAPID risk score was able to reduce, respectively, 11%, 21%, and 31% of biopsies against 1.8%, 6.2%, and 14% of missed csPCa (or 9.6%, 17%, and 26% of foregone biopsies, respectively).

3.1.3. *External validation*

Temporal validation showed adequate calibration in the model developed based on the data from 2017 to 2018 applied to 2019, with minor deviations of the final model coefficients from the original analysis. External validation of the Imperial RAPID risk score models for Gleason ≥3 + 4 on the six external cohorts showed a constant C-statistic/area under the curve (AUC) ranging from 0.80 to 0.86 and stable calibration with a maximum of 10–15% miscalibration. The decision curve analysis predominantly showed an increased net benefit over the range of threshold probabilities compared with a biopsy-all strategy, especially in populations with a lower incidence of Gleason ≥3 + 4.

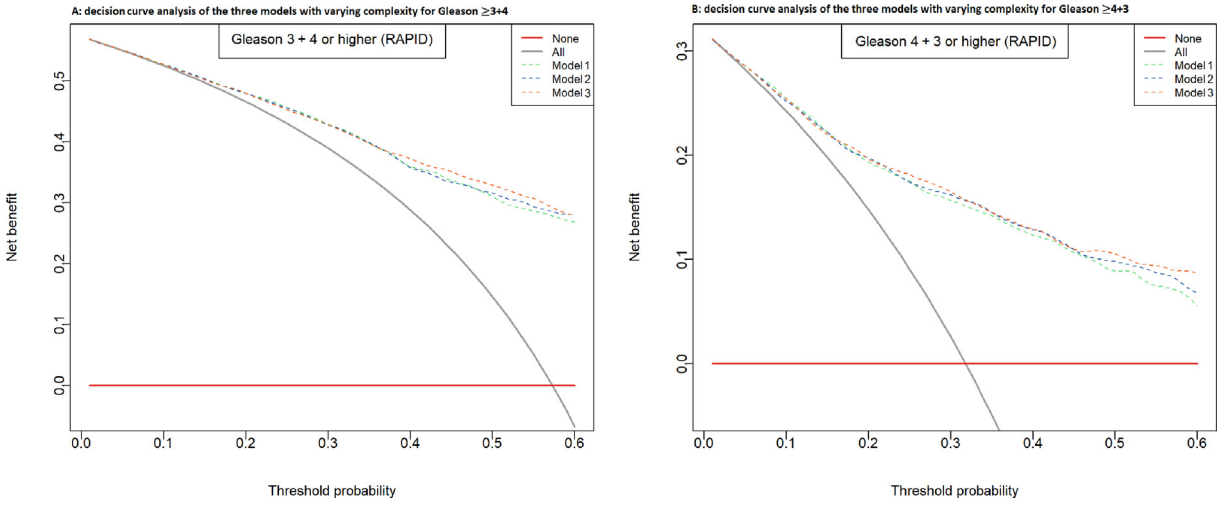

Fig. 2 – Decision curve analysis of the three models with varying complexity for the primary outcomes of (A) Gleason ≥3 + 4 and (B) Gleason ≥4 + 3 over a range of 0–60% threshold probabilities. The net benefit of the three models is depicted against a biopsy-all strategy (grey line) or biopsy-none strategy (black line). Model 1: log(PSA density), prior negative biopsy, MRI prostate volume, MRI score, and age; model 2: model 1 + black ethnicity (Gleason ≥3 + 4) or lesion size (Gleason ≥4 + 3) + positive family history; and model 3: model 2+ DRE. DRE = digital rectal examination; MRI = magnetic resonance imaging; PSA = prostate-specific antigen; RAPID = Rapid Access to Prostate Imaging and Diagnosis.

