

RESEARCH

Open Access



# Prostate cancer risk prediction based on clinical factors and prostate-specific antigen

Taewon Hwang<sup>1,2</sup>, Hyungseok Oh<sup>1</sup>, Jung Ah Lee<sup>1\*</sup> and Eo Jin Kim<sup>3\*</sup>

## Abstract

**Introduction** The incidence rate of prostate cancer (PCa) has continued to rise in Korea. This study aimed to construct and evaluate a 5-year PCa risk prediction model using a cohort with PSA < 10 ng/mL by incorporating PSA levels and individual factors.

**Methods** The PCa risk prediction model including PSA levels and individual risk factors was constructed using a cohort of 69,319 participants from the Kangbuk Samsung Health Study. 201 registered PCa incidences were observed. A Cox proportional hazards regression model was used to generate the 5-year risk of PCa. The performance of the model was assessed using standards of discrimination and calibration.

**Results** The risk prediction model included age, smoking status, alcohol consumption, family history of PCa, past medical history of dyslipidemia, cholesterol levels, and PSA level. Especially, an elevated PSA level was a significant risk factor of PCa (hazard ratio [HR]: 1.77, 95% confidence interval [CI]: [1.67–1.88]). This model performed well with sufficient discrimination ability and satisfactory calibration (C-statistic: 0.911, 0.874; Nam-D'Agostino test statistic: 19.76, 4.21 in the development and validation cohort, respectively).

**Conclusions** Our risk prediction model was effective in predicting PCa in a population according to PSA levels. When PSA levels are inconclusive, an assessment of both PSA and specific individual risk factors (e.g., age, total cholesterol, and family history of PCa) could provide further information in predicting PCa.

## Highlights

- There is a lack of prostate cancer prediction models available in primary care.
- This study suggested a 5 year risk model by combining PSA levels and clinical factors.
- This model can be used for classifying high risk individuals in clinical practice.

**Keywords** Prostate cancer, Prostate-specific antigen, Prediction model, Lifestyle risk factor, Clinical factor

\*Correspondence:

Jung Ah Lee  
ljunga00@naver.com  
Eo Jin Kim  
kej7515@hanmail.net

<sup>1</sup>Workplace Health Institute, Total Health Care Center, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, B1, 55 Sejong-daero, Jung-gu, 06521 Seoul, South Korea

<sup>2</sup>Department of Economics, Texas A&M University, 4228 TAMU, 77843 College Station, TX, USA

<sup>3</sup>Division of Hematology/Oncology, Department of Internal Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, 29 Saemunan-ro, Jongno-gu, 03181 Seoul, South Korea



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Introduction

Prostate cancer (PCa) is the second-most common cancer in men worldwide [1]. In Korea, the crude incidence rate of PCa was 58 per 100,000 people in 2018, accounting for 11.5% of all newly diagnosed cancer cases in men [2]. While this is generally lower than numbers from Western countries, the incidence rate has continued to rise in Korea [2]. This is often attributed to the increase in life expectancy and health risk behaviors such as smoking, Western dietary patterns, and sedentary lifestyles, but also to the increased monitoring of serum prostate-specific antigen (PSA) levels [3].

Given the strong association between PSA levels and PCa, PSA is considered an important biomarker for PCa screening [4]. However, PSA is not a cancer-specific marker and can be elevated in non-malignant prostatic conditions such as benign prostate hypertrophy and prostatitis [5]. Other limitations of the PSA test are related with its diagnostic cut-off points. The number of false-positive or false-negative results could increase unnecessary biopsies. To improve the efficacy of assessing PSA in predicting PCa risk, previous studies have proposed risk assessment tools based on additional measurable factors such as free PSA, precursor of PSA, and the prostate health index [6]. These factors can improve the specificity of PCa screening and lead to a more accurate estimation of PCa risk [6, 7]. However, there are few PCa risk prediction models available for utilization in a primary healthcare or community health setting [8].

Distinguished from risk models based on biomarkers, few studies have incorporated epidemiological lifestyle risk factors for PCa risk modeling in the Korean population [9, 10]. Even without the inclusion of PSA levels, these models reported satisfactory performance. However, studies that combine multiple lifestyle factors and PSA levels in risk prediction are scarce, even though PSA levels are widely used biomarkers for PCa screening. Therefore, we aimed to investigate the performance of PCa risk models with the inclusion of PSA levels as a risk factor. In particular, we attempted to develop a clinically useful PCa prediction model when PSA levels are inconclusive, especially in a primary care setting.

