Prostate volume does not provide additional predictive value to prostate health index for prostate cancer or clinically significant prostate cancer: results from a multicenter study in China

Da Huang1*, Yi-Shuo Wu2,3,7, Ding-Wei Ye4, Jun Qi4, Fang Liu4, Brian T Helfand6, Siqun L Zheng5, Qiang Ding2,3, Dan-Feng Xu1, Rong Na1,7, Jian-Feng Xu1,7, Ying-Hao Sun8

To evaluate whether prostate volume (PV) would provide additional predictive utility to the prostate health index (phi) for predicting prostate cancer (PCa) or clinically significant prostate cancer, we designed a prospective, observational multicenter study in two prostate biopsy cohorts. Cohort 1 included 595 patients from three medical centers from 2012 to 2013, and Cohort 2 included 1025 patients from four medical centers from 2013 to 2014. Area under the receiver operating characteristic curves (AUC) and logistic regression models were used to evaluate the predictive performance of PV-based derivatives and models. Linear regression analysis showed that both total prostate-specific antigen (tPSA) and free PSA (fPSA) were significantly correlated with PV (all P < 0.05). [-2]proPSA (p2PSA) was significantly correlated with PV in Cohort 2 (P < 0.001) but not in Cohort 1 (P = 0.309), while no significant association was observed between phi and PV. When combining phi with PV, phi density (PHID) and another phi derivative (PHIV, calculated as phi/PV) did not outperform phi for predicting PCa or clinically significant PCa in either Cohort 1 or Cohort 2. Logistic regression analysis also showed that phi and PV were independent predictors for both PCa and clinically significant PCa (all P < 0.05); however, PV did not provide additional predictive value to phi when combining these derivatives in a regression model (all models vs phi were not statistically significant, all P > 0.05). In conclusion, PV-based derivatives (both PHIV and PHID) and models incorporating PV did not improve the predictive abilities of phi for either PCa or clinically significant PCa.

INTRODUCTION

With the estimated 1,276,106 new cases and 358,989 deaths worldwide, prostate cancer (PCa) has become the second most common cancer and the fifth leading cause of cancer-specific death in males.1 Along with widespread prostate-specific antigen (PSA) screening and an in-depth understanding of PCa, PSA has been gradually considered an unspecific tumor biomarker, leading to large numbers of unnecessary prostate biopsies.2,3 Moreover, nonaggressive or low-grade PCa (also known as clinically insignificant disease) may not cause clinical consequences throughout the lifetime of a patient if left untreated or under surveillance, according to the results from several autopsy and active surveillance studies.4–11 These clinical issues are also known as overdiagnosis and overtreatment.

To improve PSA-based diagnostic ability, the prostate health index (phi), derived from total PSA (tPSA), free PSA (fPSA), and [-2]proPSA (p2PSA), has been introduced and shown to be a better predictor than tPSA and %PSA (fPSA/tPSA) for both PCa and clinically significant PCa. Using phi as a supplementary tool on tPSA may also reduce the number of unnecessary biopsies.15–21

Since many studies have reported that prostate volume (PV) is associated with prostate cancer and PSA density (PSAD, tPSA/PV) was introduced to adjust this influence and was shown to have better predictive value for PCa.22,23 However, phi, as a multivariable formula, did not include PV. Two previous single-center studies by Tosoian et al.24 and Druskin et al.25 demonstrated that phi density (PHID, calculated as phi/PV) outperformed phi in the diagnosis of clinically significant cancer (Gleason score, GS ≥7). A recent study by Vendrami et al.26 demonstrated that PHID has a greater predictive value than phi when prostate biopsies were guided by image fusion of magnetic resonance (MR) and transrectal ultrasound. However, due to the...
were diagnosed with PCa (42.9%) and 193 patients were diagnosed with clinically significant PCa (32.4%) in Cohort 1. In Cohort 2, 437 (42.6%) patients were diagnosed with PCa and 346 (33.8%) with clinically significant PCa.

