Asian Journal of Andrology (2019) 21, 592–597
www.asiaandro.com; www.ajandrology.com

Open Access

ORIGINAL ARTICLE

Prostate Cancer

Phi-based risk calculators performed better in the prediction of prostate cancer in the Chinese population

Yi-Shuo Wu[1,2,3,*], Xiao-Jian Fu[4,*], Rong Na[5], Ding-Wei Ye[6], Jun Qi[7], Xiao-Ling Lin[2], Fang Liu[2], Jian Gong[1,2], Ning Zhang[1,2], Guang-Liang Jiang[1,2], Hao-Wen Jiang[1,2], Qiang Ding[1,2], Jianfeng Xu[2,3], Ying-Hao Sun[8]

Risk prediction models including the Prostate Health Index (phi) for prostate cancer have been well established and evaluated in the Western population. The aim of this study is to build phi-based risk calculators in a prostate biopsy population and evaluate their performance in predicting prostate cancer (PCa) and high-grade PCa (Gleason score ≥7) in the Chinese population. We developed risk calculators based on 635 men who underwent initial prostate biopsy. Then, we validated the performance of prostate-specific antigen (PSA), phi, and the risk calculators in an additional observational cohort of 1045 men. We observed that the phi-based risk calculators (risk calculators 2 and 4) outperformed the PSA-based risk calculator for predicting PCa and high-grade PCa in the training cohort. In the validation study, the area under the receiver operating characteristic curve (AUC) for risk calculators 2 and 4 reached 0.91 and 0.92, respectively, for predicting PCa and high-grade PCa, respectively; the AUC values were better than those for risk calculator 1 (PSA-based model with an AUC of 0.81 and 0.82, respectively) (all $P < 0.001$). Such superiority was also observed in the stratified population with PSA ranging from 2.0 ng ml^{-1} to 10.0 ng ml^{-1}. Decision curves confirmed that a considerable proportion of unnecessary biopsies could be avoided while applying phi-based risk calculators. In this study, we showed that, compared to risk calculators without phi, phi-based risk calculators exhibited superior discrimination and calibration for PCa in the Chinese biopsy population. Applying these risk calculators also considerably reduced the number of unnecessary biopsies for PCa.

Asian Journal of Andrology (2019) **21**, 592–597; doi: 10.4103/aja.aja_125_18; published online: 22 March 2019

Keywords: Chinese; p2PSA; prostate biopsy; Prostate Health Index; risk calculator

INTRODUCTION

Prostate cancer (PCa) is the second most common cancer and one of the leading causes of cancer-related death among males worldwide.[1] The incidence of PCa in China is relatively low compared to that in Western countries; however, it has been progressively rising in recent decades and has become the seventh most common cancer among Chinese males.[2,3]

Prostate-specific antigen (PSA) has been introduced to prostate cancer screening and clinical practice for decades and is still the most widely used biomarker for screening and early detection of PCa. However, PSA is organ specific rather than cancer specific; therefore, PSA screening has resulted in large numbers of unnecessary biopsies and overdiagnosis of indolent cancers.[4–6]

Existing studies in Caucasians with PSA levels ranging from 2.0 ng ml^{-1} to 10.0 ng ml^{-1} have shown that [-2]proPSA (p2PSA), an isoform of PSA, and its derivative Prostate Health Index (phi) might have a better discriminating ability in predicting PCa and high-grade PCa than PSA.[7–9] Our previous study reported consistent results, and we found that phi performed better than PSA in discriminating PCa and non-PCa in patients with higher PSA levels in the Chinese population.[10]

Various risk calculators have been developed based on different clinical variables and have provided added value to the PSA test for predicting PCa or high-grade PCa. Among them, the Prostate Cancer Prevention Trial (PCPT) risk calculator[11] and the European Randomized Study of Screening for Prostate Cancer (ERSPC) risk calculator[12] have been most widely used in Caucasians. It has been proven that these risk calculators can avoid 2%–20% unnecessary biopsies at a different threshold probability of PCa.[13] The variables in the PCPT risk calculators include age, race, PSA, digital rectal examination (DRE), family history, and history of a previous negative prostate biopsy. The variables in the ERSPC risk calculators include PSA, DRE, the results of transrectal ultrasound (TRUS), prostate volume (PV), history of prostate biopsy, and the new version that incorporates

[1]Department of Urology, Huashan Hospital, Fudan University, Shanghai 200040, China; [2]Urology Research Center, Fudan University, Shanghai 200040, China; [3]Program for Personalized Cancer Care, NorthShore University HealthSystem, Evanston, IL 60201, USA; [4]Department of Surgery, Huashan Hospital, Fudan University, Shanghai 200040, China; [5]Department of Urology, Ruijin Hospital, Shanghai Jiaotong University School of Medicine, Shanghai 200025, China; [6]Department of Urology, Shanghai Cancer Center, Fudan University, Shanghai 200032, China; [7]Department of Urology, Xinhua Hospital, Shanghai Jiaotong University School of Medicine, Shanghai 200092, China; [8]Department of Urology, Changhai Hospital, The Second Military Medical University, Shanghai 200433, China.
*These authors contributed equally to this work.
Correspondence: Dr. Q Ding (qiangd_urology@126.com) or Dr. J Xu (jxu@northshore.org)
Received: 21 July 2018; Accepted: 10 December 2018

phi.[14–16] The utility of these risk calculators for predicting PCa has been externally validated in the Western population.[17,18] However, these two risk calculators are suitable for patients with PSA from 0 (or 0.5 ng ml^{-1}) to 50 ng ml^{-1}, and phi is only applicable when a patient had a recent DRE and had his PV measured using ERSPC risk calculator 3.