## Materials and methods

### Data and study population

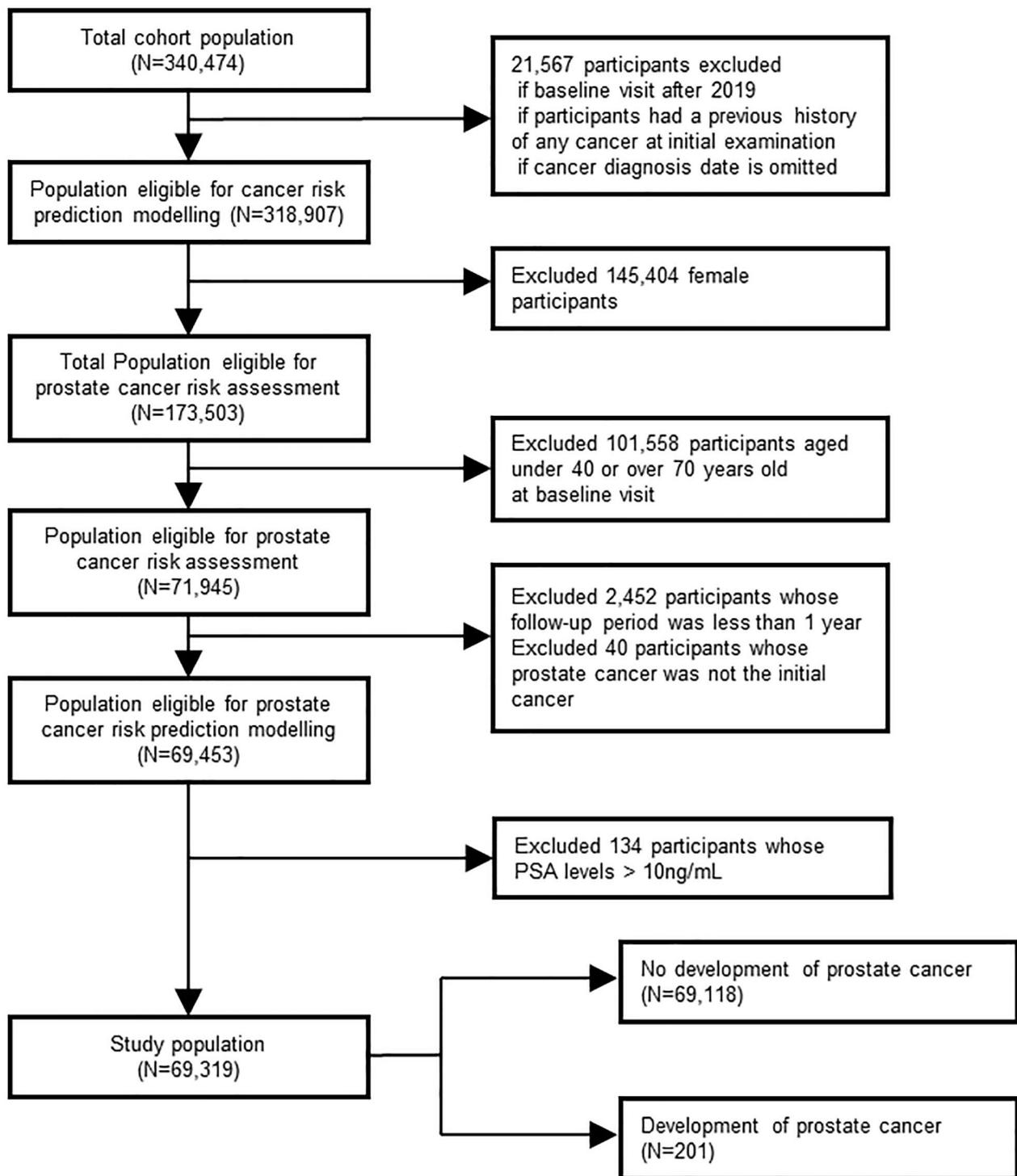
In South Korea, the Industrial Safety and Health Law mandates regular health screening examinations for all employees at no expense. Our study utilized data from the Kangbuk Samsung Health Study cohort, which is comprised of Korean men and women aged  $\geq 18$  years who received an annual/biennial health examination at one of the Kangbuk Samsung Hospital Total Healthcare Centers in Seoul or Suwon, South Korea [11]. We matched the participant data with the Korean Central

Cancer Registry provided by the National Cancer Center of Korea to identify patients with PCa up to 2019. This study identified PCa incidences as “C61 (Malignant neoplasm of prostate)” based on the 10th edition of the International Statistical Classification of Diseases and Related Health Problems (ICD-10). This study complied with the Declaration of Helsinki and was approved by the Institutional Review Board of Kangbuk Samsung Hospital, which waived the requirement for informed consent because of the use of anonymized data routinely collected as part of a health checkup program linked to mortality data from the Korea National Statistical Office (IRB No. 2011-01-030-005 for the general Kangbuk Samsung Health Study protocol and 2021-08-046 for the present study).