The correlations between serum antigen indices (tPSA, p2PSA, fPSA, and phi) and PV were evaluated by simple linear regression analysis (Supplementary Table 2). In Cohort 1, we found that both tPSA and fPSA were significantly correlated with PV ($P < 0.005$ and $P < 0.001$, respectively). However, no significant association was found between p2PSA and PV and between phi and PV ($P = 0.309$ and $P = 0.107$, respectively). In Cohort 2, tPSA, p2PSA, and fPSA were significantly correlated with PV (all $P < 0.001$). Similarly, no significant association was found between phi and PV ($P = 0.434$).

In the entire study population and separate cohorts, the median PV was approximately 40 ml (entire Cohort: 41 ml; Cohort 1: 42 ml; Cohort 2: 41 ml; Supplementary Table 1). We then performed a stratified analysis for patients with different PV ($\leq 40$ ml and $>40$ ml). When stratified using 40 ml as a threshold, patients with smaller PV had significantly lower tPSA, lower %fPSA, and higher %p2PSA in two separate cohorts (all $P < 0.05$, Supplementary Table 3 and 4). In Cohort 1, there was no significant difference in phi between the two volume groups ($P = 0.081$; Supplementary Table 3). However, marginally significant differences in phi were found between patients with PV $\leq 40$ ml and $>40$ ml in Cohort 2 ($P = 0.047$; Supplementary Table 4).

The association between phi-PV derivatives and PCa or clinically significant PCa was also evaluated. In univariable logistic regression (Table 1), both PHID and PHIV (another phi derivative, calculated as phi/PV$^{0.5}$) were significantly associated with PCa and clinically significant PCa in the two cohorts (all $P < 0.001$). Notably, PHID had higher odds ratios (ORs) than phi when predicting PCa (OR$_{\text{PHID}} = 1.90$, OR$_{\text{phi}} = 1.02$) and clinically significant PCa (OR$_{\text{PHID}} = 1.43$, OR$_{\text{phi}} = 1.01$). Similar results were observed in Cohort 1 for the four LR models (LR-1, LR-2, LR-3, and LR-4) as described below (a) model LR-1/LR-2 predicted PCa/clinically significant PCa using the variables of age, PV, and phi; (b) model LR-3/LR-4 predicted PCa/clinically significant PCa using the variables of age, PV, %p2PSA, and tPSA.

Comparisons of the AUCs among the phi-PV derivatives, models, and phi are shown in Table 3 and 4. Briefly, the AUCs of PHID and PHIV did not outperform phi for predicting PCa or clinically significant PCa (Table 3 and 4; ROC curves, Supplementary Figure 2–4). Despite the overfitting effect of the models in Cohort 1, all models in Cohort 2 did not outperform phi (all $P > 0.05$) for predicting either PCa or clinically significant PCa. Similar results are shown in Supplementary Table 5 when predicting PCa with GS $\geq 8$. These results indicated that PV would not provide additional predictive value to phi.

**DISCUSSION**

The objective of this study was to evaluate the association between phi and PV and to determine whether PV would provide additional predictive value to phi. We found that (1) p2PSA was significantly associated with PV, while there was no association between phi and PV, and (2) neither phi-PV derivatives (PHID or PHIV) nor phi-PV multivariate models outperformed phi for predicting PCa or clinically significant PCa. These results suggested that phi might not be a useful predictor for PCa.
be influenced by PV and adding PV might not provide additional predictive value to phi.

A single-center study demonstrated that PHID outperformed phi in predicting clinically significant PCAs. The highest discriminative ability was observed for PHID in predicting clinically significant disease (with an AUC of 0.84), which was significantly higher than phi (AUC = 0.76). That study only included 118 men with elevated PSA and negative DRE who underwent a phi test and prostate biopsy, while all patients had a phi test in our study. In our multicenter study, with a two-step external validation, PHID did not significantly outperform phi for predicting PCa or clinically significant PCa. The differences observed by Tosoian et al. might be due to its relatively small sample size.

Based on our results, PV would not improve the predictive abilities of phi, suggesting that regardless of PV, phi alone could independently predict PCa and clinically significant PCa. Although the molecular mechanisms are not clear, there are several assumptions that might explain these results. First, p2PSA is considered a "prostate cancer-specific antigen" rather than a prostate-specific antigen. One study showed that p2PSA had higher immunostaining in prostate tumor tissues than in benign prostate tissues. Therefore, tumor volume rather than PV would be a major factor for p2PSA value. Second, both tPSA and fPSA have a positive linear association with PV. However, the influence of PV might have been partially adjusted using tPSA or fPSA.

