Original Article

Lifestyle Risk Prediction Model for Prostate Cancer in a Korean Population

Sung Han Kim, MD, PhD[1,2]
Sohee Kim, PhD[3]
Jae Young Joung, MD, PhD[1]
Whi-An Kwon, MD, PhD[4]
Ho Kyung Seo, MD, PhD[1,5]
Jinsoo Chung, MD, PhD[1]
Byung-Ho Nam, PhD[6]
Kang Hyun Lee, MD, PhD[1]

[1]Center for Prostate Cancer, National Cancer Center, Goyang, [2]Translational Research Branch, Research Institute, National Cancer Center, Goyang, [3]Biometrics Branch, Research Institute, National Cancer Center, Goyang, [4]Department of Urology, Institute of Wonkwang Medical Science, Wonkwang University Sanbon Hospital, Wonkwang University School of Medicine, Gunpo, [5]Biomarker Branch, Research Institute, National Cancer Center, Goyang, [6]Department of Cancer Control and Policy, Graduate School of Cancer Science and Policy, National Cancer Center, Goyang, Korea

Correspondence: Byung-Ho Nam, PhD
Department of Cancer Control and Policy,
Graduate School of Cancer Science and Policy,
National Cancer Center, 323 Ilsan-ro,
Ilsandong-gu, Goyang 10408, Korea
Tel: 82-31-920-1706
Fax: 82-31-920-2799
E-mail: byunghonam@heringsglobal.com

Co-correspondence: Kang Hyun Lee, MD, PhD
Center for Prostate Cancer, National Cancer
Center, 323 Ilsan-ro, Ilsandong-gu,
Goyang 10408, Korea
Tel: 82-31-920-1501
Fax: 82-31-920-2799
E-mail: uroonco@ncc.re.kr

Received October 13, 2017
Accepted December 19, 2017
Published Online December 21, 2017

*Sung Han Kim and Sohee Kim contributed equally to this work.

Purpose
The use of prostate-specific antigen as a biomarker for prostate cancer (PC) has been controversial and is, therefore, not used by many countries in their national health screening programs. The biological characteristics of PC in East Asians including Koreans and Japanese are different from those in the Western populations. Potential lifestyle risk factors for PC were evaluated with the aim of developing a risk prediction model.

Materials and Methods
A total of 1,179,172 Korean men who were cancer free from 1996 to 1997, had taken a physical examination, and completed a lifestyle questionnaire, were enrolled in our study to predict their risk for PC for the next eight years, using the Cox proportional hazards model. The model's performance was evaluated using the C-statistic and Hosmer–Lemeshow type chi-square statistics.

Results
The risk prediction model studied age, height, body mass index, glucose levels, family history of cancer, the frequency of meat consumption, alcohol consumption, smoking status, and physical activity, which were all significant risk factors in a univariate analysis. The model performed very well (C statistic, 0.887; 95% confidence interval, 0.879 to 0.895) and estimated an elevated PC risk in patients who did not consume alcohol or smoke, compared to heavy alcohol consumers (hazard ratio [HR], 0.78) and current smokers (HR, 0.73) (p < 0.001).

Conclusion
This model can be used for identifying Korean and other East Asian men who are at a high risk for developing PC, as well as for cancer screening and developing preventive health strategies.

Key words
Prostate neoplasms, Life style, Decision aids, Forecasting

Introduction

Prostate cancer (PC) is one of the most common among cancers in worldwide, with rapidly increasing number of cases [1]. The incidence and prevalence of PC vary by race, ethnicity, and geography, the highest being in North America and the lowest in South Asia. Even though less prevalent than in the Western countries [2], incidences of PC in Korea are rapidly increasing (12.3% annual increase); second only to thyroid cancer [3]. Some of the reasons for this substantial increase in incidences of PC include better diagnosis by monitoring blood levels of prostate-specific antigen (PSA), an increase in average life expectancy, Western dietary lifestyle, nutrition, physical activity, environmental factors, and smoking [1,4].

Consistent with the diverse etiology of PC, different ethnicities, and geographic locations present PCs with different biological characteristics. Being African-American is one of the highest risk factors for an early and aggressive form of PC [5]. Korean and Japanese populations, on the other hand, develop a higher grade of PC, compared to the Caucasian population, even in men with low levels of PSA. This could be attributed to the late detection of the disease and differences specific to the Korean and Japanese population [6].