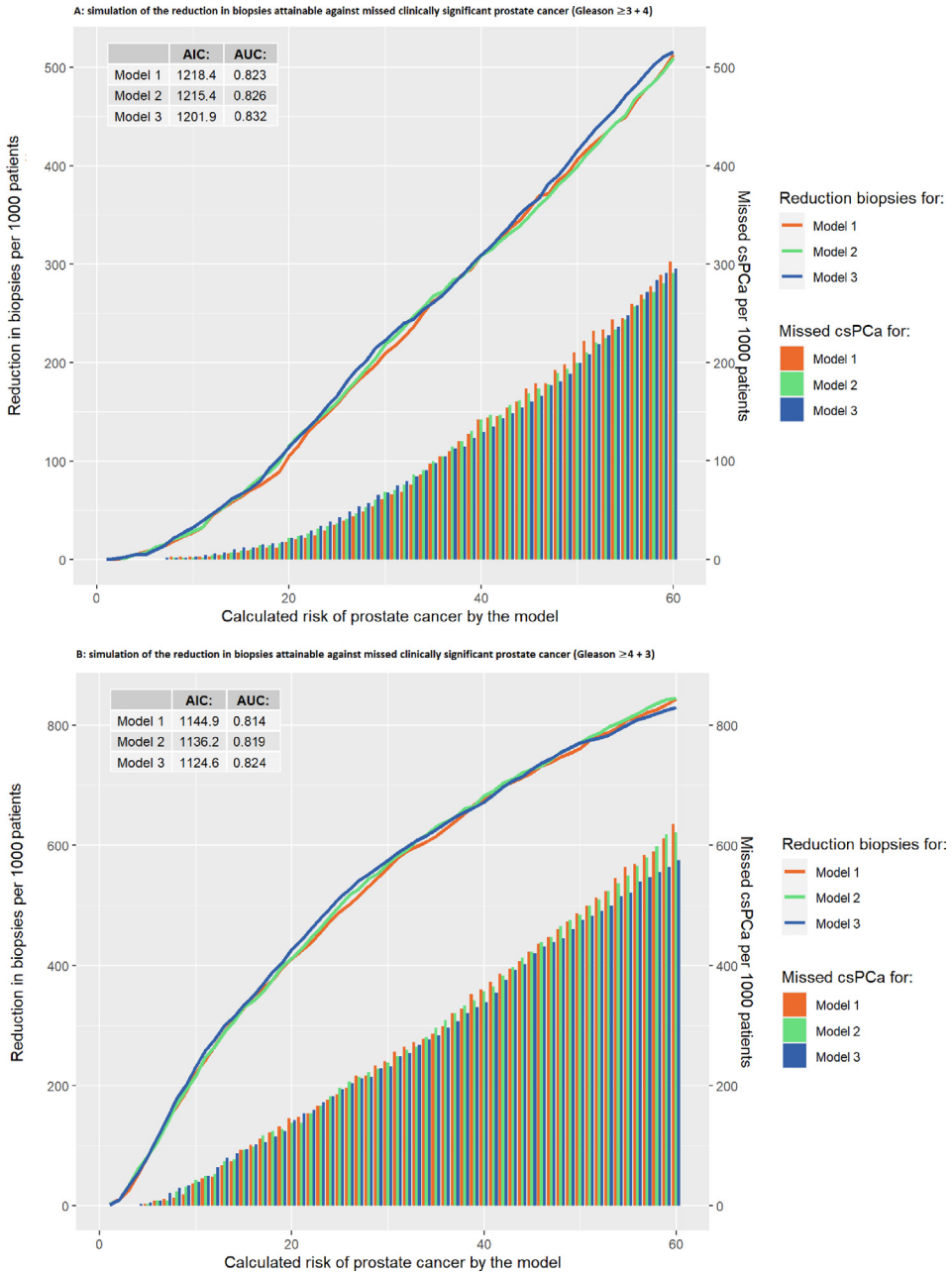

Fig. 3 – Simulation of the reduction in biopsies attainable (lines for the three separate models for the left *y* axis) against the amount of missed (A) Gleason ≥3 + 4 and (B) Gleason ≥4 + 3 clinically significant prostate cancer (csPCa; bars of the three models for the right *y* axis). The *x* axis represents the predicted probability of csPCa by the model. Example: when using model 1 on the RAPID database and filtering out the patients with ≤20% chance of Gleason ≥3 + 4 prostate cancer, you can reduce 10.5% of biopsies (or 105 per 1000 patients) while missing 1.76% (or 18 per 1000) Gleason ≥3 + 4 prostate cancers. Model 1: log (PSA density), prior negative biopsy, MRI prostate volume, MRI score, and age; model 2: model 1 + black ethnicity (Gleason ≥3 + 4) or lesion size (Gleason ≥4 + 3) + positive family history; and Model 3: model 2+ DRE. AIC = Akaike's information criterion; AUC = area under the curve; MRI = magnetic resonance imaging; PSA = prostate-specific antigen; RAPID = Rapid Access to Prostate Imaging and Diagnosis.