To construct risk prediction models for PCa, we included men between the age of 40 to 70 who had their initial health examination between January 1, 2011 and December 31, 2018 based on the US Preventive Services Task Force (USPSTF) recommendation against screening in men aged  $\geq 70$  years [12]. To assess the risk factors associated with the initial diagnosis of PCa, we excluded participants with a history of cancer at their baseline visit or who were initially diagnosed with any other type of cancer. Further, we excluded participants whose follow-up period was less than 1 year after cohort enrollment to reducing the possibility of preexisting prostate cancer at the time of cohort enrollment. Finally, we excluded individuals with elevated PSA levels ( $> 10$  ng/mL) at baseline because they were likely to be strongly recommended by their clinician for further screening. Ultimately, 69,319 participants were eligible for this study, of whom 201 had registered incidences of PCa (Fig. 1).

Data on behavioral factors and medical history were collected using a standardized, self-administered questionnaire included in the health examination. A family history of cancer, past medical history of dyslipidemia, alcohol consumption, and smoking status were collected from the questionnaire. Current smokers were defined as individuals who had a lifetime history of smoking 100 or more cigarettes and were currently smoking. Former smokers were defined as individuals who had a lifetime history of smoking 100 or more cigarettes but were “not at all” currently smoking. Alcohol consumption behavior was categorized based on the frequency and amount of alcohol consumed per drinking day; we classified alcohol consumption in two categories: non- or light drinking, and heavy drinking ( $< 20$ ,  $\geq 20$  g/day).

Participants with hyperlipidemia risk were defined as those with high total cholesterol ( $\geq 200$  mg/dL) or those currently on medication for dyslipidemia. For this study, we included PSA levels with the intent of reflecting the incremental risk of PCa according to increasing PSA levels [9, 13]. The Laboratory Medicine Department



**Fig. 1** Flowchart of the study population

at the Kangbuk Samsung Hospital in Seoul, Korea has been accredited by the Korean Society of Laboratory Medicine, the Korean Association of Quality Assurance for Clinical Laboratories and the College of American Pathologists Survey Proficiency Testing [14].

#### Statistical analysis

A conventional Cox proportional hazards regression model was used in this study to develop the PCa risk prediction model. Model performance was evaluated based on the criteria of discrimination and model calibration.

To allow for sufficient evaluation of our model, the full dataset was split at a ratio of 8:2, the former to be used for model development (the development cohort) and the latter to be used for model validation (the validation cohort). 5-fold cross validation was performed to evaluate model performance. C-statistics of both the development and validation cohorts were derived to evaluate discrimination. The Nam-D'Agostino test was used to evaluate the calibration of the model in each cohort [15]. We selected variables based on previous well-known risk factors including age, smoking status, alcohol consumption, hypercholesterolemia, and family history of PCa. The potential risk factors included in our final multivariable models were age, smoking status, alcohol consumption, past medical history of dyslipidemia, family history of PCa, and PSA levels.

The estimating equation for the probability of developing PCa within  $t$  years is given as follows:

$$P(PCa) = 1 - h_0(t|t = T)^{\exp(\sum_{i=1}^I \hat{\beta}_i * x_i)}$$

where  $x_i (i = 1, \dots, I)$  refers to risk factor  $i$ ,  $\hat{\beta}_i (i = 1, \dots, I)$  refers to the estimated coefficients from the Cox proportional hazards models, and  $h_0(t|t = T)$  refers to the baseline survival estimate at time  $t$ . Given

**Table 1** Baseline characteristics of study participants and incidences of prostate cancer (full cohort)

Risk factor	Frequency	
	Participants at baseline (%)	Incidences (%)
<b>Age (yrs)</b>		
40–50	47,731 (68.86)	51 (25.37)
50–60	15,670 (22.61)	73 (36.32)
60–70	5,918 (8.54)	77 (38.31)
<b>Smoking status</b>		
Non-smoker	9,964 (15.42)	32 (17.58)
Former smoker	29,790 (46.09)	99 (54.4)
Current smoker	24,875 (38.49)	51 (28.02)
<b>Alcohol consumption (g/day)</b>		
< 20	37,821 (57.15)	104 (58.43)
≥ 20	28,358 (42.85)	74 (41.57)
<b>The presence of hypercholesterolemia (mg/dl)</b>		
total cholesterol < 200	31,110 (44.88)	68 (33.83)
≥ 200 or on medication for hyperlipidemia	38,208 (55.12)	133 (66.17)
<b>Family history of PCa</b>		
No	64,995 (93.85)	176 (87.56)
Yes	4,259 (6.15)	25 (12.44)
<b>PSA (ng/mL)</b>		
< 3.0	67,464 (97.32)	117 (58.21)
3.0–6.9	1,678 (2.42)	69 (34.33)
7.0–10.0	177 (0.26)	15 (7.46)

Abbreviations: PCa, prostate cancer; PSA, prostate specific antigen

the time spread of the data, we estimated the risk of developing PCa in five years ( $T=5$ ).