As mentioned earlier, the introduction of PSA marker has enabled clinicians to identify high-risk PC patients and provide an earlier diagnosis with better prognostic outcomes. However, the role of PSA-based screening in PC diagnosis is still controversial, since it has not led to any significant decrease in PC mortality [7]. The results of a PSA test cannot tell if PC is present, although the risk of high-grade disease increases with increasing levels of PSA [8]. In Korea and other Asian countries, with no approved guidelines, PC is not included in the national cancer screening program. There is a clear need for the early detection of PC in Korean men due to the pathophysiologically aggressive nature of the Korean PC and the increased prevalence of PC in older men. A risk prediction model is a simple and effective way to evaluate individual cancer risks and to provide information regarding high-risk cancer.

PC is a good candidate that could use a lifetime risk prediction model to help early detection of PC, manage better prognostic outcomes, and provide important information for the development of strategic healthcare policies. In the last 20 years, several predictive tools have been developed for these purposes [5,9]. Most prediction models include PSA-derived variables, digital rectal examination (DRE) findings, and transrectal ultrasound findings, including prostate volume as a predictive variable. However, only a few studies have established risk prediction models for PC using epidemiological risk factors other than age. Therefore, this study was aimed to evaluate the lifetime epidemiologic risk factors for developing an individualized risk prediction model, independent of PSA levels, using a large Korean population-based cohort who provided a complete set of lifestyle information including social activities such as smoking, exercise, and alcohol consumption.

Materials and Methods

1. Patients population

Two independent population datasets and one dataset for PC incidence were incorporated into this study. The first dataset, which was used for model development, was from the database of the National Health Insurance Corporation (NHIC) for 1996 and 1997. The NHIC cohort consisted of Korean government employees, teachers, company employees, and their dependents, who underwent a biennial medical examination provided by NHIC. During the health examinations, participants were asked to fill out self-reported questionnaires on family history of any type of cancer, meal regularity, the frequency of meat consumption, alcohol drinking, smoking status, and physical activity. Height and weight were directly measured. Blood and urine laboratory test results were obtained, including fasting glucose levels. In the study, age, height, body mass index (BMI), glucose, family history of any cancer, the frequency of meat consumption, alcohol consumption, smoking status, and physical activity were selected for consideration in the analyses.

The second dataset, which was used for model validation, consisted of patients who participated in medical examinations in 1998 and 1999, and who were not included in the development model, using the same exclusion criteria. Details of the study design have been described in previous studies of lung, colorectal, and gastric cancers using NHIC dataset [9,10]. The last data set for PC incidence was obtained from the Korean Central Cancer Registry database up to December 31, 2007. Based on the International Classification of Disease 10th edition (ICD-10), C61 was used for PC incidence. The exclusion criteria were age less than 30 years or over 80 years, previous cancer history including alcohol- and smoking-related cancers, PC diagnosis within 2 years of baseline examination, and absence or inaccurate information regarding the analytical variables used in this study such as anthropometric parameters, family history of any type of cancer, meal regularity, frequency of meat consumption, alcohol drinking, smoking status, and physical activity.

From the development dataset, 3,482,255 men were iden-

tified who satisfied the inclusion criteria of which 825,320 men had complete information on the risk factors considered in the analyses. Because of the high rate of missing data, we complemented the data using the nearest observations imputation method, which our institution had used in the previous studies of lung and colorectal cancer [6,7]. We were able to retrieve some information from the NHIC examination data since these examinations were provided every 2 years. When the participant received examinations other than in 1996 and 1997, the data of the nearest time point was used to impute the missing values. After imputation, data from 1,179,172 men (33.9%) were available for model development. The difference in the prediction models developed based using complete data vs. imputed data was minor (S1 Fig.). The imputed, larger data set was then used for model development and validation. Similar missing data imputation was performed for the validation cohort according to the way used in the previous studies [6,7], and a total of 389,538 men were included in the validation cohort.

2. Statistical analysis

Two medical biostatisticians (S.K. and B.H.N.) performed a crude and age-adjusted analysis for each potential risk factor, to identify significant risk factors for PC in our data. A Cox proportional hazards regression model was used for developing prediction equations in the development set. Log-log survival plots were used to examine proportionality in hazards. Only factors with p-values of < 0.10 from the univariate analysis were subsequently evaluated in the Cox proportional hazards regression analysis using backward stepwise selection with a 0.10 significance level. Time to event was defined from the date of health examination at baseline to the date of first PC diagnosis. Subjects were censored at the date of death or on the end date after eight years of follow-up.