External validation results of the five cohorts originating outside the UK are depicted in Supplementary Figures 1 and 2, and Supplementary Table 1. External validation results of the UK cohort is depicted in Supplementary Figure 3 (De Long test $p = 0.61$).

External validation of the Mehralivand et al's [8] model on the RAPID data (Supplementary Fig. 4) showed a C-statistic of 0.812 (De Long test $p = 0.02$ vs model 1; De Long test model 1 vs models 2 and 3: $p = 0.12$ vs $p = 0.02$). Calibration of the model by Mehralivand et al [8] showed an overestimation of predicted probabilities in the RAPID database, which was corrected after recalibration of the intercept. Calibration on the external cohorts similarly showed an overestimation of the chance of csPCa (Gleason ≥3 + 4) of this model, which could be corrected after recalibrating the intercept (Supplementary Fig. 1).

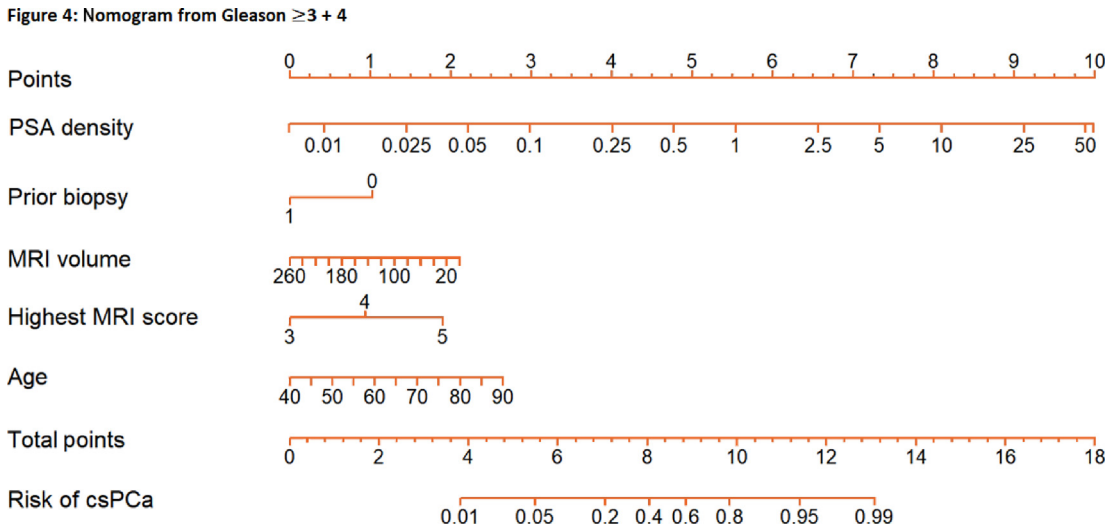

Fig. 4 – Nomogram from Gleason ≥3 + 4. The risk of clinically significant prostate cancer (csPCa; in this case Gleason ≥3 + 4) can be calculated by adding all points from the upper horizontal bar per variable after which the risk can be read from the lowest bar. For example, a patient with a PSA density of 0.1 ng/ml/ml (3 points), no previous biopsy (1 point), an MRI volume of 60 cc (1.5 points), an MRI score of 4 (1 point), and 70 yr of age (1.5 points) has a total of 8 points, translating into a risk of 40% of Gleason ≥3 + 4 PCa. MRI = magnetic resonance imaging; PCa = prostate cancer; PSA = prostate-specific antigen.

3.1.4. *Nomogram/risk calculator*

Figure 4 shows the nomogram resulting from the simplified model. An online risk calculator was created, available at https://nomogramsumcu.shinyapps.io/RAPID_Gleason_3_4/.

3.2. *Secondary outcomes: Gleason ≥4 + 3 and PROMIS definition 1*

3.2.1. *Gleason ≥4 + 3*

Results regarding Gleason ≥4 + 3 were largely concordant with Imperial RAPID risk score models for Gleason ≥3 + 4 (Supplementary Table 2), with satisfactory calibration and C-statistics of 0.814, 0.819, and 0.824 for models 1, 2, and 3, respectively. Temporal validation and external validation in the six external cohorts showed similar performance (Supplementary Fig. 5–7 and Supplementary Table 1). Larger net benefit differences were obtained at lower threshold probabilities than a biopsy-all strategy (Fig. 2B and 3B, and Supplementary Table 3).