The model was validated based on general indices of discrimination and calibration. Harell's C-statistics were calculated to assess the models' accuracy in ranking those with higher likelihood of developing PCa. The receiver operating characteristic (ROC) curves provided intuitive insight into the predictive ability of the two models based on the occurrence distribution at  $T=5$ . The calibration was measured using the Nam-D'Agostino test, executed as follows: ten subgroups were generated based on the estimated risk of PCa, and the proportion of observed occurrence and the average estimated predictions were compared using a  $\chi^2$  distribution. Hence, the Nam-D'Agostino test statistic allows the assessment of whether predicted values adhere to the observed occurrence data. All analyses were completed using STATA version 17.0 (StataCorp, College Station, TX, USA).

## Results

### Baseline characteristics of study participants

A predominant proportion of the study cohort received their first medical examination at the Kangbuk Samsung Hospital Total Healthcare Centers in their forties. The baseline visits of 15,670 and 5,918 participants were in their fifties and sixties. The incidence of PCa between the age groups showed a sharp increase with age. That is, the incidence per 100,000 person-years was approximately 16 times higher in the 60–70 age group than in the 40–50 age group, totaling an average incidence of approximately 45 per 100,000 person-years. Furthermore, approximately 48% (96 incidences) of PCa diagnoses were made within five years of the initial health examination.

### Evaluation of risk factors for prostate cancer

We showed the descriptive statistics of variables and incidence of PCa at the participants' first visit in Table 1. We showed the incidence of PCa according to age, the presence of hyperlipidemia, family history of PCa, and PSA levels among the full study cohort. Utilizing these risk factors, we constructed a multivariable Cox proportional hazards regression model using our development cohort. As shown in Table 2, it is evident that PSA is a significant and substantial risk factor associated with increasing hazard ratios for increasing PSA levels. When PSA level increased 1 ng/dL, the HR ratio for PCa increased by 1.774 (95% CI: [1.673–1.881]). Age, inclusion in a hyperlipidemia risk patient group, and family history of PCa, and PSA were statistically significant ( $p < 0.05$ ) risk factors of PCa. Although smoking status and alcohol consumption were not strong risk factors, we included those factors in multivariable analyses since they were assessed to contribute to a better global model fit.

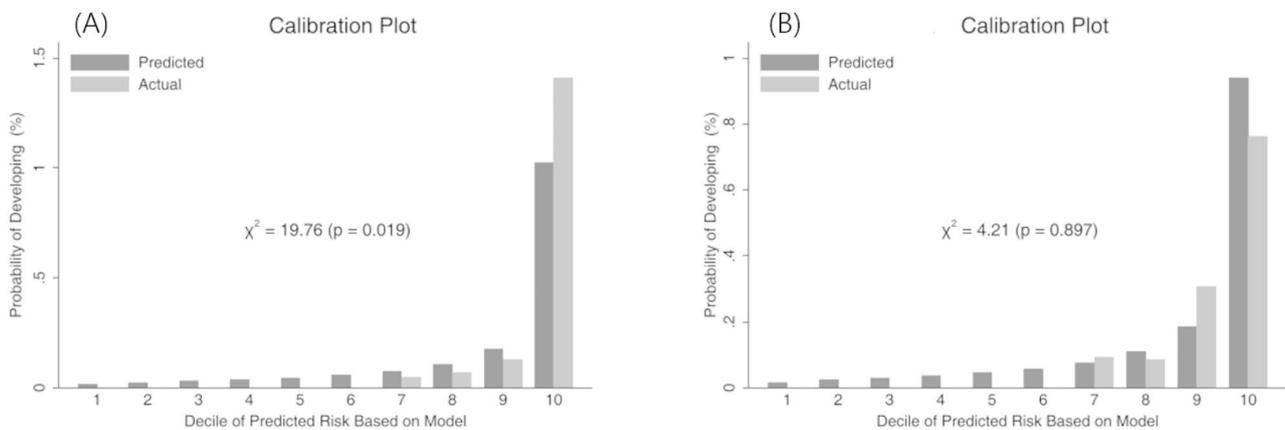
**Table 2** Multivariable regression analysis (development cohort)