There were several limitations of this study. First, the PVs were all calculated through a transrectal approach (TRUS), which might cause subjective error among different ultrasonologists. However, all volumes of prostate were measured by skilled ultrasonologists, with a minimum of 5 years of working experience in our study. A recent study demonstrated that PHID appears to have greater predictive performance than phi when prostate biopsies were guided by image fusion of MR and transrectal ultrasound. However, we were not able to perform similar analyses in the present study due to the lack of MRI data from our study subjects. MR-TRUS fusion biopsy will be applied in future studies to address this problem. Second, all medical centers participating in the present study were located in Shanghai, a large city in East China, which may cause selection bias. However, individuals all over the country seek the services of these tertiary hospitals. In conclusion, PV-based derivatives (both PHIV and PHID) and correlative models do not improve the predictive abilities of phi for both PCa and clinically significant PCa.

**AUTHOR CONTRIBUTIONS**

RN, DFX, JFX, and YHS conceived and designed the study. DWY, JQ, FL, BTH, SLZ, and QD contributed materials and collected the data.
Table 3: Area under receiver operating characteristic curves of different measurements for predicting prostate cancer in entire cohorts and subsets grouped by total prostate-specific antigen

| Measurements | Cohort 1 (entire, n=595) | Cohort 2 (entire, n=1025) | Cohort 1 (tPSA 2–10 ng ml⁻¹, n=211) | Cohort 2 (tPSA 2–10 ng ml⁻¹, n=433) | Cohort 1 (tPSA 10–20 ng ml⁻¹, n=171) | Cohort 2 (tPSA 10–20 ng ml⁻¹, n=243) |
|--------------|--------------------------|--------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| IPSA         | 0.82 (0.78–0.85)         | 0.78 (0.76–0.81)         | 0.56 (0.45–0.66)                   | 0.58 (0.52–0.64)                   | 0.74 (0.69–0.81)                   | 0.59 (0.52–0.67)                   |
| %PSA         | 0.65 (0.60–0.69)         | 0.71 (0.68–0.74)         | 0.58 (0.48–0.68)                   | 0.73 (0.68–0.78)                   | 0.61 (0.53–0.70)                   | 0.60 (0.52–0.67)                   |
| %p2PSA       | 0.84 (0.81–0.88)         | 0.87 (0.84–0.89)         | 0.68 (0.58–0.75)                   | 0.88 (0.84–0.92)                   | 0.78 (0.71–0.85)                   | 0.81 (0.75–0.87)                   |
| PSAD         | 0.86 (0.83–0.89)         | 0.84 (0.79–0.85)         | 0.65 (0.56–0.75)                   | 0.67 (0.61–0.73)                   | 0.74 (0.67–0.82)                   | 0.68 (0.61–0.75)                   |
| phi          | 0.88 (0.85–0.91)         | 0.91 (0.89–0.93)         | 0.69 (0.58–0.81)                   | 0.89 (0.85–0.93)                   | 0.79 (0.72–0.86)                   | 0.82 (0.77–0.88)                   |
| PHID         | 0.89 (0.86–0.92)         | 0.89 (0.87–0.91)         | 0.72 (0.62–0.82)                   | 0.84 (0.81–0.88)                   | 0.80 (0.74–0.87)                   | 0.79 (0.73–0.85)                   |
| PHIV         | 0.89 (0.86–0.92)         | 0.91 (0.89–0.92)         | 0.72 (0.61–0.83)                   | 0.88 (0.84–0.91)                   | 0.80 (0.74–0.87)                   | 0.81 (0.75–0.87)                   |
| LR-1         | 0.91 (0.87–0.93)         | 0.90 (0.89–0.92)         | 0.78 (0.69–0.88)                   | 0.88 (0.84–0.91)                   | 0.84 (0.78–0.90)                   | 0.84 (0.77–0.88)                   |
| LR-3         | 0.90 (0.88–0.94)         | 0.90 (0.88–0.92)         | 0.78 (0.69–0.88)                   | 0.86 (0.82–0.90)                   | 0.84 (0.78–0.90)                   | 0.81 (0.76–0.87)                   |