3.2.2. *PROMIS definition 1*

Models for PROMIS definition 1 were similar to the models for Gleason ≥3 + 4 and ≥4 + 3 with C-statistics of 0.826, 0.830, and 0.831 for models 1, 2, and 3, respectively. Temporal validation showed good calibration and minor deviations from the original model. For all outcomes, see the Supplementary material for the statistical output, or included code and anonymised data (upon request).

4. Discussion

We report the development and external validation of a contemporary risk tool to determine which patients with suspicious MRI (score ≥3) might defer immediate biopsy, given the probability of csPCa. Routinely collected clinical factors were used, allowing for a broad clinical application. We observed that the simplified five-item Imperial RAPID risk score showed consistency after external validation utilising six datasets from multiple countries using different MRI scoring systems and diverse reference standards. We also examined the trade-off between avoided biopsy and missed detection of csPCa if the model was deployed as a decision tool in such patients. Such evaluation is vital in the process of gaining informed consent prior to recommending a prostate biopsy or not.

The prevalence of 57% of Gleason ≥3 + 4 is high in the RAPID dataset compared with the literature; consequently, the net benefit is increased most at higher threshold probabilities compared with other models [1,12]. At lower thresholds, for Gleason ≥3 + 4, the models' net benefit is more modest in the RAPID database. For example, an overview of prediction models incorporating MRI shows an average of 20–50% of men not biopsied when a threshold probability of 20% is adopted [14], against 11% using the RAPID risk score (125/1189). Missed csPCa accounts for 9.6% (12/125) of biopsies foregone, equivalent to the numbers found in the literature (∼10%, range ∼1.5–15%), but lower when looking relatively to all csPCa cases (1.8% of all Gleason ≥3 + 4 in RAPID, which compares favourably with the approximate 10% in the literature, ranging from 1% to 12%). Although the net benefit in RAPID is achieved at higher threshold probabilities, the models are applicable to populations with a varying prevalence of the endpoint as well, as visible in the external validation sets (where the incidence of Gleason ≥3 + 4 varied from 27% up to 65%). It is also clear that in these cohorts, the net benefit difference is more apparent at lower threshold probabilities. A comprehensive externally validated model from the literature [8] showed overprediction across a wide range of probabilities including the RAPID data, indicating the better applicability of the RAPID risk score across populations.

Additionally, for the endpoint with a lower incidence (Gleason ≥4 + 3), the models lead to a higher net benefit at lower threshold probabilities.

Our findings support the observations made by others that upfront risk stratification of patients with an MRI lesion could reduce harms from biopsy [6]. Previous studies have developed models including MRI findings that have reported similar performance characteristics but use either a single MRI scoring system and/or transrectal biopsies as a reference standard [7–11]. The Imperial RAPID risk score overcomes the limitations of previous models developed using a combination of PIRADSv2.0 or Likert scores, and validated across multiple independent multinational external datasets. The findings report high performance characteristics across different csPCa prevalences during external validation. To increase the reliability of MRI predictors, the protocol required MRI scans to be double reported. This was required in the RAPID pathway to address the widely reported issue of interobserver mpMRI variability [30].

We recognise that using MRI to develop risk scores is not new. Other groups have equally shown the possible clinical utility of a risk model incorporating MRI; an overview can be found in the study of Schoots and Roobol [14]. The European Randomized study of Screening for Prostate Cancer (ERSPC) risk calculator was added with MRI PIRADSv1 information (MRI-ERSPC-RC 3 and 4), increasing the AUC to 0.84 and saving 33–48% of biopsies against 7–10% missed Gleason $\geq$3 + 4 [10] at a risk threshold of 20%. Radtke et al [7] also improved the ERSPC risk calculator from a single-centre retrospective cohort determining better accuracy using PIRADSv1.0 MRI scores. The model was later validated externally [31]; however, predictive accuracy varied, resulting in an overestimation/miscalibration when a lower csPCa prevalence was observed. Miscalibration is also frequently observed when the MRI-ERSPC 3–4 risk calculators (and other models) are validated externally [32,33], as was additionally seen when the model by Mehralivand et al [8] was applied to our data. Uniquely, our study reports high performance characteristics for models developed using a combination of PIRADSv2.0 and Likert scores, and validated across multiple independent multinational external datasets. The Imperial RAPID risk score appeared robust across different csPCa prevalences during external validation.