Risk factor	HR	[95% CI]	p-value
<b>(Age-40)</b>	1.126	[1.102,1.150]	< 0.001
<b>Smoking status</b>			
Non-smoker	1	-	-
Former smoker	1.147	[0.691,1.901]	0.596
Current smoker	1.143	[0.652,2.00]	0.642
<b>Alcohol consumption (g/day)</b>			
< 20	1	-	-
≥ 20	1.158	[0.821,1.631]	0.403
<b>The presence of hypercholesterolemia (mg/dl)</b>			
total cholesterol < 200	1	-	-
≥ 200 or on medication for hyperlipidemia	1.802	[1.257,2.585]	0.001
<b>Family history of PCa</b>			
No	1	-	-
Yes	1.648	[0.944, 2.878]	0.079
<b>PSA (ng/mL)</b>	1.774	[1.673, 1.881]	< 0.001

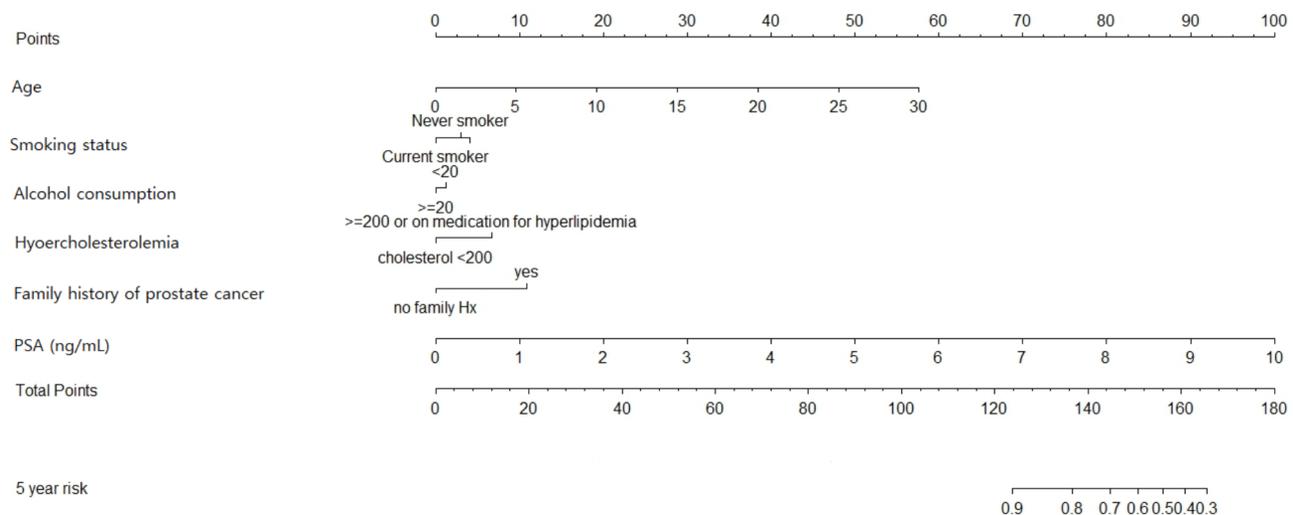
Abbreviations: HR, hazard ratio; CI, confidence interval; PCa, prostate cancer; PSA, prostate specific antigen

**Model performance and comparisons**

The general indices of calibration and discrimination were investigated to confirm the validity of the model. The C-statistics for the development and validation cohorts were calculated: 0.922, 0.874 for the development cohort and validation cohort, respectively. From these results, we concluded that the ability of the risk prediction model to accurately rank higher-and-lower risk groups was sufficient. The calibration plot is found in Fig. 2. The Nam-D’Agostino test statistic was 19.76 ( $p=0.019$ ) and 4.21 ( $p=0.897$ ) for the development and validation cohorts. These results indicate that our model showed satisfactory calibration with the real PCa occurrence data. Also, we added the nomogram for calculating the probability of developing PCa in Fig. 3.



**Fig. 2** Discrimination and calibration plots of the cancer risk prediction model of the validation cohort. **A**, deciles of predicted and real incidence rates in the validation cohort, **B** deciles of predicted and real incidence rates in the validation cohort. Nam-D’Agostino test statistics are displayed in plots **A** and **B**



**Fig. 3** Nomogram for calculating the probability of developing PCa

## Discussion

The present study suggests an effective PCa prediction model that could be practically used in primary care or a community health setting. Our model had sufficient discrimination ability in both development and validation cohorts, suggesting high efficiency in PCa prediction. In addition, our risk prediction model can provide useful information for clinical decisions on whether further evaluation for PCa should be performed when PSA levels are inconclusive.