The potential risk factors considered in the analysis were age, age squared, height, BMI, glucose levels, family history of any type of cancer, meal regularity, the frequency of meat consumption, alcohol consumption, smoking status, and physical activity. All risk factors except age were included as categorical variables in the model. Further descriptions of the categorization rationale for these variables can also be found in previous studies [3,10].

The probability of developing PC within t years (t=8) for an individual with covariate values $x=(x_{1'}=ex_k)$ can be estimated using the following equation:

$$P(PC)=1-S_0(t)^{\exp[f(x,M)]}$$

, where $f(x,M)=\beta_1(x_1-M_1)+\beta_2(x_2-M_2)+\beta_k(x_k-M_k)$.

Here, β_1, re, β_k are the estimated coefficients from the Cox proportional hazards model, and M_1 and M_k are the mean values for each risk factor in the study population. $S_0(t)$ is the baseline survival estimate at time t (t=8 years), when all risk factors are at their mean values.

The developed models were validated by evaluating their performance in terms of discrimination and calibration. Harrell's C-statistics for survival data were measured for discrimination [11]. This value represents the odds of the predicted probability of developing PC being higher for those who actually develop PC in 8 years compared to those who do not develop the disease. An ROC curve was created based on the event distribution at time t=8.

Calibration is related to prediction accuracy. The Hosmer–Lemeshow type chi-square statistics was used for calibration [12]. To calculate the chi-square statistics, data were divided into 10 disjoint subgroups based on the predicted probabilities of developing PC from the developed model. The average predicted probabilities (expected) and the actual event rate measured by the Kaplan-Meier estimate (observed) were then compared. Values exceeding 20 indicated a significant lack of calibration [13]. All the analyses were completed using the SAS ver. 9.2 (SAS Institute Inc., Cary, NC) and the STATA ver. 13 (Stata Corp., College Station, TX) software by two medical biostatisticians (S.K and B.H.N).

3. Ethical statement

This study was approved by the Institutional Review Board of the National Cancer Center in Korea (IRB no. NCCNCS 09-305). Participants' informed consent was waived by the institutional review board because this study involved routinely collected medical data that were anonymously managed in all stages, including the stages of data cleaning and statistical analyses.

Results

1. Baseline characteristics and risk factors

During the 8-year follow-up, 2,747 (0.23%) and 846 (0.22%) patients developed PC in the development and validation cohorts, respectively. The mean±standard deviation age at diagnosis in the development cohort was 44.3±10.13 years.

In the univariate analyses after age-adjustment, the significant risk factors for PC were height, BMI, glucose levels, the presence of a family history of any cancer, meal regularity, the frequency of meat consumption, alcohol consumption, smoking status, and physical activity (p < 0.05) (Table 1).

Table 1. Risk factor distributions between cancer patients and cancer-free participants, and age-adjusted univariable in the developing cohort