Most risk calculators (included PCPT and ERSPC risk calculators) were based on screening populations with low average PSA levels (most participants in PCPT had PSA below 6.0 ng ml^{-1}, and the participants in ERSPC had a mean PSA of 1.7 ng ml^{-1}).[11,13,19] However, PSA screening is not often utilized in China, and most of the Chinese patients come to urologists for elevated PSA, lower urinary tract symptoms, and abnormal findings in DRE or TRUS in their comprehensive physical examination. Such population has a higher risk for PCa than the screening population. Therefore, these risk calculators might not be appropriate for such different clinical settings in China. In addition, according to a previous study, PCPT and ERSPC risk calculators overestimate the probability of PCa and high-grade disease (Gleason score ≥7).[20] Furthermore, we considered that phi might have more widespread use in Chinese biopsy populations according to our former study (*e.g.*, for patients with PSA ≥10.0 ng ml^{-1}).[10] In consideration of the limited data on the performance of PCa risk calculators that include phi in the Chinese population, here, we built risk calculators (with or without phi) based on a prostate biopsy population and followed with an observational study to validate the performance of our risk calculators.

PATIENTS AND METHODS

Cohort of the training study

Subjects (n = 635) who underwent initial prostate biopsy from April 2012 to August 2013 in three tertiary medical centers in Shanghai (including Huashan Hospital, Fudan University; Shanghai Cancer Center, Fudan University; and Xinhua Hospital, Shanghai Jiao Tong University School of Medicine) were recruited as the training cohort. All clinical information was collected before biopsy. The patients were excluded if any essential clinical data on age, PSA, %fPSA (free PSA divided by PSA), p2PSA, or PV were missing. The characteristics of tertiary health institutes in China had been described in our previous study.[10]

Cohort of the validation study

Patients (n = 1045) who underwent initial prostate biopsy from August 2013 to December 2014 in four tertiary medical centers in Shanghai (including Huashan Hospital, Fudan University; Shanghai Cancer Center, Fudan University; Xinhua Hospital, Shanghai Jiao Tong University School of Medicine; and Changhai Hospital, the Second Military Medical University) were enrolled in our observational validation study. Participants who completed all the examinations (PSA, fPSA, p2PSA, and TRUS) were included in the final analysis. All clinical data were entered into the risk calculator to generate everyone's risk index and were recorded. This process did not influence the decision-making for prostate biopsy in the validation population. In both the training and validation studies, all subjects underwent an ultrasound-guided transperineal needle prostate biopsy with 10–14 cores. The indications for prostate biopsy at our institute were the following: (1) PSA >10.0 ng ml^{-1}; (2) PSA >4.0 ng ml^{-1} with confirmation after 2–3 months; (3) PSA level ranging from 4.0 ng ml^{-1} to 10.0 ng ml^{-1}, with suspicious fPSA/PSA <0.16 or PSA density >0.15 (PSAD = PSA/PV, PV (ml) = height (cm) × length (cm) × width (cm) × 0.52); and (4) positive findings from a DRE, TRUS, or magnetic MRI with any level of PSA.

Sample collection

All biopsy specimens were diagnosed by pathologists from the pathology department of each hospital. All blood samples were collected before biopsy for the measurement of PSA, fPSA, and p2PSA using the Beckman Coulter D×I 800 Immunoassay System (Beckman Coulter, Brea, CA, USA). The protocol of the current study was reviewed and approved by the institutional review board of each hospital in Shanghai, China. Both written informed consent and verbal informed consent were obtained from the patients for their participation in the study.

Statistical analyses

Phi was calculated using the following formula: (p2PSA/fPSA) × √PSA. PSA and phi were log-transformed for further statistical analysis due to their nonnormalized distributions. Based on the training cohort (n = 635), we built four risk calculators. Risk calculator 1 (RC1) was built according to the rules of a logistic regression model based on age, lgPSA (logarithm of PSA), and %fPSA. Risk calculator 2 (RC2) was built according to the rules of a logistic regression model based on age and lgphi (logarithm of phi). Risk calculator 3 (RC3) was built according to the rules of a logistic regression model based on age, lgPSA, %fPSA, and PV. Risk calculator 4 (RC4) was built according to the rules of a logistic regression model based on age, lgphi, and PV. Then, we evaluated the performances of phi, RC1, RC2, RC3, and RC4 in predicting PCa and high-grade PCa (Gleason score ≥7) in the training cohort, the validation cohort, and their subgroups (i.e., the population with PSA ranging from 2.0 ng ml^{-1} to 10.0 ng ml^{-1}).

The baseline characteristics (age, PV, PSA, and %fPSA) between two cohorts were compared using Student's t-test. The performance of each risk calculator was evaluated statistically on the basis of its discrimination and calibration according to established framework.[21] Receiver operating characteristic (ROC) curves were used to measure the discriminating abilities of different predictors (phi and the four risk calculators). A Z-test was performed to evaluate the differences in area under the ROC curves (AUCs) for phi and the four risk calculators (RC1 was used as a reference). We performed a decision curve analysis (DCA) to evaluate the potential clinical usefulness of making decisions based on phi and the four risk calculators. A two-sided test with P = 0.05 was used. Statistical analyses were performed using SPSS 19.0 (Statistical Product and Service Solutions, IBM Corporation, Armonk, NY, USA) and R version 3.4.0 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

A total of 635 patients were included in the training study, and 1045 patients were enrolled in the validation study. The positive biopsy rates were 42.8% and 43.0% in the training and validation cohorts, respectively, which were comparable.

Training study

The characteristics of the cohort and the stratified subgroups (PSA ranging from 2.0 ng ml^{-1} to 10.0 ng ml^{-1}) are shown in **Table 1**. In the training cohort and its subgroup with PSA ranging from 2.0 ng ml^{-1} to 10.0 ng ml^{-1}, the mean age, PSA, and phi were statistically higher in men diagnosed with PCa than in men without PCa, whereas the mean PV and %fPSA were lower in the PCa group (all P < 0.05). When the patients were categorized by a Gleason score ≥7 for high-grade PCa and others, the differences in age, PSA, phi, and PV remained significant, whereas the differences in %fPSA were insignificant (data not shown).

