Based on biomedical index data
Risk prediction model for prostate cancer
Hanxu Guo, BS, a, Xianjie Jia, MS, b, Hao Liu, PhD, c, ∗

Abstract
To explore the influencing factors of prostate cancer occurrence, set up risk prediction model, require reference for the preliminary diagnosis of clinical doctors, this model searched database through the data of prostate cancer patients and prostate hyperplasia patients National Clinical Medical Science Data Center.

With the help of Stata SE 12.0 and SPSS 25.0 software, the biases between groups were balanced by propensity score matching. Based on the matched data, the relevant factors were further screened by stepwise logistic regression analysis, the key variable and artificial neural network model are established. The prediction accuracy of the model is evaluated by combining the probability of test set with the area under receiver operating characteristic curve (ROC).

After 1:2 PSM, 339 pairs were matched successfully. There are 159 cases in testing groups and 407 cases in training groups. And the regression model was $P = 1 / (1 + e^{(0.122 \times \text{age} + 0.083 \times \text{Apo lipoprotein C3} + 0.371 \times \text{total prostate specific antigen (tPSA)} - 0.227 \times \text{Apo lipoprotein C2} - 6.093 \times \text{free calcium (iCa)} + 0.428 \times \text{Apo lipoprotein E} - 1.246 \times \text{triglyceride-1.919} \times \text{HDL cholesterol} + 0.083 \times \text{creatine kinase isoenzyme (CKMB)})})$. The logistic regression model performed very well (ROC, 0.963; 95% confidence interval, 0.951 to 0.978) and artificial neural network model (ROC, 0.983; 95% confidence interval, 0.964 to 0.997). High degree of Apo lipoprotein E (Apo E) (Odds Ratio, [OR], 1.535) in blood test is a risk factor and high triglyceride (TG) (OR, 0.288) is a protective factor.

It takes the biochemical examination of the case as variables to establish a risk prediction model, which can initially reflect the risk of prostate cancer and bring some references for diagnosis and treatment.

Abbreviations: Apo A1 = Apo lipoprotein A1, Apo A2 = Apo lipoprotein A2, Apo B2 = Apo lipoprotein B2, Apo C2 = Apo lipoprotein C2, Apo C3 = Apo lipoprotein C3, Apo E = Apo lipoprotein E, CKMB = creatine kinase isoenzyme, iCa = free calcium, TG = triglyceride, tPSA = total prostate specific antigen.

Keywords: logistic regression model, neural network model, propensity score matching, prostate cancer, risk factor

1. Introduction

1.1. Background
Prostate cancer (PCa) is a common malignant tumor of the genitourinary system in elderly men, and its incidence has obvious ethnic and regional differences. Scientific research shows that prostate cancer accounts for the first place in the total number of new cancers in the United States in 2019, and prostate cancer deaths account for 10% of all cancer deaths. Prostate cancer screening has always been a controversial health topic. According to the recommendations of the American Urological Association, prostate specific antigen (prostate specific antigen, PSA) screening was generally carried out in the mid-1990s. However, in recent years, many national guidelines have been updated and revised in accordance with the recommendations of U.S. Preventive Services Task Force (USPSTF), which recommended against PSA screening and pointed out that patients may be over diagnosed and over treated. Nowadays PSA has not shown the ability to discriminate clinically important cancers from low-risk tumors. The Prostate Cancer Intervention versus Observation Trial (PIVOT) trial showed no survival benefit from radical prostatectomy in men with PSA ≤ 10 μg/L, which are based on D’Amico criteria as a combination of PSA < 10 μg/L, stage ≤ T2A and Gleason score ≤ 6. It also was reported that more than half of these men underwent unnecessary treatments in Australia. Moreover, PSA screening as currently practiced in the United States provides little to no reduction in prostate cancer mortality or morbidity, does not decrease any-cause mortality, and results in substantial diagnostic and treatment harms and large health-care expenditures. The health importance of prostate cancer and the financial costs to society require improved detection and treatment strategies and more rational use of current options. Until then, men and their health-care providers can make a wise health-care choice by saying no to the PSA test.
incidence is increasing rapidly year by year. Some prediction model has been set up, Zhu et al has tested these models that was overestimated by approximately 20% for a wide range of predicted probabilities. Therefore, the construction of prostate cancer risk prediction model for Chinese is an effective and necessary method. There are also many reports about prostate cancer risk prediction in China, but the main factors considered are diagnosis and treatment methods, pathological grading, magnetic resonance imaging, fluorescent probe, etc.