* P value, statistical analysis (DeLong method) between AUCs of different variables. ROC: receiver operating characteristic; AUCs: area under ROC curves; PSA: prostate-specific antigen; tPSA: total PSA; fPSA: free PSA; p2PSA: [-2]proPSA; PSAD: PSA density; phi: prostate health index; PHID: phi density; PHIV: phi/PV; LR-1: the first logistic regression model; LR-3: the third logistic regression model; CI: confidence interval.

Table 4: Area under receiver operating characteristic curves of different measurements for predicting clinically significant prostate cancer (Gleason score ≥ 7) in entire cohorts and subsets grouped by total prostate-specific antigen

| Measurements | Cohort 1 (entire, n=595) | Cohort 2 (entire, n=1025) | Cohort 1 (tPSA 2–10 ng ml⁻¹, n=211) | Cohort 2 (tPSA 2–10 ng ml⁻¹, n=433) | Cohort 1 (tPSA 10–20 ng ml⁻¹, n=171) | Cohort 2 (tPSA 10–20 ng ml⁻¹, n=243) |
|--------------|--------------------------|--------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| IPSA         | 0.85 (0.81–0.88)         | 0.81 (0.79–0.84)         | 0.59 (0.44–0.73)                   | 0.60 (0.53–0.67)                   | 0.63 (0.54–0.73)                   | 0.59 (0.50–0.67)                   |
| %PSA         | 0.64 (0.59–0.68)         | 0.70 (0.67–0.74)         | 0.58 (0.47–0.70)                   | 0.68 (0.61–0.75)                   | 0.63 (0.52–0.73)                   | 0.63 (0.55–0.70)                   |
| %p2PSA       | 0.83 (0.79–0.87)         | 0.87 (0.84–0.89)         | 0.72 (0.59–0.86)                   | 0.88 (0.85–0.92)                   | 0.72 (0.62–0.81)                   | 0.81 (0.75–0.87)                   |
| PSAD         | 0.87 (0.84–0.90)         | 0.84 (0.82–0.87)         | 0.59 (0.46–0.73)                   | 0.71 (0.65–0.78)                   | 0.80 (0.73–0.87)                   | 0.67 (0.60–0.75)                   |
| phi          | 0.88 (0.85–0.91)         | 0.92 (0.90–0.93)         | 0.73 (0.61–0.86)                   | 0.90 (0.87–0.94)                   | 0.73 (0.64–0.82)                   | 0.82 (0.76–0.88)                   |
| PHID         | 0.88 (0.85–0.91)         | 0.90 (0.88–0.92)         | 0.74 (0.63–0.85)                   | 0.87 (0.83–0.91)                   | 0.77 (0.69–0.86)                   | 0.79 (0.73–0.86)                   |
| PHIV         | 0.89 (0.86–0.92)         | 0.91 (0.90–0.93)         | 0.75 (0.63–0.87)                   | 0.90 (0.87–0.93)                   | 0.77 (0.68–0.85)                   | 0.81 (0.75–0.87)                   |
| LR-2         | 0.90 (0.87–0.92)         | 0.92 (0.90–0.93)         | 0.80 (0.70–0.90)                   | 0.91 (0.87–0.94)                   | 0.79 (0.72–0.86)                   | 0.83 (0.77–0.89)                   |
| LR-4         | 0.90 (0.88–0.94)         | 0.90 (0.89–0.92)         | 0.77 (0.67–0.87)                   | 0.88 (0.84–0.92)                   | 0.81 (0.74–0.88)                   | 0.81 (0.75–0.87)                   |

* P value, statistical analysis (DeLong method) between AUCs of different variables. ROC: receiver operating characteristic; AUC: area under ROC curve; PSA: prostate-specific antigen; tPSA: total PSA; fPSA: free PSA; p2PSA: [-2]proPSA; PSAD: PSA density; phi: prostate health index; PHID: phi density; PHIV: phi/PV; LR-2: the second logistic regression model; LR-4: the fourth logistic regression model; CI: confidence interval.
DH, YSW, RN, and DFX analyzed the data. DH, YSW, RN, and DFX wrote the manuscript. RN, DFX, JFX, and YHS supervised the study. DH and YSW contributed equally to this study. All authors have read and approved the final manuscript.