Risk factor	Frequency		Age-adjusted univariable model	
	No. of participants at baseline (n=1,179,172)	No. of event (n=2,747)	HR (95% CI)	p-value
Age, mean±SD (yr)	44.3±10.13			
Height (cm)				
≤ 165	350,864 (29.8)	1,106 (40.3)	1.000	
165.1-168	225,760 (19.2)	545 (19.8)	1.157 (1.043-1.283)	0.006
168.1-172	325,984 (27.7)	669 (24.4)	1.297 (1.176-1.431)	< 0.001
> 172	273,817 (23.3)	427 (15.5)	1.315 (1.172-1.476)	< 0.001
BMI (kg/m²)				
< 18.5	27,521 (2.3)	63 (2.3)	0.675 (0.523-0.871)	0.003
18.5-22.9	482,264 (41.0)	1,039 (37.8)	1.000	
23.0-24.9	334,361 (28.4)	803 (29.2)	1.192 (1.087-1.307)	< 0.001
≥ 25.0	332,279 (28.2)	842 (30.7)	1.326 (1.210-1.453)	< 0.001
Glucose (mg/dL)				
< 126	1,108,417 (94.2)	2,549 (82.8)	1.000	
≥ 126	68,008 (5.8)	198 (7.2)	0.838 (0.725-0.968)	0.017
Family history of cancer				
No	993,152 (84.4)	2,272 (82.7)	1.000	
Yes	183,273 (15.6)	475 (17.3)	1.395 (1.263-1.54)	< 0.001
Meal regularity				
Regular	688,195 (58.5)	2,007 (73.1)	1.000	
Intermediate	388,450 (33.0)	622 (22.6)	0.841 (0.768-0.920)	< 0.001
Irregular	99,780 (8.5)	118 (4.3)	0.750 (0.623-0.904)	0.003
Frequency of meat consumption (per week)				
≤ 1 time	543,211 (46.2)	1,182 (43)	1.000	
2-3 times	572,245 (48.6)	1,336 (48.6)	1.192 (1.102-1.289)	< 0.001
≥ 4 times	60,969 (5.2)	229 (8.3)	1.085 (0.941-1.252)	0.261
Alcohol consumption (g/day)				
0	346,313 (29.4)	1,131 (41.2)	1.000	
1-14.9	341,906 (29.1)	756 (27.5)	1.027 (0.936-1.128)	0.569
15-24.9	209,677 (17.8)	390 (14.2)	0.988 (0.879-1.110)	0.836
≥ 25	278,529 (23.7)	470 (17.1)	0.784 (0.704-0.874)	< 0.001
Smoking status				
Never	345,848 (29.4)	1,130 (41.1)	1.000	
Former	166,960 (14.2)	535 (19.5)	0.973 (0.878-1.078)	0.600
Current	663,617 (56.4)	1,082 (39.4)	0.730 (0.671-0.794)	< 0.001
Physical activity				
None	560,188 (47.6)	1,192 (43.4)	1.000	
Light	188,988 (16.1)	410 (14.9)	1.182 (1.056-1.322)	0.004
Moderate	348,598 (29.6)	856 (31.2)	1.340 (1.227-1.464)	< 0.001
Heavy	78,651 (6.7)	289 (10.5)	1.212 (1.066-1.379)	0.004

Values are presented as number (%) unless otherwise indicated. HR, hazard ratio; CI, confidence interval; SD, standard deviation; BMI, body mass index.

Table 2. Multivariable regression model: risk prediction model

Risk factor	Mean values in development set	Multivariable model			
	M	β	HR	95% CI	p-value
(Age-Mean_age)	0	0.231	1.260	1.245-1.276	< 0.001
(Age-Mean_age)^2	102.5224	–0.004	0.996	0.996-0.997	< 0.001
Height (cm)					
≤ 165	0		1.000		
165.1-168	0.1919	0.081	1.084	0.978-1.202	0.126
168.1-172	0.2770	0.198	1.219	1.106-1.344	< 0.001
> 172	0.2326	0.240	1.272	1.135-1.426	< 0.001
BMI (kg/m²)					
< 18.5	0.0234	–0.260	0.771	0.598-0.996	0.046
18.5-22.9	0		1.000		
23.0-24.9	0.2842	0.090	1.094	0.997-1.200	0.059
≥ 25.0	0.2825	0.176	1.193	1.088-1.308	< 0.001
Glucose (mg/dL)					
< 126	0		1.000		
≥ 126	0.0578	–0.242	0.785	0.679-0.908	0.001
Family history of cancer					
No	0		1.000		
Yes	0.1558	0.286	1.331	1.205-1.470	< 0.001
Frequency of meat consumption (per week)					
≤ 1 time	0		1.000		
2-3 times	0.4864	0.181	1.199	1.107-1.297	< 0.001
≥ 4 times	0.0519	0.162	1.176	1.018-1.359	0.028
Alcohol consumption (g/day)					
0	0		1.000		
1-14.9	0.2906	–0.010	0.990	0.902-1.088	0.840
15-24.9	0.1781	–0.036	0.964	0.857-1.085	0.547
≥ 25	0.2366	–0.224	0.800	0.715-0.894	< 0.001
Smoking status					
Never	0		1.000		
Former	0.1420	–0.053	0.949	0.855-1.052	0.320
Current	0.5637	–0.257	0.774	0.710-0.843	< 0.001
Physical activity					
None	0		1.000		
Light	0.1606	0.069	1.071	0.957-1.200	0.233
Moderate	0.2964	0.180	1.198	1.095-1.310	< 0.001
Heavy	0.0669	0.129	1.138	1-1.294	0.051

HR, hazard ratio; CI, confidence interval; BMI, body mass index.