1.2. Objectives
And most of the models do not take into account common clinical indicators such as biochemical examinations. But it is convenient for us to obtain these results from blood test, we can highly increase the efficiency of screening prostate cancer by this innovative model. It may be the first one to use biochemical index to build a prediction model for prostate cancer. In addition, the co-variables of most models will affect the accuracy of the model in varying degrees. Through the innovative method of propensity score matching, this scientific study is expected to better match the case group with the control group, so as to make the co-variables more balanced this model can be used for identifying Chinese who are at a high risk for developing prostate cancer, as well as for cancer screening and developing preventive health strategies.

Artificial neural network, which was first proposed by David and James of San Diego State University. It theoretically proves that any continuity function in a closed interval constantly adjusts the connection strength criterion between neurons in the training of the network. So that the difference between the calculated output dependent variable vector of the network and the dependent variable vector of the known training sample is minimum (that is, the prediction effect is the best).

2. Materials and methods

2.1. Source of data & sample size
The data of the study were collected from PSA and biochemical examination data of 3000 patients in the National Clinical Medical Science Data Center (301 Hospital), which including 2771 hyperplasia of prostate patients, 112 both prostate hyperplasia and prostate cancer patients and 117 prostate cancer patients. And 112 both prostate hyperplasia and prostate cancer patients recognized as prostate cancer patients with 117 prostate cancer patients are experiment group. 2771 hyperplasia of prostate patients are defined as control group.

2.2. Participants & missing data
With the help of Excel 2010 software to sort out the data, we selected a number of biomedical index variables with data, which the missing values are less than 5% of the original data, and this experiment uses SPSS version 25.0 (SPSS Inc. Chicago, IL) software to deal with the missing values of a small number of variables with MCAR method and extracts relevant variables from the diagnostic information of PSA and biochemical tests.

2.3. Predictors
The propensity score matching was achieved on Stata version 12.0 SE (Stata College Station, Texas 77845 USA). In the data after screening by SPSS software, 2771 non-prostate cancer patients were used as the control group, and 229 prostate cancer patients were entered into the Stata database as the experimental group. Based on whether the patient has prostate cancer as a grouping factor, the remaining variables [AGE WEIGHT Body Mass Index Apo lipoprotein A1 (Apo A1) Apo lipoprotein A2 (Apo A2) Apo lipoprotein B2 (Apo B2) Apo lipoprotein C2 (Apo C2) Apo lipoprotein C3 (Apo C3) Apo lipoprotein E (Apo E) albumin alkaline phosphatase lactate dehydrogenase creatine kinase creatine kinase isoenzyme (CKMB) triglyceride high-density lipoprotein cholesterol low-density lipoprotein cholesterol sodium calcium inorganic phosphorus free calcium (iCa) potassium chloride creatinine (Cre) total prostate specific antigen (tPSA) free prostate specific antigen] are included in the Ps model as covariates. The caliper value is set to 0.05 according to previous scientific research. In this example, it is set to 1:2 matching. After that, the results need to process the validation, which is called the equilibrium Test. If the bias were negative number, it is necessary to process the test again till the matched data fits well.

2.4. Ethical statement
This study was approved by the Institutional Review Board of the Chinese PLA General Hospital. 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.

2.5. Statistical analysis methods
2.5.1. Multivariable logistic regression analysis and artificial neural network. Study was used to set the non-prostate cancer patients in the matching database as the control group and prostate cancer patients as the case group. The case of prostate cancer was assigned a value of 1 and the case of no prostate cancer was assigned a value of 0. Based on the data, literature expert experience and clinical knowledge, the obtained patient information was screened: age, weight, body mass index, Apo A1, Apo A2, Apo B2, Apo C2, Apo C3, Apo E, serum albumin, alkaline phosphatase, lactate dehydrogenase, CKMB, triglyceride, high-density lipoprotein cholesterol and low-density lipoprotein cholesterol. These variables were compared and analyzed for the above-mentioned indicators of these 2 types of patients. Various factors related to prostate cancer were determined using univariable logistic regression analysis methods. Related factors (P<.05) were included in multivariable stepwise Logistic. Meanwhile, data with filtered variables were also included in artificial neural network for training and testing. After analysis, meaningful correlation factors were found, and the degree of impact on prostate cancer were based on significant differences.

3. Results
3.1. Participants
After PSM, the result of the first equilibrium test show that the fitting effects of the variables Na, ALP, and LDLC are not good. As shown in Table 1, the 3 variables are excluded for the second matching. The second equilibrium test shows that all variables have a good fit, such as in Supplementary Table 1, http://links.
In summary, the variables included in the model at this time are: age, weight, height, body mass index, Apo A1, Apo A2, Apo C3, Apo E, serum albumin, Cre, CKMB, triglyceride, high-density lipoprotein cholesterol, iCa, tPSA.