COMPETING INTERESTS

In the present study, we declare that Beckman Coulter, Inc., provided the tests for tPSA, fPSA, and p2PSA. All the sample tests, data analyses, and manuscript writing were performed by the researchers, independent from Beckman Coulter, Inc. There are no other potential competing interests to be declared.

ACKNOWLEDGMENTS

We thank all subjects included in this study. We thank Beckman Coulter, Inc., for the measurements of tPSA, fPSA, and p2PSA. This work was in part supported by grants from the innovation grant by Shanghai Hospital Development Center (SHDC12015105) to Jianfeng Xu, the National Natural Science Foundation, Shanghai Municipal Education Commission-Gaofeng Clinical Center (Grant No. 18QA1402800), the “Chen Guang” project supported by Shanghai Municipal Education Commission, and Shanghai Education Development Foundation, Shanghai Municipal Education Commission-Gaofeng Clinical Medicine Grant Support (Grant No. 20181701) to Rong Na.

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

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68: 394–424.
2. Welch HG, Schwartz LM, Woloshin S. Prostate-specific antigen levels in the United States: implications of various definitions for abnormal. J Natl Cancer Inst 2005; 97: 1132–7.
3. Ainkerst DP, Thompson IM. Sensitivity and specificity of prostate-specific antigen for prostate cancer detection with high rates of biopsy verification. Arch Ital Urol Androl 2006; 78: 125–9.
4. Sakr WA, Grignon DJ, Crissman JD, Heilbrun LK, Cassin BJ, et al. High grade prostatic intraepithelial neoplasia (HGPIN) and prostatic adenocarcinoma between the ages of 20-69: an autopsy study of 249 cases. In Vivo (Athens, Greece) 1994; 8: 439–43.
5. Sánchez-Chapado M, Olmedilla G, Cabeza M, Donat E, Ruiz A. Prevalence of prostate cancer and prostatic intraepithelial neoplasia in Caucasian Mediterranean males: an autopsy study. Prostate 2003; 54: 238–47.
6. Soos G, Tkács I, Szántó J, Turzo C, Haas PG, et al. The prevalence of prostate carcinoma and its precursor in Hungary: an autopsy study. Eur Urol 2008; 48: 739–44.
7. Stamatouli K, Alexopoulos I, Perimenis D, Sofras F, Agapitos E. Frequency of impalpable prostate adenocarcinoma and precancerous conditions in Greek male population: an autopsy study. Prostate Cancer Prostatic Dis 2006; 9: 45.
8. Haas GP, Delongchamps NB, Jones RF, Chandan V, Serio AM, et al. Needle biopsies on autopsy samples: sensitivity of cancer detection based on true prevalence. J Natl Cancer Inst 2007; 99: 1484–9.
9. Zlotta AR, Egawa S, Pushkar D, Grafonov A, Kimura T, et al. Prevalence of prostate cancer on autopsy: cross-sectional study on unscreened Caucasian and Asian men. J Natl Cancer Inst 2013; 105: 1050–8.
10. Drazin G, Boer R, Otto SJ, van der Cruysen IW, Damhuis RA, et al. Lead times and overdiagnosis due to prostate-specific antigen screening: estimates from the European randomised study of screening for prostate cancer. J Natl Cancer Inst 2003; 95: 686–78.
11. Zappa M, Ciatto S, Bonardi R, Mazzotta A. Overdiagnosis of prostate carcinoma by screening: an estimate based on the results of the Florence Screening Pilot Study.
### Supplementary Table 1: Characteristics of Cohort 1 and 2 and comparison of each variable between cohorts