Most previous prediction models for PCa have focused on the measures of PSA (freePSA, %fPSA, PSAD, p2PSA etc.) that maximized the accuracy of PCa detection [16–22]. These studies achieved high predictive accuracy and area under the curve (AUC) measures. However, the applicability of these models in a clinical setting is limited because they often focused on narrow PSA ranges (2.5–4.0 ng/mL) or on biomarkers less available in primary care or health examinations. On the other hand, few studies on PCa risk prediction investigated individual risk factors associated with PCa carcinogenesis [9, 10, 23]. However, considering that serum PSA is a commonly performed screening test in health examinations, a model consisting of only questionnaire-based risk factors is limited in providing additional information for people with marginal PSA levels.

This study proposed a PCa risk prediction model for clinical implementation by combining PSA levels and PCa risk factors. The efficacy of our model was evaluated in terms of both discrimination and calibration, and we established that our model was sufficient in producing statistically meaningful inferences. From the multivariable model, when PSA level increased 1ng/mL, the risk for PCa development increased 1.7 times higher after adjusting for other risk factors. Combined with other PCa risk factors, our model can provide useful information for clinical decisions on whether further evaluation for PCa should be performed when PSA levels are inconclusive. Moreover, since our model provides estimates based on predictions of 5-year risk of PCa, it may aid in the process of classifying individuals that require further close observation.

We performed the validity of several conventionally contested PCa risk factors including age, smoking status, alcohol consumption, past medical history of dyslipidemia, cholesterol levels, and family history. Aging is a main risk factor for PCa [24]. In our study, we found that a one-year increment in age after 40 years is associated with an approximately 20% increased risk of PCa. We also included age squared with the intent of capturing a quadratic increase in PCa risk with age but found little evidence to substantiate this hypothesis. Smoking status and alcohol consumption are risk factors often investigated for the risk of PCa [25]. Although those factors

were not statistically significant in multivariable analysis, they were well-known carcinogen and included in our prediction model.

A family history of PCa appeared to have a significant relationship with PCa carcinogenesis. These findings were consistent with those reported by Lesko et al. (1996) and support the literature on the genetic predispositions of PCa (e.g. BRCA1, BRCA2 mutations) [26]. Finally, our model suggested the presence of hypercholesterolemia as a risk factor, defined by an assessment of total cholesterol ( $\geq 200$  mg/dL) or dyslipidemia medication administration. It was associated with an approximately 67% increase in PCa risk, even after controlling for other factors. Previous studies have reported that elevated cholesterol levels could affect cell proliferation, inflammation, and lipid accumulation in the prostate [27]. The rapid growth of cancer cells might require a sufficient amount of cholesterol [27]. A recent study showed that androgen-independent PCa cell growth could be influenced by extracellular lipid levels and low-density lipoprotein-cholesterol availability [28]. Since many cancers develop as a consequence of chronic inflammation, persistent inflammation induced by cholesterol may increase the risk of PCa.

Our study had several limitations. First, the incidence of PCa was relatively lower than that of the National Cancer Statistics in Korea and it was not associated with several individual risk factors debated in previous studies. Direct comparisons with previous studies are unwarranted, however, because our study participants were younger than the general Korean population. Further, regarding our study cohort, it should be recognized that the participants at the Kangbuk Samsung Total Healthcare Centers tend to be relatively healthier and have higher socioeconomic statuses than the general Korean population [29]. Since the association between socioeconomic status on the prognosis of PCa is notable from previous studies, this aspect may limit the generalizability of this study [16]. Second, the follow-up period may be not sufficient to observe the development of PCa in this study. Hence, the risk of several factors could have been underestimated. Finally, it should be noted that lifestyle attributes, such as smoking status, alcohol consumption, and family history of PCa, were extracted from a self-reported questionnaire. Resulting recall biases may lead to an underestimation of the magnitude of these risk factors. However, as evidence to negate this claim, we found that the association between cotinine levels and smoking status was significant in our data, and further refer to Kerber and Slattery (1997) to suggest that recall bias for PCa is relatively small [30].

Despite the limitations above, we propose that the rich quality and quantity of our data, and the accuracy of the PCa registry data from the National Cancer Center are

features of our study that bolster the credibility of our results. Moreover, a prediction model developed in our study are sufficiently effective, suggesting that the potential risk factors that were included in the models are collectively strong predictors of PCa risk. We observed that individuals with increased PSA levels were associated with a significant increase in PCa risk within 5 years. Also, our prediction model can provide useful clinical information to classify high risk population with inconclusive PSA levels.