### 3.2. Model specification

Matched data were entered into univariable logistic regression to screen variables. And pick up which one has statistical significance ($P < .05$). The analysis results are shown in Supplementary Table 2, http://links.lww.com/MD/G38. As for PSA and calcium parameter, only one of each type has been selected. The above table shows that the total PSA is greater than the free PSA in the Wald test. And in previous studies, total PSA always reflects the serum antigen value, which has not been included in the model.

### 3.3. Model performance

The Logistic regression equation was obtained as follows: $P = 1 / (1 + e^{- (0.122 \cdot \text{age} + 0.083 \cdot \text{triglyceride} + 0.227 \cdot \text{iCa} + 0.428 \cdot \text{tPSA} + 0.179 \cdot \text{ALB} + 0.371 \cdot \text{Apo C3} + 0.327 \cdot \text{Apo E} + 0.35 \cdot \text{serum albumin} + 0.083 \cdot \text{BMI} + 0.083 \cdot \text{ip}\text{PSA} + 0.083 \cdot \text{Apopt lipoprotein C2} + 0.083 \cdot \text{Apo C3} + 0.083 \cdot \text{PSM})})$.

The new model was compared with the simulated probability of total diagnosis of prostate cancer only by PSA, and the diagnostic efficiency was judged according to the ROC curve (Fig. 1). It can be seen that the prediction efficiency of the new model (measure of area: 0.963) is significantly higher than that of only PSA (measure of area: 0.785) as a single factor.

Data with 9 related factors were resulted from univariable logistic regressions were put into artificial neural network model to simulate the diagnosis of prostate cancer only by PSA, and the performance of the new model was compared with the simulated probability of total diagnosis of prostate cancer only by PSA, and the diagnostic efficiency was judged according to the ROC curve (Fig. 1). It can be seen that the prediction efficiency of the new model (measure of area: 0.963) is significantly higher than that of only PSA (measure of area: 0.785) as a single factor.

At the same time, compare the logistic regression model with the artificial neural network model to simulate the diagnosis efficiency, and judge the diagnosis efficiency according to the

### Table 1

Basic demographics.

| Variable | Non-PCa n=458 | P value | Relative bias | Non-PCa n=458 | P value | Relative bias | Bias (%) |
|----------|---------------|---------|---------------|---------------|---------|---------------|----------|
| AGE      | 54.733        | .000    | 119.3         | 66.899        | .000    | 11.3          | 94.6     |
| WEIGHT   | 77.019        | .000    | 119.3         | 72.367        | .000    | 11.3          | 92.8     |
| BMI      | 25.882        | .000    | 119.3         | 29.6          | .000    | 11.3          | 99.5     |
| Apo C3   | 11.801        | .007    | 119.3         | 10.334        | .000    | 11.3          | 95.3     |
| Apo A2   | 27.925        | .000    | 119.3         | 26.674        | .000    | 11.3          | 92.8     |
| Apo C2   | 5.1609        | .000    | 119.3         | 4.0751        | .000    | 11.3          | 99.8     |
| Apo E    | 4.2864        | .000    | 119.3         | 4.9898        | .000    | 11.3          | 96.3     |
| ALB      | 43.535        | .000    | 119.3         | 41.68         | .000    | 11.3          | 89.4     |
| CKMB     | 12.48         | .000    | 119.3         | 16.548        | .000    | 11.3          | 87.3     |
| tPSA     | 0.4132        | .000    | 119.3         | 1.1049        | .000    | 11.3          | 87.9     |
| Ca       | 2.2898        | .000    | 119.3         | 2.2423        | .000    | 11.3          | 87.9     |
| CL       | 105.05        | .000    | 119.3         | 102.93        | .000    | 11.3          | 83.7     |
| iCa      | 1.1793        | .000    | 119.3         | 1.1378        | .000    | 11.3          | 81.5     |
| LDH      | 150.65        | .000    | 119.3         | 158.13        | .000    | 11.3          | 84.8     |
| CK       | 103.83        | .000    | 119.3         | 101.64        | .000    | 11.3          | 84.3     |
| Cre      | 80.962        | .000    | 119.3         | 85.771        | .000    | 11.3          | 84.3     |
| TG       | 1.9869        | .000    | 119.3         | 1.4111        | .000    | 11.3          | 90.0     |
| HDL-C    | 1.1229        | .000    | 119.3         | 1.2167        | .000    | 11.3          | 93.4     |
| Apo A1   | 1.1679        | .000    | 119.3         | 1.3451        | .000    | 11.3          | 93.4     |
| K        | 4.0508        | .000    | 119.3         | 4.0239        | .000    | 11.3          | 90.7     |
| Apo B    | 0.8912        | .000    | 119.3         | 0.99848       | .000    | 11.3          | 58.0     |
| Na       | 141.9         | .000    | 119.3         | 141.79        | .000    | 11.3          | 58.5     |
| ALP      | 68.444        | .000    | 119.3         | 66.53         | .000    | 11.3          | 46.8     |
| LDL-C    | 2.8569        | .000    | 119.3         | 2.919         | .000    | 11.3          | 45.5     |