| Variables          | Entire cohort (n=1620) | Cohort 1 (n=595) | Cohort 2 (n=1025) | P     |
|--------------------|------------------------|------------------|-------------------|-------|
|                    | Median (range)         |                  |                   |       |
| Age (year)         | 69 (62–74)             | 69 (62–76)       | 68 (62–74)        | 0.014 |
| PV (ml)            | 41 (31–58)             | 42 (33–57)       | 41 (31–59)        | 0.045 |
| tPSA (ng ml\(^{-1}\)) | 12.16 (7.29–26.60)    | 13.14 (7.60–30.62) | 11.65 (7.08–25.51) | 0.031 |
| p2PSA (pg ml\(^{-1}\)) | 21.68 (12.43–58.30)   | 23.74 (13.21–84.89) | 20.78 (12.02–48.44) | 0.001 |
| %fPSA              | 0.14 (0.10–0.20)       | 0.14 (0.10–0.21) | 0.13 (0.09–0.19)  | 0.121 |
| %p2PSA             | 14.99 (9.70–24.00)     | 15.33 (10.28–25.61) | 14.79 (9.47–22.6) | 0.031 |
| phi                | 46.35 (28.94–108.37)   | 48.45 (30.81–139.70) | 44.73 (27.47–98.07) | 0.009 |
| # (%) of positive  | 692 (42.7)             | 255 (42.9)       | 437 (42.6)        | 0.930 |
| PCa (GS ≥7)        | 539 (33.3)             | 193 (32.4)       | 346 (33.8)        | 0.587 |
| PCa (GS ≥8)        | 287 (17.7)             | 102 (17.1)       | 185 (18.0)        | 0.645 |

PV: prostate volume; PSA: prostate-specific antigen; tPSA: total PSA; fPSA: free PSA; p2PSA: [-2]proPSA; phi: prostate health index; PCa: prostate cancer; GS: Gleason score

### Supplementary Table 2: Simple linear regression between serum indices and prostate volume in entire cohort and two separate cohorts

| Variables | Cohort | n  | Coefficient (95% CI) | R\(^2\) | P     |
|-----------|--------|----|----------------------|--------|-------|
| tPSA      | 1      | 595| 0.359 (0.107–0.611)   | 0.013  | 0.005 |
|           | 2      | 1025| 0.501 (0.359–0.642)   | 0.045  | <0.001|
|           | Entire | 1620| 0.469 (0.343–0.594)   | 0.032  | <0.001|
| p2PSA     | 1      | 595| 0.169 (–0.157–0.495)  | 0.002  | 0.309 |
|           | 2      | 1025| 0.566 (0.382–0.751)   | 0.034  | <0.001|
|           | Entire | 1620| 0.470 (0.307–0.633)   | 0.020  | <0.001|
| fPSA      | 1      | 595| 0.544 (0.299–0.789)   | 0.031  | <0.001|
|           | 2      | 1025| 0.759 (0.629–0.889)   | 0.114  | <0.001|
|           | Entire | 1620| 0.709 (0.591–0.827)   | 0.079  | <0.001|
| phi       | 1      | 595| –0.195 (–0.433–0.042) | 0.004  | 0.107 |
|           | 2      | 1025| 0.058 (–0.087–0.202)  | 0.0006 | 0.434 |
|           | Entire | 1620| –0.004 (–0.128–0.119) | 3.0×10^-6 | 0.944 |

PSA: prostate-specific antigen; tPSA: total PSA; fPSA: free PSA; p2PSA: [-2]proPSA; phi: prostate health index; CI: confidence interval
### Supplementary Table 3: Characteristics of Cohort 1 and comparison of each variable between different groups by prostate volume

| Variables | Entire cohort (n=595) | Volume (ml) | P* |
|-----------|-----------------------|-------------|----|
|           | ≤40 (n=275)           | >40 (n=320) |    |
| Age (year)| 69 (62–76)            | 69 (60–75)  | 70 (63–77) | 0.017 |
| tPSA (ng ml⁻¹) | 13.14 (7.60–30.62) | 12.11 (6.64–24.22) | 14.34 (8.44–38.21) | 0.003 |
| p2PSA (pg ml⁻¹) | 23.74 (13.21–84.89) | 21.87 (11.73–73.66) | 25.73 (14.38–114.46) | 0.062 |
| %fPSA     | 0.14 (0.10–0.21)      | 0.13 (0.09–0.18) | 0.15 (0.11–0.22) | 0.001 |
| %p2PSA    | 15.33 (10.28–25.61)   | 18.53 (11.81–26.07) | 13.58 (8.91–24.01) | <0.001 |
| phi       | 48.45 (30.81–139.70)  | 58.78 (32.93–136.53) | 43.30 (30.19–146.57) | 0.081 |
| # (%) of positive PCa | 255 (42.9) | 137 (49.8) | 118 (36.9) | 0.001 |
| PCa (GS ≥7) | 193 (32.4) | 101 (36.7) | 92 (28.8) | 0.038 |
| PCa (GS ≥8) | 102 (17.1) | 50 (18.0) | 52 (16.3) | 0.533 |