Multivariate logistic regression analyses were performed to evaluate each factor in the training cohort. We observed that age, PSA, %fPSA, p2PSA, and PV were all associated with PCa (all P < 0.01), while %fPSA

was not associated with high-grade PCa. On the basis of the results from the multivariate logistic regression analyses and the consensus risk factors (*i.e.*, age, PSA, %fPSA, p2PSA, and PV), we constructed the risk calculators as follows:

For PCa:

i. Risk calculator 1 (RC1): risk points = −4.928 + 0.023 × age + 2.789 × lgPSA − 1.962 × %fPSA
ii. Risk calculator 2 (RC2): risk points = −9.597 + 0.026 × age + 4.154 × lgphi
iii. Risk calculator 3 (RC3): risk points = −4.566 + 0.035 × age − 0.038 × PV + 3.236 × lgPSA − 1.735 × %fPSA;
iv. Risk calculator II (RC4): risk points = −8.919 + 0.036 × age − 0.029 × PV + 4.147 × lgphi.

For high-grade PCa:

i. Risk calculator 1 (RC1): risk points = −5.547 + 0.012 × age + 3.058 × lgPSA
ii. Risk calculator 2 (RC2): risk points = −8.791 + 0.016 × age + 3.664 × lgphi
iii. Risk calculator 3 (RC3): risk points = −5.193 + 0.021 × age + 3.351 × lgPSA − 0.029 × PV
iv. Risk calculator II (RC4): risk points = −8.315 + 0.021 × age + 3.642 × lgphi − 0.017 × PV.

In further analysis, the prediction accuracy of PSA, phi, and risk calculators 1–4 were evaluated in the training cohort. The AUCs of the different PCa risk calculators in the training cohorts and its subgroups are shown in **Table 2**. In the training cohort, when predicting PCa, the AUCs for phi, RC2, and RC4 were 0.88, 0.89, and 0.90, respectively, which indicated that phi and phi-based risk calculators (RC2 and RC4) all performed better than RC1 (AUC = 0.83, all $P < 0.05$). When predicting high-grade PCa (a Gleason sore ≥7), there was no significant difference among the AUCs of RC1 (AUC = 0.84), phi (AUC = 0.89), and RC2 (AUC = 0.89) in the training cohort, while RC4 (AUC = 0.89) performed better than RC1 ($P < 0.05$). However, phi, RC2, and RC4 did not outperform RC1 in the subgroup with PSA ranging from 2.0 ng ml^{-1} to 10.0 ng ml^{-1}, and similar results were also observed when predicting high-grade PCa in the same subgroup (**Table 3**).

Validation study

In the subsequent validation study, we evaluated the predictive performance of PSA, phi, and RC1-4 in the validation cohort. The characteristics of the validation cohort and its subgroup (PSA ranging from 2.0 ng ml^{-1} to 10.0 ng ml^{-1}) are shown in **Table 4,** and they were consistent with those of the training group.

When evaluating the prediction abilities of the risk calculators in the validation cohort and its subgroup, phi (**Tables 2 and 3**), RC2, and RC4 all outperformed RC1 (0.91, 0.91, and 0.91, respectively, *vs* 0.81 in the whole validation cohort and 0.89, 0.90, and 0.89, respectively, *vs* 0.71 in its subgroup) in predicting PCa (all $P < 0.01$). When predicting high-grade PCa, such superiority was still observed in the RCs that included phi. The calibration of the risk calculators was also assessed in the validation cohort (**Table 5**).

The ROC curves for phi and RC 1–4 in predicting PCa in two cohorts and in the subgroups with PSA ranging from 2.0 ng ml^{-1} to 10.0 ng ml^{-1} are shown in **Supplementary Figure 1**. The DCA revealed that any phi-based risk calculator performed better than PSA-based RC1 in both training and validation cohorts (**Supplementary Figures 2** and **3**). A potential net reduction in the number of biopsies was seen at PCa risk thresholds of approximately 20% for PCa in both cohorts.

Table 1: Clinical characteristics of the training cohort

Variables	*All PSA*				*PSA: 2.0–10.0 ng ml^{-1}*			
	All	*PCa*	*Non-PCa*	*P**	*All*	*PCa*	*Non-PCa*	*P**
Patient (*n*)	635	272	363		222	39	183	
Age (year), median (IQR)	69.0 (61.0–76.0)	72.0 (65.0–78.0)	66.0 (59.0–74.0)	5.11×10^{-10}	65.0 (58.8–72.2)	69.0 (63.0–75.0)	65.0 (58.0–72.0)	0.015
Volume (ml), median (IQR)	41.0 (32.0–56.1)	38.0 (29.0–51.0)	45.0 (34.0–62.0)	1.44×10^{-7}	40.0 (33.0–50.0)	36.0 (27.0–43.0)	40.0 (33.0–51.0)	6.0×10^{-3}
PSA (ng ml^{-1}), median (IQR)	13.3 (7.6–31.5)	31.8 (14.0–145.7)	9.4 (6.2–15.1)	3.06×10^{-39}	6.6 (4.7–8.1)	6.9 (5.2–8.4)	6.6 (4.7–8.1)	0.55
%fPSA (%), median (IQR)	14.0 (10.0–21.0)	12.0 (9.0–16.0)	17.0 (11.0–23.0)	3.55×10^{-10}	16.6 (11.6–22.0)	13.7 (11.3–19.7)	17.0 (12.4–22.2)	0.09
Phi, median (IQR)	48.4 (30.6–143.5)	158.7 (64.6–448.7)	34.0 (25.9–47.4)	1.44×10^{-59}	33.2 (24.5–41.5)	44.4 (33.3–64.5)	31.6 (23.9–39.3)	3.9×10^{-5}

*The *P* values were calculated using Mann–Whitney U test to observe whether there is any significant difference between the distributions of two groups. PCa: prostate cancer; IQR: interquartile range; PSA: prostate-specific antigen; %fPSA: percentage of free prostate-specific antigen; phi: prostate health index

Table 2: Evaluation of the area under the receiver operation curves of different risk calculators in all patients