Being tall, greater BMI, lower glucose levels, presence of a familial cancer history, higher frequency of meat consumption, and heavier physical activity were poor risk factors with a hazard ratio (HR) greater than 1.0, whereas more irregular meal habit, lifestyle of heavy alcohol consumption > 25 g per day were favorable risk factors with a HR less than 1.0 (p < 0.05) (Table 1).

In the multivariate analyses for risk factors for PC, the sig-nificant risk factors with similar directional hazard ratios in the univariate analysis were still significant, except for meal regularity (p < 0.05) (Table 2). Based on the multivariate analysis, age, height, BMI, glucose levels, family history of any cancer, the frequency of meat consumption, alcohol con-sumption, smoking status, and physical activity were all included in the development of the risk prediction model.
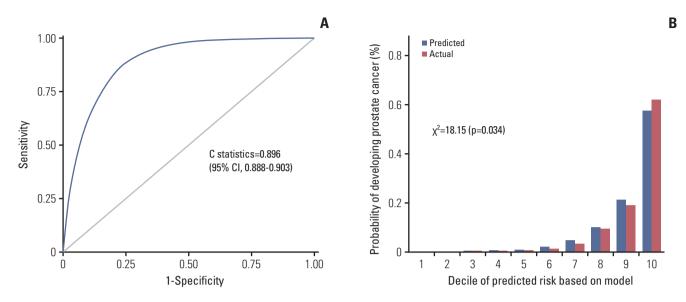
Fig. 1. Discrimination and calibration plots in the development cohort. (A) Discrimination. (B) Calibration. CI, confidence interval.
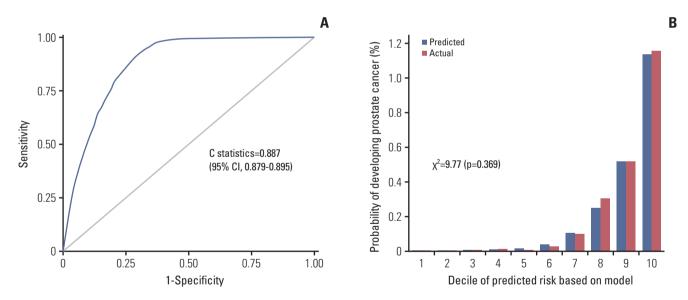

Fig. 2. Discrimination and calibration plots in the validation cohort. (A) Discrimination. (B) Calibration. CI, confidence interval.

2. Model performance

When calculating the probability (P) and estimating the baseline survival probability ($S_0(t)$) at the time (t)=8 years, for the mean values of the risk factors in the model, the $S_0(t)$ estimate at t=8, is 0.9998329. The ability of the risk prediction model to discriminate was measured using C-statistics in both the development and validation datasets and was found to be 0.896 (95% confidence interval [CI], 0.888 to 0.903) in the development dataset and 0.887 (95% CI, 0.879 to 0.895) in validation datasets. Figs. 1 and 2 show the calibration plots for the PC risk prediction model. The Hosmer–Lemeshow-type chi-square value was 18.15 for the development cohort and 9.77 for the validation cohort.

Discussion

With an increase in life expectancy and more people living well into their old age, comes the fear of being diagnosed with cancer later in life. This concern has driven the development of not only new diagnostic and screening tools but also models to predict the lifetime risk of developing cancer. In an effort to predict the risk of developing cancer in a healthy cohort and the risk of progressing to an aggressive disease in cancer patients, researchers have focused on the analysis of a wide spectrum of underlying factors ranging from genetics to lifestyle. This study evaluated several epidemiological risk factors suspected to have a role in predisposing men to PC and determined some significant variables to develop a risk prediction model of PC in Korean men.

The prevalence and outcomes of PC in Asians including the Koreans are different from those seen in the Western ethnic populations, suggesting that the epidemiological factors involved could also be different. This study is clinically important since it focuses on the risk factors for PC in the Korean population. The findings of this study will be helpful in deciding on future health policies and preventive strategies for PC in Korea. This study is the first to develop a risk prediction model of PC, based on lifestyle information. Our risk prediction model not only showed excellent discrimination in both the development dataset (C statistic, 0.896; 95% CI, 0.888 to 0.903) and the validation dataset (C statistic, 0.887; 95% CI, 0.879 to 0.895), but it also exhibited good calibration. Based on C statistics, our model is an excellent discriminator, compared with previous models, including those that combine PSA values with PSA derivatives and prostate volume.