Other models are developed using (transrectal ultrasonography–guided systematic) biopsies for patients with nonsuspicious MRI, which are subject to spectrum and verification bias when comparing results with the RAPID risk score. We do not know in which direction model performance would change when a different reference test is used. In case of a difference between these biopsy techniques, as reported in the literature, transrectal biopsies often perform worse in detecting clinically significant disease than MRI targeted and systematic transperineal biopsies [34]. However, the effect on diagnostic accuracy can vary in such cases [35]. The issue of a spectrum bias can result in decreased applicability of the model in populations with less severe disease characteristics. The effect on diagnostic accuracy once again varies between studies [35].

The Imperial RAPID risk score reflects clinical practice in many centres and countries, where patients with MRI scores of 1 and 2 alongside those with an MRI score of 3 and a low PSA density are not biopsied immediately [13].

As a result, we expect any risk calculator such as ours, in this setting, to have an incremental impact. Indeed, given the higher prevalence of significant cancers in patients with suspicious MRI, the rates of biopsy avoidance that can be achieved are in the order of 10%; such an incremental impact contributes to the reduction in biopsy-related harms.

There are limitations to our study. RAPID-Online is an observation cohort and not a clinical trial. Nonetheless, it has a standardised protocol for delivering the pathway across multiple sites and reports outcomes of its large dataset in a uniform manner with all consecutive cases collected. Although the model was validated externally in cohorts from four different countries with varying prevalence of the target cancer endpoints, the results may not be generalisable to all populations. The model is not valid for the transrectal route of biopsy. We did not incorporate other descriptors of MRI lesions (eg, size, diffuse vs focal, capsular abutment, and ADC values). Rather than relying on individual radiological criteria that may not be representative of the full diagnostic potential of the imaging technique, we utilised an MRI score (PIRADS or Likert) that combines multiple imaging findings.

There have been many previous validated nomograms that have been published and made available for physicians and the public. The impact that these models have on physicians' and patients' decision-making process, or the overall impact these have on a healthcare service is unclear. To this end, we intend to prospectively evaluate the clinical utility of the Imperial RAPID risk score regarding the decision on biopsy as an embedded randomisation within the IP3-PROSPECT study (ClinicalTrials.gov NCT04400656), which uses the cohort multiple randomised controlled study design [36]. Furthermore, we will evaluate model performance and potentially adapt the model based on newly acquired data from RAPID in the future. We encourage further validation on external cohorts.

5. Conclusions

The Imperial RAPID risk score provides a standardised tool for the prediction of different definitions of csPCa in patients with an MRI-detected PIRADS/Likert $\geq$3 lesion, and can support patients and physicians to make decisions regarding the need for prostate biopsy across a range of probabilities.

Author contributions: Max Peters had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Peters, Eldred-Evans, Winkler, Ahmed, Shah, Connor, Reddy, Bass.

Acquisition of data: Peters, Eldred-Evans, Winkler, Ahmed, Shah, Connor, Reddy, Bass.

Analysis and interpretation of data: Peters, Kurver, van Rossum, Eldred-Evans, Falagario.

Drafting of the manuscript: Peters, Eldred-Evans, Ahmed, Shah.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Peters, Kurver, van Rossum, Eldred-Evans, Falagario.

Obtaining funding: Ahmed, Winkler, Shah, Eldred-Evans.

Administrative, technical, or material support: Ahmed, Winkler, Connor, Reddy, Bass, Gordon, Qazi, Falagario, Taimen, Aronen, Knaapila, Perez, Ettala, Stabile, Gandaglia, Fossati, Martini, Cucchiara, Briganti, Lantz, Picker, Haug, Nordström, Tanaka, Wong, Tam, Boström, Jambor.
Supervision: Ahmed, Winkler, Shah.
Other: None.