Calculators	*PCa (n=268)[a]*			*High-grade PCa (n=207)[a]*			*PCa (n=438)[b]*			*High-grade PCa (n=347)[b]*		
	AUC	*95% CI*	*P**	*AUC*	*95% CI*	*P**	*AUC*	*95% CI*	*P**	*AUC*	*95% CI*	*P**
phi	0.88	0.85–0.91	0.015	0.89	0.86–0.91	0.06	0.91	0.89–0.93	1.3×10^{-9}	0.92	0.90–0.93	3.9×10^{-9}
RC1[c]	0.83	0.79–0.86	NA	0.84	0.81–0.88	NA	0.81	0.79–0.84	NA	0.82	0.79–0.85	NA
RC2[d]	0.89	0.86–0.91	8.1×10^{-3}	0.89	0.86–0.92	0.055	0.91	0.90–0.93	1.1×10^{-10}	0.92	0.90–0.94	8.9×10^{-10}
RC3[e]	0.86	0.83–0.89	0.085	0.86	0.83–0.9	0.35	0.84	0.81–0.86	0.17	0.85	0.83–0.88	0.084
RC4[f]	0.895	0.87–0.92	1.5×10^{-3}	0.89	0.86–0.92	0.032	0.91	0.89–0.93	1.3×10^{-9}	0.92	0.90–0.94	6.1×10^{-10}

**P* referred to the significance between RC1 and other risk calculator, [a]in training cohort (*n*=635); [b]in validation cohort (*n*=1045); [c]RC1 was built in the rule of logistic regression RC based on age, PSA, and %fPSA (%fPSA were excluded while predicting high-grade PCa), [d]RC2 were built in the rule of logistic regression RC based on age and phi; [e]RC3 was built in the rule of logistic regression RC based on age, PSA, %fPSA (%fPSA were excluded while predicting high-grade PCa), and prostate volume; [f]RC4 was built in the rule of logistic regression RC based on age, phi, and prostate volume. PCa: prostate cancer; PSA: prostate-specific antigen; phi: prostate health index; AUC: area under the receiver operating characteristic curve; CI: confidence interval; NA: not available; RC: risk calculator; %fPSA: percentage of free prostate-specific antigen

Table 3: Evaluation of the area under the receiver operation curves of different risk calculators in patients with prostate-specific antigen 2.0 ng ml^{-1}–10.0 ng ml^{-1}

Calculators	*PCa (n=39)*[a]			*High-grade PCa (n=24)*[a]			*PCa (n=105)*[b]			*High-grade PCa (n=17)*[b]		
	AUC	*95% CI*	*P**	*AUC*	*95% CI*	*P**	*AUC*	*95% CI*	*P**	*AUC*	*95% CI*	*P**
Phi	0.71	0.60–0.82	0.2	0.75	0.64–0.87	0.14	0.89	0.86–0.93	2.3×10^{-8}	0.9	0.85–0.95	2.0×10^{-3}
RC1[c]	0.61	0.51–0.72	NA	0.61	0.47–0.76	NA	0.71	0.65–0.76	NA	0.69	0.56–0.81	NA
RC2[d]	0.72	0.60–0.83	0.19	0.74	0.62–0.87	0.18	0.9	0.87–0.94	2.2×10^{-9}	0.91	0.87–0.96	6.7×10^{-4}
RC3[e]	0.69	0.6–0.79	0.28	0.63	0.5–0.76	0.87	0.73	0.67–0.78	0.62	0.68	0.55–0.80	0.93
RC4[f]	0.74	0.64–0.84	0.085	0.75	0.63–0.86	0.16	0.89	0.86–0.92	1.9×10^{-8}	0.9	0.84–0.95	2.1×10^{-3}

**P* referred to the significance between RC 1 and other risk calculators. [a]In training cohort (*n*=222); [b]In validation cohort (*n*=443); [c]RC1 were built in the rule of logistic regression RC based on age, PSA, and %fPSA (%fPSA were excluded while predicting high-grade PCa); [d]RC2 were built in the rule of logistic regression RC based on age and phi; [e]RC3 were built in the rule of logistic regression RC based on age, PSA, %fPSA (%fPSA were excluded while predicting high-grade PCa), and prostate volume; [f]RC4 were built in the rule of logistic regression RC based on age, phi and prostate volume. PCa: prostate cancer; PSA: prostate-specific antigen; phi: prostate health index; AUC: area under the receiver operating characteristic curve; CI: confidence interval; NA: not available; RC: risk calculator; %fPSA: percentage of free prostate-specific antigen

Table 4: Clinical characteristics of the validation cohort

Variables	*All PSA*				*PSA: 2.0 ng ml^{-1}–10.0 ng ml^{-1}*			
	All	*PCa*	*Non-PCa*	*P**	*All*	*PCa*	*Non-PCa*	*P**
Patient (*n*)	1045	449	596		443	106	337	
Age (year), median (IQR)	68.0 (62.0–74.0)	71.0 (65.0–76.0)	66.0 (61.0–72.0)	8.0×10^{-16}	66.0 (61.0–72.0)	69.0 (63.0–74.0)	65.0 (60.0–71.0)	9.2×10^{-4}
Volume (ml), median (IQR)	41.0 (31.2–58.3)	38.0 (30.2–53.2)	43.1 (31.2–62.0)	7.5×10^{-4}	40.0 (31.2–55.0)	33.7 (25.0–44.6)	42.2 (31.2–57.2)	7.4×10^{-5}
PSA (ng ml^{-1}), median (IQR)	11.7 (7.0–25.7)	24.2 (11.0–93.3)	8.8 (5.7–13.9)	9.7×10^{-58}	6.89 (5.33–8.57)	7.3 (5.9–8.7)	6.6 (5.0–8.5)	7.2×10^{-3}
%fPSA (%), median (IQR)	13.3 (9.2–19.3)	11.2 (8.1–13.8)	16.7 (11.1–23.0)	1.4×10^{-32}	17.0 (11.5–22.8)	12.3 (9.2–16.0)	18.8 (13.6–24.0)	4.8×10^{-13}
Phi, median (IQR)	45.3 (27.6–99.2)	114.0 (61.6–301.8)	30.9 (21.4–42.8)	1.5×10^{-114}	31.5 (21.4–45.5)	56.2 (43.5–70.1)	27.3 (19.7–37.1)	4.9×10^{-34}