Prediction models for PC incidence have been reported previously [14,15]. Most of them have focused on increasing the PSA test accuracy for PC detection and tumor stage prediction. The area under the curve (AUC) of PSA testing for predicting any PC has ranged from 0.53 to 0.83. High predictive accuracy and discrimination is achieved with some prediction models, such as Finne (AUC, 0.74), Karakiewcz (AUC, 0.74), Chun (AUC, 0.76), ERSPC RC3 (AUC, 0.79), and Prostaclass I (AUC, 0.79) [16]. However, due to no clear cutoff values associated with high specificity and sensitivity, the PSA test has limited value, often leading to over diagnosis and overtreatment [17]. Moreover, due to the absence of the PSA screening program in most Asian countries, multiple other additive parameters such as free PSA, DRE, and prostate volume were added to improve the predictive accuracy of PSA testing in the developmental prediction model.

This study did not utilize PSA levels, instead focused on the epidemiological lifestyle factors for assessing an individual's lifetime risk for developing PC. Some previous studies have dealt with a few epidemiologic factors, such as age, ethnicity, or family history of cancer in their prediction models. Lifestyle factors such as obesity, height, weight, BMI, and lean body mass, have also been evaluated but were found to have limited effects on risk for PC [10,18]. In addition to confirming these risk factors (Figs. 1 and 2), this large national case-control study found that age, height, higher BMI, lower glucose levels, presence of family history of any cancer, regularity of meals, higher meat consumption, and intense physical activity are all significant factors (p < 0.05) (Table 2), that help predict lifetime risk of PC with higher accuracy. Findings based on these factors were inconsistent in the previous studies [18,19].

One of the perplexing results of this study is the inverse association between smoking or alcohol consumption and increased PC risk, which is contrary to the commonly accepted risk factors for cancer development. This study showed that no alcohol and smoking were significantly associated with an elevated risk for PC risk, while heavy alcohol consumption (≥ 25 g per day; HR, 0.78; p < 0.001) and current smoking (HR, 0.73; p < 0.001), showed a poor association. Alcohol is an established risk factor for many cancers including PC and acts by altering circulating sex steroid hormone concentrations [20] and causing higher free radical generation [21]. Some meta-analyses have found modest risk increases associated with alcohol use [22,23]. Dennis [23] reported an insignificant overall pooled estimate (relative risk [RR], 1.05), and Bagnardi et al. [22] showed a RR of 1.19 (95% CI, 1.03 to 1.37) in his meta-analysis when comparing people who consumed 100 g alcohol per day to non-drinkers. This study stratified alcohol consumption into four groups with the highest being 25 g per day. However, studies of a relation between PC and alcohol consumption have yielded inconsistent results, with most studies showing no association. Rohrmann et al. [24] reported no association between alcohol consumption and PC in a cohort of 142,607 European men. Overall, neither alcohol consumption at baseline nor average lifetime alcohol consumption was associated with the risk for PC in the EPIC study, which, however, found a strong association between alcohol consumption and advanced or high-grade PC than with total PC risk. This is in agreement with previous studies [25].

There could be several reasons for the inverse relation between heavy alcohol consumption and lifetime PC incidence. First, the heavy drinkers in this study cohort were mostly young men (mean age, 44.3 years; 20-year follow-up study) with a lower incidence of PC. Individuals in their late fifties, on the other hand, consumed less than 15 g per day (not shown in tables) of alcohol and were more regular with their medical check-ups, which would account for the increase in the number of PC diagnoses among this group. As most PC cases develop in older individuals, cessation of

smoking and reduction of alcohol consumption to less than moderate levels may be expecting behaviors in those who are deeply concerned about their health; these individuals live relatively longer than others and have greater possibility of developing PC. Second, most patients with a past history of alcohol-related cancers, such as liver, esophageal, and stomach cancers, were excluded from this study cohort. Lastly, since heavy drinkers are likely to have other moderate to severe health problems and lower PSA levels, they may have lower chances of being diagnosed with PC [19]. Hence, there may be many individuals who die from other alcohol-related problems before the onset of PC. In addition, some of the individuals who were diagnosed with a chronic disease, such as hypertension, hypercholesterolemia, or diabetes, usually discontinue smoking or consuming alcohol, even if they were heavy drinkers or smokers previously. Due to the short latency period between risk measure and incidence of PC identification, these individuals can be classified as ex-smokers or non-drinkers (less drinkers) of which a great portion of heavy smokers or drinkers seemed to be excluded in the heavy smoker and drinker group to result in better than expected survival outcomes.