## Conclusion

Our prediction model can provide useful clinical information to classify high risk population with inconclusive PSA levels. These results are helpful in a clinical setting where PSA screening is less costly and hence frequently practiced. Furthermore, the development of advanced PCa risk prediction models could provide useful information to discriminate high-risk groups of PCa.

## Abbreviations

CI	Confidence interval
HR	Hazard ratio
ICD-10	International Classification of Diseases and Related Health Problems 10th Revision
PCa	Prostate cancer
PSA	Prostate-specific antigen
USPSTF	US Preventive Services Task Force

## Acknowledgements

The authors thank Sunwoo Cho for advice on study design and statistical analysis.

## Authors' contributions

JA Lee and EJ Kim conceptualized and designed this work. T Hwang and JA Lee conducted all analyses. H Oh, JA Lee, and EJ Kim provided clinical expertise and aided in interpreting the results. T Hwang and JA Lee drafted the manuscript, figure, and tables. All authors reviewed and edited the final manuscript.

## Funding

This work was supported by the Medical Research Funds from Kangbuk Samsung Hospital (Grant no. I20016, 2021).

## Data Availability

The data are not publicly shared because we do not have permission from the Institutional Review Board to distribute the data. The analytic methods are available from the corresponding authors upon reasonable request.

## Declarations

### Ethical approval

This study complied with the Declaration of Helsinki and was approved by the Institutional Review Board of Kangbuk Samsung Hospital, which waived the requirement for informed consent because of the use of anonymized data routinely collected as part of a health checkup program linked to mortality data from the Korea National Statistical Office (IRB No. 2011-01-030-005 for the general Kangbuk Samsung Health Study protocol and 2021-08-046 for the present study).

### Consent for publication

Not applicable.

### Competing interest

The authors declare that they have no conflict of interests.

Received: 6 October 2022 / Accepted: 24 April 2023

Published online: 03 June 2023

## References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and Mortality Worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209–49.
- Ahorsu DK, Lin C-Y, Imani V, Saffari M, Griffiths MD, Pakpour AHJljomh, et al. The fear of COVID-19 scale: development and initial validation. 2020:1–9.
- Zanini S, Renzi S, Limongi AR, Bellavite P, Giovinozzo F, Bermanno G. A review of lifestyle and environment risk factors for pancreatic cancer. *Eur J Cancer.* 2021;145:53–70.
- Pezaro C, Woo HH, Davis ID. Prostate cancer: measuring PSA. *Intern Med J.* 2014;44(5):433–40.
- Sindhvani P, Wilson CM. Prostatitis and serum prostate-specific antigen. *Curr Urol Rep.* 2005;6(4):307–12.
- Optenberg SA, Clark JY, Brawer MK, Thompson IM, Stein CR, Friedrichs P. Development of a decision-making tool to predict risk of prostate cancer: the Cancer of the prostate risk index (CAPRI) test. *Urology.* 1997;50(5):665–72.
- Lazzeri M, Haese A, de la Taille A, Palou Redorta J, McNicholas T, Lughezzani G, et al. Serum isoform [-2]proPSA derivatives significantly improve prediction of prostate cancer at initial biopsy in a total PSA range of 2–10 ng/ml: a multicentric european study. *Eur Urol.* 2013;63(6):986–94.
- Aladwani M, Lophatananon A, Ollier W, Muir K. Prediction models for prostate cancer to be used in the primary care setting: a systematic review. *BMJ Open.* 2020;10(7):e034661.
- Kim SH, Kim S, Jung JY, Kwon WA, Seo HK, Chung J, et al. Lifestyle risk prediction model for prostate Cancer in a Korean Population. *Cancer Res Treat.* 2018;50(4):1194–202.
- Yeo Y, Shin DW, Lee J, Han K, Park SH, Jeon KH, et al. Personalized 5-Year prostate Cancer Risk Prediction Model in Korea based on Nationwide Representative Data. *Journal of Personalized Medicine.* 2022;12(1):2.
- Chang Y, Jung H-S, Cho J, Zhang Y, Yun KE, Lazo M, et al. Metabolically healthy obesity and the development of nonalcoholic fatty liver disease. *Official journal of the American College of Gastroenterology | ACG.* 2016;111(8):1133–40.
- Grossman DC, Curry SJ, Owens DK, Bibbins-Domingo K, Caughey AB, Davidson KW, et al. Screening for prostate Cancer: US Preventive Services Task Force Recommendation Statement. *Jama.* 2018;319(18):1901–13.
- Huncharek M, Haddock KS, Reid R, Kupelnick B. Smoking as a risk factor for prostate cancer: a meta-analysis of 24 prospective cohort studies. *Am J Public Health.* 2010;100(4):693–701.
- Chang Y, Ryu S, Choi Y, Zhang Y, Cho J, Kwon MJ, et al. Metabolically healthy obesity and development of chronic kidney disease: a Cohort Study. *Ann Intern Med.* 2016;164(5):305–12.
- D'Agostino RB, Nam B-H. Evaluation of the performance of survival analysis models: discrimination and calibration measures. *Handbook of statistics.* 2003;23:1–25.
- Cheng I, Witte JS, McClure LA, Shema SJ, Cockburn MG, John EM, et al. Socioeconomic status and prostate cancer incidence and mortality rates among the diverse population of California. *Cancer Causes Control.* 2009;20(8):1431–40.
- Mearini L, Ferri C, Lazzeri M, Bini V, Nunzi E, Fiorini D, et al. Evaluation of prostate-specific antigen isoform p2PSA and its derivatives, %p2PSA, prostate health index and prostate dimension-adjusted related index in the detection of prostate cancer at first biopsy: an exploratory, prospective study. *Urol Int.* 2014;93(2):135–45.
- Fossati N, Buffi NM, Haese A, Stephan C, Larcher A, McNicholas T, 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: results from a multicentric european prospective study. *Eur Urol.* 2015;68(1):132–8.
- Abrate A, Lughezzani G, Gadda GM, Lista G, Kinzikeeva E, Fossati N, et al. Clinical use of [-2]proPSA (p2PSA) and its derivatives (%p2PSA and prostate Health Index) for the detection of prostate cancer: a review of the literature. *Korean J Urol.* 2014;55(7):436–45.
- Porcaro AB, Caruso B, Terrin A, De Luyk N, Cacciamani G, Corsi P, et al. The preoperative serum ratio of total prostate specific antigen (PSA) to