**P* values were calculated using Mann–Whitney U test to observe whether there is any significant difference between the distributions of two groups. PCa: prostate cancer; IQR: interquartile range; PSA: prostate-specific antigen; %fPSA: percentage of free prostate-specific antigen; phi: prostate health index

Table 5: Assessments of the performance of the risk calculators in validation cohort (n=1045)

Performance measures	*Prostate cancer*				*High-grade disease*			
	RC1	*RC2*	*RC3*	*RC4*	*RC1*	*RC2*	*RC3*	*RC4*
Overall								
R^2 (Nagelkerke)	0.39	0.62	0.43	0.61	0.40	0.56	0.44	0.57
Calibration								
Predicted outcome (%)	35.77	31.87	37.33	34.89	19.49	21.25	25.34	22.51
Observed outcome (%)	43.00	43.00	43.00	43.00	34.16	34.16	34.16	34.16
Calibration slope (95% CI)	1.018 (0.924–1.111)	1.124 (1.056–1.192)	0.970 (0.889–1.051)	1.088 (1.023–1.152)	1.024 (0.934–1.114)	1.123 (1.05–1.197)	1.018 (0.935–1.101)	1.130 (1.058–1.202)
Hosmer–Lemeshow test	χ^2=26.09, *P*<0.001	χ^2=23.10, *P*=0.003	χ^2=10.27, *P*=0.25	χ^2=7.09, *P*=0.53	χ^2=9.63, *P*=0.29	χ^2=83.02, *P*<0.001	χ^2=13.61, *P*=0.09	χ^2=72.6, *P*<0.001

CI: confidence interval; RC: risk calculator

We specifically evaluated the potential reduction of unnecessary biopsies using phi-based risk calculators compared with RC1 at the PCa risk threshold of 40% (approximately the positive biopsy rate in the Chinese biopsy population) (**Table 6**). For instance, with a sensitivity of 90%, phi, RC2, and RC4 could spare 28.0%, 27.8%, and 25.60%, respectively, of the patients in the validation cohort who did not have PCa from undergoing an unnecessary invasive procedure.

DISCUSSION

Our previous study showed that %p2PSA and its derivative phi provide additional value to predicting PCa and high-grade PCa in the Chinese biopsy population.[10] This is the first study evaluating phi-based risk calculators for PCa in the Chinese population. All risk calculators were constructed based on the same previous population and were validated in an independent multicenter population. Our results indicated that the phi-based risk calculators (RC2 and RC4) exhibited superior discrimination and calibration in predicting PCa and high-grade PCa compared to the risk calculators without phi (RC1 and RC3). In addition, we also demonstrated that phi-based risk calculators provided added value for sparing unnecessary biopsies through a DCA.

The goal of the PCa diagnosis is to identify the presence of clinically significant disease while minimizing unnecessary biopsies. Combining PSA, PSA derivatives, other PCa biomarkers and patients' clinical information would predict the risk of PCa much more precisely than using PSA alone. Risk calculators for PCa had been developed and

Table 6: Number of unnecessary biopsies reduced by different prostate cancer risk calculators comparing with risk calculator 1[a]

RCs	Biopsies reduced in training cohort* (%)		Biopsies reduced in validation cohort*	
	Sensitivity=90%	Sensitivity=95%	Sensitivity=90%	Sensitivity=95%
Phi	21.00	11.20	28.00	16.40
RC2[b]	19.50	10.40	27.80	16.90
RC3[c]	8.90	7.60	9.80	0
RC4[d]	25.90	14.80	25.60	17.70

*The numbers of biopsies reduced were calculated at the sensitivity of 90 and 95 with the PCa risk threshold of 40. [a]RC1 was built in the rule of logistic regression RC based on age, PSA, and fPSA (fPSA were excluded while predicting high-grade PCa); [b]RC2 was built in the rule of logistic regression RC based on age and phi; [c]RC3 was built in the rule of logistic regression RC based on age, PSA, fPSA (fPSA were excluded while predicting high-grade PCa), and prostate volume; [d]RC4 was built in the rule of logistic regression RC based on age, phi, and prostate volume. PCa: prostate cancer; PSA: prostate-specific antigen; fPSA: free prostate-specific antigen; phi: prostate health index; RCs: risk calculators

validated in a variety of populations (*e.g.*, 0Caucasian, Hispanics, African Americans, and Chinese).[17,22,23]

P2PSA testing and phi have been approved for decision-making involving prostate biopsy in patients with PSA levels ranging from 2.0 ng ml^{-1} to 10.0 ng ml^{-1} (or 4.0–10.0 ng ml^{-1}) in the United States and Europe. Recently, the integration of phi into risk calculators was proven to further increase the accuracy of risk stratification and better inform the decision for prostate biopsy.[15,24–26] However, few studies in Chinese that deal with phi and the risk calculator have reported the relationship with high-grade PCa (Gleason score ≥7) or the predictive performance in the PSA gray zone (PSA 2.0–10.0 ng ml^{-1}).[10,26] These two questions might be of importance for PCa diagnosis as it is associated with aggressive clinical intervention and treatment. Thus, we looked into these two issues throughout the study and performed a comprehensive analysis based on our data.