Financial disclosures: Max Peters certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: Martin J. Connor reports grant funding from University College Hospitals London (UCLH) Charity for his prostate cancer research. Pekka Taimen reports research grant paid to institution by Finnish Cancer Foundation; consulting fees from Faron Pharmaceuticals; honoraria for lecture from Roche Finland; being a data management committee member at ProScreen Prostate Cancer Screening Trial. Hannu J. Aronen reports novel MRI techniques for noninvasive detection and characterisation of prostate cancer research grant from the Turku University Central Hospital. Otto Ettala reports academic grants from Finnish Cancer Society and Sakari Alhopuro Foundation. Alberto Martini reports grants from Intuitive Surgical (Sunnyvale, CA, USA). Alberto Briganti reports academic grant from Italian Ministry of Health. Tobias Nordström reports consulting fees from AstraZeneca; payment/honoraria from Ipsen; stock or stock options at A3P Biomedical. Deepika Reddy reports funding for research from Prostate Cancer UK; funding for attendance to conferences from Sonblate Corp and Imperial Healthcare Charity. Peter J. Boström reports grant from Cancer Society of Finland; lecture fees from Astellas; being a member of the advisory board of Pfizer. Hashim U. Ahmed reports support from Wellcome Trust; grants from Cancer Research UK, National Institute of Health Research, Medical Research Council, the Urology Foundation, Prostate Cancer UK, the BMA Foundation, Sophiris Biocorp, and Sonablate Corp; consulting fees for teaching/training of HIFU prostate procedure from Sonablate Corp and teaching/training of cryotherapy and Rezum prostate procedure from Boston Scientific; payment/honoraria for teaching/training courses for HIFU from Sonablate Corp and teaching/training courses for Rezum from Boston Scientific; support for attending meetings and/or travels from Sonablate Corp; participation for trial design in Francis Medical; chair at NCRI Prostate Research Group (unpaid); specialist adviser for NICE Diagnostics Assessment Panel on Transperineal Freehand Biopsy (unpaid).

***Funding/Support and role of the sponsor*:** This work was supported by Wellcome Trust Senior Research Fellowship (ref: 204998/Z/16/Z) and NHS England Cancer Transformation Fund. The funders had no role in any of the aspects of the conduction of the current study.

Peer Review Summary and supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eururo.2022.07.022.

References

[1] Drost FJH, Osses DF, Nieboer D, et al. Prostate MRI, with or without MRI-targeted biopsy, and systematic biopsy for detecting prostate cancer. Cochrane Database Syst Rev 2019;4:CD012663.
[2] Barentsz JO, Weinreb JC, Verma S, et al. Synopsis of the PI-RADS v2 guidelines for multiparametric prostate magnetic resonance imaging and recommendations for use. Eur Urol 2016;69:41–9.
[3] NICE. NICE guidance—prostate cancer: diagnosis and management. BJU Int 2019;124:9–26.
[4] Mottet N, Bellmunt J, Bolla M, et al. EAU-ESTRO-SIOG guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent. Eur Urol 2017;71:618–29.
[5] Kasivisvanathan V, Rannikko AS, Borghi M, et al. MRI-targeted or standard biopsy for prostate-cancer diagnosis. New Engl J Med 2018;378:1767–77.
[6] Schoots IG, Padhani AR, Rouvière O, Barentsz JO, Richenberg J. Analysis of magnetic resonance imaging–directed biopsy strategies for changing the paradigm of prostate cancer diagnosis. Eur Urol Oncol 2020;3:32–41.
[7] Radtke JP, Wiesenfarth M, Kesch C, et al. Combined clinical parameters and multiparametric magnetic resonance imaging for advanced risk modeling of prostate cancer—patient-tailored risk stratification can reduce unnecessary biopsies. Eur Urol 2017;72:888–96.
[8] Mehralivand S, Shih JH, Rais-Bahrami S, et al. A magnetic resonance imaging–based prediction model for prostate biopsy risk stratification. JAMA Oncol 2018;4:678.
[9] Perez IM, Jambor I, Kauko T, et al. Qualitative and quantitative reporting of a unique biparametric MRI: towards biparametric MRI-based nomograms for prediction of prostate biopsy outcome in men with a clinical suspicion of prostate cancer (IMPROD and MULTI-IMPROD trials). J Magn Reson Imaging 2020;51:1556–67.
[10] Alberts AR, Roobol MJ, Verbeek JFM, et al. 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 2019;75:310–8.
[11] Distler FA, Radtke JP, Bonekamp D, et al. The value of PSA density in combination with PI-RADS™ for the accuracy of prostate cancer prediction. J Urol 2017;198:575–82.
[12] Ahmed HU, Bosaily AE, Brown LC, et al. The PROMIS study: a paired-cohort, blinded confirmatory study evaluating the accuracy of multi parametric MRI and TRUS biopsy in men with an elevated PSA. J Clin Oncol 2016;34:5000.
[13] NICE. Prostate cancer: diagnosis and management; 2021. https://www.nice.org.uk/guidance/ng131.
[14] Schoots IG, Roobol MJ. Multivariate risk prediction tools including MRI for individualized biopsy decision in prostate cancer diagnosis: current status and future directions. World J Urol 2020;38:517–29.
[15] NHS England. Implementing a timed prostate cancer diagnostic pathway; 2018. https://www.england.nhs.uk/wp-content/uploads/2018/04/implementing-timed-prostate-cancer-diagnostic-pathway.pdf.
[16] Bass E, Eldred-Evans D, Connor M, et al. MP13-06 The Rapid Access Prostate Imaging and Diagnosis (RAPID) Pathway—a multicentre update of 1719 patients undergoing multi-parametric MRI as a triage test. J Urol 2020;203:e186.
[17] Connor M, Eldred-Evans D, van Son M, et al. A multicentre study of the clinical utility of non-targeted systematic transperineal prostate biopsies in patients undergoing pre-biopsy mpMRI. J Urol 2020;204:1195–201.
[18] Khoo CC, Eldred-Evans D, Peters M, et al. Likert vs PI-RADS v2: a comparison of two radiological scoring systems for detection of clinically significant prostate cancer. BJU Int 2020;125:49–55.
[19] Gönen M, Panageas KS, Larson SM. Statistical issues in analysis of diagnostic imaging experiments with multiple observations per patient. Radiology 2001;221:763–7.
[20] Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS prostate imaging–reporting and data system: 2015, version 2. Eur Urol 2016;69:16–40.
[21] Brizmohun Appayya M, Adshead J, Ahmed HU, et al. National implementation of multi-parametric magnetic resonance imaging for prostate cancer detection—recommendations from a UK consensus meeting. BJU Int 2018;122:13–25.
[22] Khoo CC, Eldred-Evans D, Peters M, et al. A comparison of prostate cancer detection between visual estimation (cognitive registration) and image fusion (software registration) targeted transperineal prostate biopsy. J Urol 2021;205:1075–81.
[23] Epstein JI, Allsbrook Jr WC, Amin MB, Egevad LL, Committee IG. The 2005 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma. Am J Surg Pathol 2005;29:1228–42.
[24] Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. BMJ 2009;338:b604.