Smoking is also believed to increase the risk of many cancers. However, many studies report no association between smoking and PC incidence, and some European cohort studies have shown that smoking results in a small but significant reduction in the risk for PC risk, consistent with our results (Table 2) [26]. The European cohort study also showed an inverse association between smoking developing aggressive cancer, whereas other studies have shown no definite association between these two factors [27]. Another Asian cohort study that followed 14,450 males for 1 year, showed the age-adjusted relative risks of past and current smokers at entry to be 0.60 (95% CI, 0.34 to 1.06) and 0.70 (95% CI, 0.43 to 1.13), respectively. These findings are consistent with our study (Table 1), suggesting that cigarette smoking may not be a risk factor for PC [26]. Paradoxically, smoking, in particular, heavy smoking, has been associated with a significant increase in the risk of death from PC [28]. Another meta-analysis has shown a statistically significant and consistent increase in PC incidence and risk of death from this disease with increased smoking [4]. In this study, we found that men who never smoked had an elevated PC risk compared to current smokers (HR, 0.77; p < 0.001). As with alcohol consumption described earlier, patients with a high possibility of developing smoking-related upper digestive tract and respiratory tract cancers were excluded from this study cohort. Other current or former smokers in this cohort are more likely to die sooner from smoking relating diseases such as coronary heart disease, chronic obstructive pulmonary disease, nasopharyngeal cancer, bladder cancer and lung cancer before PC diagnosis [29].

This study has a few potential limitations such as the retrospective design and recall biases based on the questionnaires. Other limitations were a clear lack of information regarding baseline PSA and PC pathology, since all data were collected through routine physical examinations; the short periods between risk measure and incidence of PC identification; and the exclusion of additional unmeasured or unexamined variables. However, we controlled for numerous potential confounders relating to the development of PC, and none appear to have substantially affected our risk estimates. Lastly, the information on consumption patterns is difficult to validate. It is therefore not clear whether binge drinking increased the RR for PC, as was seen among the participants of the Health Professionals Follow-up Study [30]. Despite these limitations, this is the first significant study of clinical prediction modeling assessing the incidence risk of PC by routine lifestyle epidemiological and anthropometric parameters in Koreans. We expect this model to play an important role in improving decision-making and defining groups at high risk for PC and applying cancer prevention strategies in the field of health welfare policy in Asian countries including Korea, where the PSA screening program is not included in the routine nationwide health screening program.

Electronic Supplementary Material

Supplementary materials are available at Cancer Research and Treatment website (http://www.e-crt.org).

Conflicts of Interest

Conflict of interest relevant to this article was not reported.

Acknowledgments

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA Cancer J Clin. 2015;65: 87-108.

2. Park SK, Sakoda LC, Kang D, Chokkalingam AP, Lee E, Shin HR, et al. Rising prostate cancer rates in South Korea. Prostate. 2006;66:1285-91.

3. 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;100:693-701.

4. Thompson IM, Ankerst DP, Chi C, Goodman PJ, Tangen CM, Lucia MS, et al. Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. J Natl Cancer Inst. 2006;98:529-34.

5. Jeong IG, Dajani D, Verghese M, Hwang J, Cho YM, Hong JH, et al. Differences in the aggressiveness of prostate cancer among Korean, Caucasian, and African American men: a retrospective cohort study of radical prostatectomy. Urol Oncol. 2016;34:3.e9-14.

6. Andriole GL, Crawford ED, Grubb RL 3rd, Buys SS, Chia D, Church TR, et al. Prostate cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: mortality results after 13 years of follow-up. J Natl Cancer Inst. 2012;104:125-32.

7. Thompson IM, Pauler DK, Goodman PJ, Tangen CM, Lucia MS, Parnes HL, et al. Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter. N Engl J Med. 2004;350:2239-46.

8. Roobol MJ, Steyerberg EW, Kranse R, Wolters T, van den Bergh RC, Bangma CH, et al. A risk-based strategy improves prostate-specific antigen-driven detection of prostate cancer. Eur Urol. 2010;57:79-85.