A study has shown that the summary AUC for the PCPT risk calculator (without phi) was 0.66 (95% CI: 0.62–0.70), and the AUC for the ERSPC risk calculator (RC3 without phi) was 0.79 (95% CI: 0.77–0.81).[22] These two risk calculators had also been validated in Chinese populations with better predictive ability, which might be attributable to the fact that these populations were at higher risk for PCa.[20,27] Monique *et al*.[28] found that ERSPC + phi had an AUC of 0.72 for all PCa and 0.68 for clinically relevant PCa, and our RC2 and RC4 (phi-based) reached an AUC of 0.9 for both PCa and clinically relevant PCa in our validation cohort. This might be caused by two factors: the high risk of the study population mentioned above and the good cancer discriminating ability of phi in patients with PSA ≥10.0 ng ml^{-1} in the Chinese population. Because there is no complete PSA screening among the Chinese due to the large population and great health burden, most of the patients receive a PSA test under certain circumstances (*e.g.*, patients come to urologists for lower urinary tract symptoms, men with reliable health insurance, and men who pay out of pocket for a comprehensive physical examination). Basically, the biopsy population consists of patients with abnormal findings in the above situations. Therefore, we considered that the PCa risk calculators should be tailored more to a biopsy population rather than to a screening population in China. Thus, our study could represent the clinical situation in which Chinese urologists actually meet and help make the decision of biopsy together with their patients.

For the two important issues we mentioned above (performance of risk calculators in predicting high-grade PCa and in the PSA gray zone), we could observe that superiority was not obvious in the training cohort. However, when looking at the results in the larger validation cohort, the AUC reached 0.92 for high-grade PCa and 0.89–0.90 for PCa in the PSA gray zone (PSA 2.0 ng ml^{-1}–10.0 ng ml^{-1}). Therefore, we suggest that phi and risk calculators that include phi are useful tools for predicting high-grade PCa and for discriminating PCa or high-grade PCa from other conditions in patients with PSA levels ranging from 2.0 ng ml^{-1} to 10.0 ng ml^{-1}.

In addition, we constructed a DCA curve while applying phi-based risk calculators and demonstrated the net reduction in the number of biopsies. The net reduction in the number of biopsies was observed at PCa risk thresholds of approximately 20%, and it would be higher in a population with a higher risk of PCa. This indicates that these risk calculators might be suitable for clinically based cohorts.

Based on the detailed calculations, using phi-based risk calculators led to a reduction in the number of unnecessary biopsies by 20% while maintaining a sensitivity of 90%. While maintaining a sensitivity of 95%, the reduction rate was still up to 10%–18% in the training and validation cohorts. This finding indicated that a large proportion of unnecessary biopsies could be avoided by adding phi into the prebiopsy consideration. The number of biopsies reduced was calculated at the threshold risk of 40% because it approximated to the PCa detection rate in a Chinese biopsy population.[29]

The current study had several strengths: (i) this study was based on a large-scale multicenter biopsy population; (ii) we built logistical models based on a biopsy population, and we validated the risk calculators in a large independent biopsy population, which made our results solid and firm; and (iii) owing to the population size, we were able to evaluate the performance of phi and phi-based risk calculators in predicting high-grade PCa and in patients with PSA ranging from 2.0 ng ml^{-1} to 10.0 ng ml^{-1}, which was seldom reported in the Chinese population.

There are several limitations to our study. First, there might be inclusion bias because the four medical centers involved in this study were all located in Shanghai. However, all these medical centers were tertiary health institutes, and nearly half of the patients were from various provinces in China other than Shanghai. Second, the study population was at a relatively high risk for PCa, as we had mentioned previously (with a risk threshold of 40%); thus, the risk calculators built in the current study might be more applicable to other biopsy populations rather than to screening populations. Third, although we showed that phi-based risk calculators exhibited superior discrimination and calibration than risk calculators without phi, they did not perform better than phi alone. This could be attributed to either the relatively small sample size of the training cohort or the relatively small contribution of age and PV to the PCa risk. These risk calculators will be improved in further studies with larger populations.

CONCLUSION

The phi-based risk calculators performed better than PSA-based risk calculators in the Chinese population. In addition, a considerable reduction in the number of unnecessary biopsies for PCa was achieved by applying these risk calculators.

AUTHOR CONTRIBUTIONS

YSW, QD, JX, and YHS conceived and designed the study. RN, XLL, FL, NZ, and JG performed the experiments. YSW, XJF, and RN analyzed the data. DWY, JQ, HWJ, and YHS contributed materials and analysis tools. YSW, XJF, JX, and QD wrote the manuscript. All authors have read and approved the final version of the manuscript.

COMPETING INTERESTS

All authors declared no competing interests.