[25] Rubin DB. Multiple imputation for nonresponse in surveys. John Wiley & Sons; 2004.
[26] Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and interpreting decision curve analysis: a guide for investigators. Eur Urol 2018;74:796–804.
[27] Delong ER, Delong DM, Clarkepearson DI. Comparing the areas under 2 or more correlated receiver operating characteristic curves—a nonparametric approach. Biometrics 1988;44:837–45.
[28] Steyerberg EW. Clinical prediction models. A practical approach to development, validation, and updating. Springer; 2009.
[29] Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162:W1–W73.
[30] Sonn GA, Fan RE, Ghanouni P, et al. Prostate magnetic resonance imaging interpretation varies substantially across radiologists. Eur Urol Focus 2019;5:592–9.
[31] Pullen L, Radtke JP, Wiesenfarth M, et al. External validation of novel magnetic resonance imaging-based models for prostate cancer prediction. BJU Int 2020;125:407–16.
[32] Lee HJ, Lee A, Yang XY, et al. External validation and comparison of magnetic resonance imaging-based predictive models for clinically significant prostate cancer. Urol Oncol 2021;39:783.e1–783.e10.
[33] Deniffel D, Healy GM, Dong X, et al. Avoiding unnecessary biopsy: MRI-based risk models versus a PI-RADS and PSA density strategy for clinically significant prostate cancer. Radiology 2021;300:369–79.
[34] Hossack T, Patel MI, Huo A, et al. Location and pathological characteristics of cancers in radical prostatectomy specimens identified by transperineal biopsy compared to transrectal biopsy. J Urol 2012;188:781–5.
[35] Whiting PF, Rutjes AWS, Westwood ME, Mallett S. QUADAS-2 Steering Group. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. J Clin Epidemiol 2013;66:1093–104.
[36] Bass EJ, Klimowska-Nassar N, Sasikaran T, et al. PROState pathway embedded comparative trial: the IP3-PROSPECT study. Contemp Clin Trials 2021;107:106485.