*Difference in continuous variables is evaluated by Mann–Whitney U-test, and categorical variables by Chi-squared test. PSA: prostate-specific antigen; tPSA: total PSA; fPSA: free PSA; p2PSA: [-2]proPSA; phi: prostate health index; PCa: prostate cancer; GS: Gleason score.

### Supplementary Table 4: Characteristics of Cohort 2 and comparison of each variable between different groups by prostate volume

| Variables | Entire cohort (n=1025) | Volume (ml) | P* |
|-----------|-----------------------|-------------|----|
|           | ≤40 (n=500)           | >40 (n=525) |    |
| Age (year)| 68 (62–74)            | 68 (61–73)  | 68 (63–74) | 0.011 |
| tPSA (ng ml⁻¹) | 11.65 (7.08–25.51) | 10.68 (5.96–22.83) | 12.48 (8.19–28.13) | <0.001 |
| p2PSA (pg ml⁻¹) | 20.78 (12.02–48.44) | 18 (10.32–44.91) | 23.23 (14.99–52.61) | <0.001 |
| %fPSA     | 0.13 (0.09–0.19)      | 0.12 (0.08–0.16) | 0.15 (0.11–0.22) | <0.001 |
| %p2PSA    | 14.79 (9.47–22.6)     | 17.01 (11.18–24.42) | 12.49 (8.61–21.21) | <0.001 |
| phi       | 44.73 (27.47–98.07)   | 49.75 (29.46–98.06) | 42.35 (25.93–98.07) | 0.047 |
| # (%) of positive PCa | 437 (42.6) | 239 (47.8) | 198 (37.7) | 0.001 |
| PCa (GS ≥7) | 346 (33.8) | 188 (37.6) | 158 (30.1) | 0.011 |
| PCa (GS ≥8) | 185 (18.0) | 92 (18.4) | 93 (17.7) | 0.775 |

*Difference in continuous variables is evaluated by Mann–Whitney U-test, and categorical variables by Chi-squared test. PSA: prostate-specific antigen; tPSA: total PSA; fPSA: free PSA; p2PSA: [-2]proPSA; phi: prostate health index; PCa: prostate cancer; GS: Gleason score.
Supplementary Table 5: Area under receiver operating characteristic curves of different measurements for predicting prostate cancer (Gleason score ≥8) in entire cohorts and subsets grouped by total prostate-specific antigen