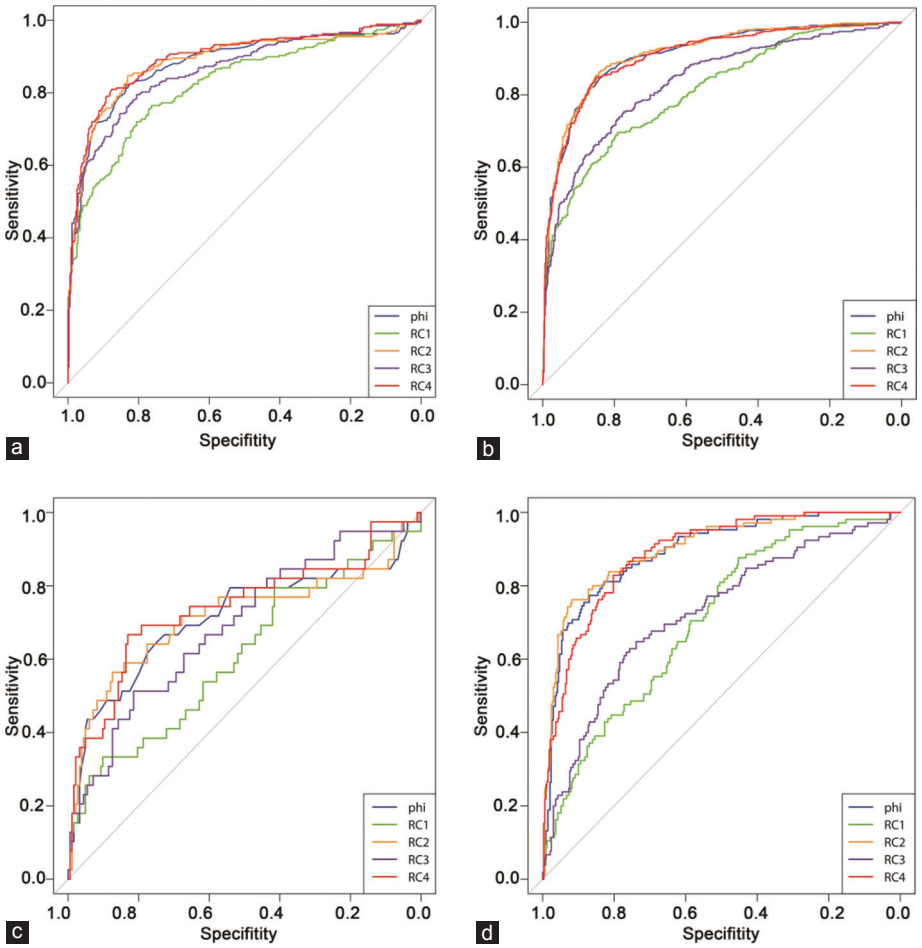
ACKNOWLEDGMENTS

This work was funded in part by grants from the Clinical Science and Technology Innovation Project of Shanghai Shen Kang Hospital Development Center to QD (No. SHDC12015105), the National Key Research and Development Plan of China, Precision Medicine Project (No. 2016YFC0902202), and the Scientific Research Project of Huashan Hospital, Fudan University to YSW (No. 2016QD079).

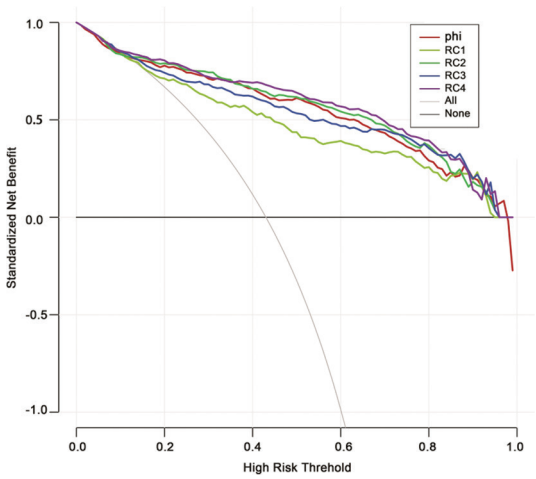
Supplementary information is linked to the online version of the paper on the *Asian Journal of Andrology* website.