9. Park S, Nam BH, Yang HR, Lee JA, Lim H, Han JT, et al. Individualized risk prediction model for lung cancer in Korean men. PLoS One. 2013;8:e54823.

10. Shin A, Joo J, Yang HR, Bak J, Park Y, Kim J, et al. Risk prediction model for colorectal cancer: National Health Insurance Corporation study, Korea. PLoS One. 2014;9:e88079.

11. Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA. 1982;247:2543-6.

12. D'Agostino RB, Nam BH. Evaluation of the performance of survival analysis models: discrimination and calibration measures. Handb Stat. 2003;23:1-25.

13. D'Agostino RB Sr, Grundy S, Sullivan LM, Wilson P; CHD Risk Prediction Group. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. JAMA. 2001;286:180-7.

14. Schroder F, Kattan MW. The comparability of models for predicting the risk of a positive prostate biopsy with prostate-specific antigen alone: a systematic review. Eur Urol. 2008;54: 274-90.

15. Shariat SF, Karakiewicz PI, Roehrborn CG, Kattan MW. An updated catalog of prostate cancer predictive tools. Cancer. 2008;113:3075-99.

16. 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.

17. Thompson IM, Ankerst DP, Chi C, Lucia MS, Goodman PJ, Crowley JJ, et al. Operating characteristics of prostate-specific antigen in men with an initial PSA level of 3.0 ng/ml or lower. JAMA. 2005;294:66-70.

18. Kolonel LN. Fat, meat, and prostate cancer. Epidemiol Rev. 2001;23:72-81.

19. Oesterle S, Hill KG, Hawkins JD, Guo J, Catalano RF, Abbott RD. Adolescent heavy episodic drinking trajectories and health in young adulthood. J Stud Alcohol. 2004;65:204-12.

20. Sierksma A, Sarkola T, Eriksson CJ, van der Gaag MS, Grobbee DE, Hendriks HF. Effect of moderate alcohol consumption on plasma dehydroepiandrosterone sulfate, testosterone, and estradiol levels in middle-aged men and postmenopausal women: a diet-controlled intervention study. Alcohol Clin Exp Res. 2004;28:780-5.

21. Poschl G, Seitz HK. Alcohol and cancer. Alcohol Alcohol. 2004;39:155-65.

22. Bagnardi V, Blangiardo M, La Vecchia C, Corrao G. A meta-analysis of alcohol drinking and cancer risk. Br J Cancer. 2001;85:1700-5.

23. Dennis LK. Meta-analysis for combining relative risks of alcohol consumption and prostate cancer. Prostate. 2000;42:56-66.

24. Rohrmann S, Linseisen J, Key TJ, Jensen MK, Overvad K, Johnsen NF, et al. Alcohol consumption and the risk for prostate cancer in the European Prospective Investigation into Cancer and Nutrition. Cancer Epidemiol Biomarkers Prev. 2008;17:1282-7.

25. Schoonen WM, Salinas CA, Kiemeney LA, Stanford JL. Alcohol consumption and risk of prostate cancer in middle-aged men. Int J Cancer. 2005;113:133-40.

26. Bae JM, Li ZM, Shin MH, Kim DH, Lee MS, Ahn YO. Cigarette smoking and prostate cancer risk: negative results of the Seoul Male Cancer Cohort Study. Asian Pac J Cancer Prev. 2013;14: 4667-9.

27. Sawada N, Inoue M, Iwasaki M, Sasazuki S, Yamaji T, Shimazu T, et al. Alcohol and smoking and subsequent risk of prostate cancer in Japanese men: the Japan Public Health Center-based prospective study. Int J Cancer. 2014;134:971-8.

28. Rohrmann S, Linseisen J, Allen N, Bueno-de-Mesquita HB, Johnsen NF, Tjonneland A, et al. Smoking and the risk of prostate cancer in the European Prospective Investigation into Cancer and Nutrition. Br J Cancer. 2013;108:708-14.

29. Iribarren C, Tekawa IS, Sidney S, Friedman GD. Effect of cigar smoking on the risk of cardiovascular disease, chronic obstructive pulmonary disease, and cancer in men. N Engl J Med. 1999;340:1773-80.

30. Platz EA, Leitzmann MF, Rimm EB, Willett WC, Giovannucci E. Alcohol intake, drinking patterns, and risk of prostate cancer in a large prospective cohort study. Am J Epidemiol. 2004;159:444-53.