- free testosterone (FT), PSA/FT index ratio, and prostate cancer. Results in 220 patients undergoing radical prostatectomy. *Arch Ital Urol Androl.* 2016;88(1):17–22.
21. Chen R, Xie L, Cai X, Huang Y, Zhou L, Ma L, et al. Percent free prostate-specific antigen for prostate cancer diagnosis in chinese men with a PSA of 4.0–10.0 ng/mL: results from the chinese prostate Cancer Consortium. *Asian J Urol.* 2015;2(2):107–13.
  22. Boegemann M, Stephan C, Cammann H, Vincendeau S, Houlgatte A, Jung K, et al. The percentage of prostate-specific antigen (PSA) isoform [-2]proPSA and the prostate Health Index improve the diagnostic accuracy for clinically relevant prostate cancer at initial and repeat biopsy compared with total PSA and percentage free PSA in men aged  $\leq 65$  years. *BJU Int.* 2016;117(1):72–9.
  23. Park SY, Murphy SP, Wilkens LR, Henderson BE, Kolonel LN. Fat and meat intake and prostate cancer risk: the multiethnic cohort study. *Int J Cancer.* 2007;121(6):1339–45.
  24. Leitzmann MF, Rohrmann S. Risk factors for the onset of prostatic cancer: age, location, and behavioral correlates. *Clinical epidemiology.* 2012;4:1.
  25. Bostwick DG, Burke HB, Djakiew D, Euling S, Ho SM, Landolph J, et al. Human prostate cancer risk factors. *Cancer.* 2004;101(10 Suppl):2371–490.
  26. Lesko SM, Rosenberg L, Shapiro S. Family history and prostate cancer risk. *Am J Epidemiol.* 1996;144(11):1041–7.
  27. Pelton K, Freeman MR, Solomon KR. Cholesterol and prostate cancer. *Current opinion in pharmacology.* 2012;12(6):751–9.
  28. Raftopoulos NL, Washaya TC, Niederprüm A, Egert A, Hakeem-Sanni MF, Varney B, et al. Prostate cancer cell proliferation is influenced by LDL-cholesterol availability and cholesteryl ester turnover. *Cancer Metab.* 2022;10(1):1.
  29. Hur HW, Ryu SY, Park J, Choi SW. Relationship between socioeconomic status and prevalent prostate Cancer in the South Korea. *Asian Pac J Cancer Prev.* 2019;20(10):3137–44.
  30. Kerber RA, Slattery ML. Comparison of self-reported and database-linked family history of cancer data in a case-control study. *Am J Epidemiol.* 1997;146(3):244–8.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.