REFERENCES

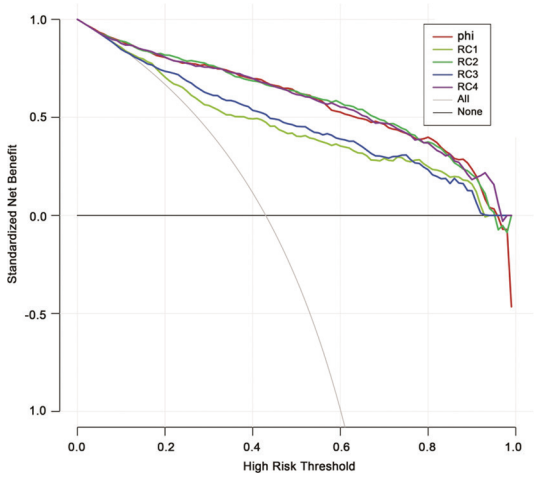
1 Jemal A, Bray F, Center MM, Ferlay J, Ward E, *et al*. Global cancer statistics. *CA Cancer J Clin* 2011; 61: 69–90.
2 Liu MW, Zhang YG, Zhu SC, Lu ZH, Wan B. [Detection of urological and male genital tumors diagnosed in Beijing hospital 1995-2004]. *Zhonghua Yi Xue Za Zhi* 2007; 87: 3. [Article in Chinese].
3 Chen W, Zheng R, Zeng H, Zhang S. The updated incidences and mortalities of major cancers in China, 2011. *Chin J Cancer* 2015; 34: 502–7.
4 Catalona WJ, Smith DS, Ratliff TL, Dodds KM, Coplen DE, *et al*. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *N Engl J Med* 1991; 324: 1156–61.
5 Loeb S, Bjurlin MA, Nicholson J, Tammela TL, Penson DF, *et al*. Overdiagnosis and overtreatment of prostate cancer. *Eur Urol* 2014; 65: 1046–55.
6 Hugosson J, Carlsson S, Aus G, Bergdahl S, Khatami A, *et al*. Mortality results from the Göteborg randomised population-based prostate-cancer screening trial. *Lancet Oncol* 2010; 11: 725–32.
7 Guazzoni G, Nava L, Lazzeri M, Scattoni V, Lughezzani G, *et al*. Prostate-Specific Antigen (PSA) isoform p2PSA significantly improves the prediction of prostate cancer at initial extended prostate biopsies in patients with total PSA between 2.0 and 10 ng/ml: results of a prospective study in a clinical setting. *Eur Urol* 2011; 60: 214–22.
8 Catalona WJ, Partin AW, Sanda MG, Wei JT, Klee GG, *et al*. A multicenter study of [-2]pro-prostate specific antigen combined with prostate specific antigen and free prostate specific antigen for prostate cancer detection in the 2.0 to 10.0 ng/ml prostate specific antigen range. *J Urol* 2011; 185: 1650–5.
9 Guazzoni G, Lazzeri M, Nava L, Lughezzani G, Larcher A, *et al*. Preoperative prostate-specific antigen isoform p2PSA and its derivatives, % p2PSA and prostate health index, predict pathologic outcomes in patients undergoing radical prostatectomy for prostate cancer. *Eur Urol* 2012; 61: 455–66.
10 Na R, Ye D, Liu F, Chen H, Qi J, *et al*. Performance of serum prostate-specific antigen isoform [-2]proPSA (p2PSA) and the prostate health index (PHI) in a Chinese hospital-based biopsy population. *Prostate* 2014; 74: 1569–75.
11 Thompson IM, Ankerst DP, Chi C, Goodman PJ, Tangen CM, *et al*. Assessing prostate cancer risk: results from the prostate cancer prevention trial. *J Natl Cancer Inst* 2006; 98: 529–34.
12 Roobol MJ, Steyerberg EW, Kranse R, Wolters T, van den Bergh RC, *et al*. A risk-based strategy improves prostate-specific antigen-driven detection of prostate cancer. *Eur Urol* 2010; 57: 79–85.
13 Cavadas V, Osorio L, Sabell F, Teves F, Branco F, *et al*. Prostate cancer prevention trial and European randomized study of screening for prostate cancer risk calculators: a performance comparison in a contemporary screened cohort. *Eur Urol* 2010; 58: 551–8.
14 Ankerst DP, Boeck A, Freedland SJ, Thompson IM, Cronin AM, *et al*. Evaluating the PCPT risk calculator in ten international biopsy cohorts: results from the Prostate Biopsy Collaborative Group. *World J Urol* 2012; 30: 181–7.
15 Roobol MJ, Vedder MM, Nieboer D, Houlgatte A, Vincendeau S, *et al*. Comparison of two prostate cancer risk calculators that include the prostate health index. *Eur Urol Focus* 2015; 1: 185–90.
16 Foley RW, Maweni RM, Gorman L, Murphy K, Lundon DJ, *et al*. European Randomised Study of Screening for Prostate Cancer (ERSPC) risk calculators significantly outperform the Prostate Cancer Prevention Trial (PCPT) 2.0 in the prediction of prostate cancer: a multi-institutional study. *BJU Int* 2016; 118: 706–13.
17 Eyre SJ, Ankerst DP, Wei JT, Nair PV, Regan MM, *et al*. Validation in a multiple urology practice cohort of the Prostate Cancer Prevention Trial calculator for predicting prostate cancer detection. *J Urol* 2009; 182: 2653–8.
18 Parekh DJ, Ankerst DP, Higgins BA, Hernandez J, Canby-Hagino E, *et al*. External validation of the Prostate Cancer Prevention Trial risk calculator in a screened population. *Urology* 2006; 68: 1152–5.
19 Roobol MJ, Zhu X, Schroder FH, van Leenders GJ, van Schaik RH, *et al*. A calculator for prostate cancer risk 4 years after an initially negative screen: findings from ERSPC rotterdam. *Eur Urol* 2013; 63: 627–33.
20 Zhu Y, Wang JY, Shen YJ, Dai B, Ma CG, *et al*. External validation of the prostate cancer prevention trial and the European randomized study of screening for prostate cancer risk calculators in a Chinese cohort. *Asian J Androl* 2012; 14: 738–44.
21 Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, *et al*. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010; 21: 128–38.
22 Louie KS, Seigneurin A, Cathcart P, Sasieni P. Do prostate cancer risk models improve the predictive accuracy of PSA screening? A meta-analysis. *Ann Oncol* 2015; 26: 848–64.
23 Nguyen CT, Yu C, Moussa A, Kattan MW, Jones JS. Performance of prostate cancer prevention trial risk calculator in a contemporary cohort screened for prostate cancer and diagnosed by extended prostate biopsy. *J Urol* 2010; 183: 529–33.
24 Foley RW, Gorman L, Sharifi N, Murphy K, Moore H, *et al*. Improving multivariable prostate cancer risk assessment using the Prostate Health Index. *BJU Int* 2016; 117: 409–17.
25 Zhu Y, Han CT, Zhang GM, Liu F, Ding Q, *et al*. Development and external validation of a prostate health index-based nomogram for predicting prostate cancer. *Sci Rep* 2015; 5: 15341.
26 Chiu PK, Roobol MJ, Teoh JY, Lee WM, Yip SY, *et al*. Prostate Health Index (PHI) and Prostate-Specific Antigen (PSA) predictive models for prostate cancer in the Chinese population and the role of digital rectal examination-estimated prostate volume. *Int Urol Nephrol* 2016; 48: 1631–7.
27 Wu YS, Zhang N, Liu SH, Xu JF, Tong SJ, *et al*. The Huashan risk calculators performed better in prediction of prostate cancer in Chinese population: a training study followed by a validation study. *Asian J Androl* 2016; 18: 925–9.
28 Roobol MJ, Vedder MM, Nieboer D, Houlgatte A, Vincendeau S, *et al*. Comparison of two prostate cancer risk calculators that include the prostate health index. *Eur Urol Focus* 2015; 1: 185-90.
29 Na R, Wu Y, Xu J, Jiang H, Ding Q. Age-specific prostate specific antigen cutoffs for guiding biopsy decision in Chinese population. *PLoS One* 2013; 8: e67585.



Supplementary Figure 1: ROC curves of phi, RC1, RC2, RC3, and RC4 for predicting prostate cancer in two cohorts and in subgroups with PSA ranged from 2.0 ng ml^{-1} to 10.0 ng ml^{-1}. (**a**). ROC curves of phi, RC1, RC2, RC3, and RC4 for predicting prostate cancer in the training cohort. (**b**) ROC curves of phi, RC1, RC2, RC3, and RC4 for predicting prostate cancer in the validation cohort. (**c**) ROC curves of phi, RC1, RC2, RC3, and RC4 for predicting prostate cancer in the subgroup of the training cohort with PSA ranging from 2.0 ng ml^{-1} to 10.0 ng ml^{-1}. (**d**) ROC curves of phi, RC1, RC2, RC3, and RC4 for predicting prostate cancer in the subgroup of the validation cohort with PSA ranging from 2.0 ng ml^{-1} to 10.0 ng ml^{-1}. ROC: receiver operating characteristic; PSA: prostate-specific antigen.



Supplementary Figure 2: Net benefits of phi and phi-based risk calculators for predicting prostate cancer in the training cohort.



Supplementary Figure 3: Net benefits of phi and phi-based risk calculators for predicting prostate cancer in the validation cohort.