| Measurements | Cohort 1 (entire, n=595) | AUC (95% CI) | P | Cohort 2 (entire, n=1025) | AUC (95% CI) | P | Cohort 1 (tPSA 2–10, n=211) | AUC (95% CI) | P | Cohort 2 (tPSA 2–10, n=433) | AUC (95% CI) | P | Cohort 1 (tPSA 10–20, n=171) | AUC (95% CI) | P | Cohort 2 (tPSA 10–20, n=243) | AUC (95% CI) | P |
|--------------|--------------------------|--------------|---|--------------------------|--------------|---|--------------------------|--------------|---|--------------------------|--------------|---|--------------------------|--------------|---|--------------------------|--------------|---|
| tPSA         | 0.88 (0.85–0.91)         | 0.87 (0.84–0.90) | P | 0.75 (0.48–1.00)         | 0.65 (0.50–0.80) | P | 0.693 (0.537–0.849)       | 0.65 (0.53–0.77) | P |
| %fPSA        | 0.59 (0.52–0.65)         | 0.66 (0.63–0.70) | P | 0.63 (0.37–0.89)         | 0.51 (0.37–0.66) | P | 0.589 (0.402–0.775)       | 0.62 (0.49–0.75) | P |
| %p2PSA       | 0.81 (0.76–0.86)         | 0.83 (0.80–0.87) | P | 0.59 (0.26–0.92)         | 0.67 (0.49–0.86) | P | 0.666 (0.503–0.829)       | 0.76 (0.65–0.87) | P |
| PSAD         | 0.89 (0.86–0.92)         | 0.87 (0.84–0.90) | P | 0.94 (0.89–0.99)         | 0.007         | P | 0.61 (0.47–0.76)          | 0.433         | P | 0.831 (0.753–0.910)       | 0.116         | P | 0.75 (0.66–0.85)          | 0.679         | P |
| phi          | 0.87 (0.83–0.91)         | Reference     | P | 0.90 (0.87–0.92)         | Reference     | P | 0.51 (0.16–0.86)          | Reference     | P | 0.70 (0.53–0.88)          | Reference     | P | 0.689 (0.527–0.852)       | Reference     | P |
| PHID         | 0.87 (0.83–0.91)         | 0.981         | P | 0.88 (0.85–0.91)         | 0.003         | P | 0.69 (0.44–0.94)          | 0.003         | P | 0.67 (0.50–0.83)          | 0.240         | P | 0.740 (0.590–0.890)       | 0.216         | P |
| PHIV         | 0.87 (0.83–0.91)         | 0.474         | P | 0.89 (0.86–0.92)         | 0.163         | P | 0.61 (0.29–0.93)          | <0.001        | P | 0.69 (0.51–0.86)          | 0.345         | P | 0.724 (0.570–0.878)       | 0.200         | P |
| LR-2         | 0.87 (0.83–0.91)         | 0.807         | P | 0.89 (0.86–0.92)         | 0.281         | P | 0.72 (0.47–0.98)          | 0.008         | P | 0.72 (0.56–0.88)          | 0.434         | P | 0.742 (0.609–0.874)       | 0.265         | P |
| LR-4         | 0.89 (0.86–0.92)         | 0.058         | P | 0.89 (0.87–0.92)         | 0.715         | P | 0.84 (0.68–0.99)          | 0.004         | P | 0.74 (0.60–0.88)          | 0.439         | P | 0.780 (0.681–0.880)       | 0.150         | P |

*P: P value, statistical analysis (DeLong Method) between AUCs of different variables. PSA: prostate-specific antigen; tPSA: total prostate-specific antigen; fPSA: free prostate-specific antigen; p2PSA: [-2]proPSA; PSAD: PSA density; phi: prostate health index; PHID: phi density; PHIV: phi/PV0.5; LR-n: the nth logistic regression model; PCa: prostate cancer; GS: Gleason score; CI: confidence interval; ROC: receiver operating characteristic; AUC: area under ROC curve
Supplementary Figure 1: Flowchart of study population enrollment based on inclusion and exclusion criteria.
Supplementary Figure 2: ROC curves of different measurements in a subset (tPSA 2–10 ng ml$^{-1}$) (1) for PCa in Cohort 1; (2) for PCa in Cohort 2; (3) for PCa (GS $\geq$7) in Cohort 1; (4) for PCa (GS $\geq$7) in Cohort 2. PCa: prostate cancer; PSA: prostate-specific antigen; ROC: receiver operating characteristic; GS: Gleason score.
Figure 3: ROC curves of different measurements in a subset (IPSA 10–20 ng ml\(^{-1}\)) (1) for PCa in Cohort 1; (2) for PCa in Cohort 2; (3) for PCa (GS ≥ 7) in Cohort 1; (4) for PCa (GS ≥ 7) in Cohort 2. PCa: prostate cancer; PSA: prostate-specific antigen; ROC: receiver operating characteristic; GS: Gleason score.
Supplementary Figure 4: ROC curves of different measurements in entire cohorts (1) for PCa in Cohort 1; (2) for PCa in Cohort 2; (3) for PCa (GS ≥ 7) in Cohort 1; (4) for PCa (GS ≥ 7) in Cohort 2. PCa: prostate cancer; ROC: receiver operating characteristic; GS: Gleason score.