Values are presented as mean. Apo C2 = Apo lipoprotein C2, Apo C3 = Apo lipoprotein C3, PSM = propensity score matching.
The diagnosis efficiency of neural network model (ROC, 0.963; 95% confidence interval, 0.951–0.978) is slightly higher than that of logistic regression model (ROC, 0.983; 95% confidence interval, 0.964–0.997) (Supplemental Table 3, http://links.lww.com/MD/G39).

### 4. Discussion

With more debate on the accuracy of the PSA screening on prostate cancer, a large number of researchers found that PSA screening may lead to overtreatments and overdiagnoses. This concern driven the process of not only new diagnostic and prognostic tools but also models to predict the risk of prostate cancer. Based on the fundamental realities of the country, this...
study innovatively evaluated several biomedical risk factors suspected to have a role in predisposing men to prostate cancer and determined some significant variables to develop a risk prediction model of prostate cancer in China.

The findings of this study will be helpful in deciding on future health policies and preventive strategies for prostate cancer in China. This study is the first to develop a risk prediction model of prostate cancer based on biomedical information. With the 96.3% diagnostic efficiency of logistic regression model and 98.3% that of neural network model, our model is an excellent discriminator, compared with former models, including those that combine PSA values with PSA relatives and prostate volume.

Former prediction models for prostate cancer have been reported.\cite{24,25} Most of previous models have concentrated on the PSA test for prostate cancer screening and ignore that the cut-off values of specificity and sensitivity are indistinct. High predictive accuracy and discrimination is completed in several models, for instance, Prostaclass I (AUC, 0.79), Chun (AUC, 0.76), Karakiewcz (AUC, 0.74), and Finne (AUC, 0.74).\cite{26} But they still limited in high probability of over diagnosis and overtreatment.\cite{27} In the case of the absence of PSA screening in China, multiple other additive parameters such as MRI, DRE, and prostate volume were added to increase the predictive accuracy of PSA testing in the developmental prediction model.

This study not only adds the PSA also include multiple biomedical parameters, which are easily obtained in Chinese blood test report. Relevant reports are few, but risk factors above are demonstrated in experiments that are explored by researchers.

4.1. Interpretation

4.1.1. Apo lipoprotein E, Apo lipoprotein C2, and Apo lipoprotein C3. Despite recognized risk factors such as age and PSA,\cite{28,29} which have also been demonstrated in the experiment, some risk factors that are still controversial also be found in this experiment. Apo lipoprotein E is also an important cholesterol regulatory protein. The main genetic subtypes of Apo lipoprotein E in the body are E3/E3, E3/E4, E2/E3, E2/E2, and E4/E4. The Apo lipoprotein E is also an important cholesterol regulatory protein. At present, studies at home and abroad have shown that the relationship between the invasion and Gleason score of prostate cancer cells and their genotypes in vivo is controversial. In an earlier study, Liu et al, through a case-control study,\cite{30} indicated that the E4 genotype and its allele were not associated with the pathogenesis and prognosis of prostate cancer, but could not explain the experiment conducted by Ifere et al\cite{31} to prove that Apo lipoprotein E2/E4 is a risk factor for prostate cancer. In recent years, Yencilek et al\cite{32} believe that the presence of E4 may reduce the possibility of prostate cancer, but still believe that E3/E3 is a major risk factor for prostate cancer and affecting Gleason scores. Recently, a research by Asare et al\cite{33} had proved Apo E could potentially be a discriminating biomarker for prostate cancer. Our study has supported this opinion.

Logistic stepwise analysis showed that Apo lipoprotein E was a risk factor for prostate cancer, with an OR of 1.535. From the side to verify the relationship between Apo lipoprotein E and prostate cancer, in clinical work, can guide patients to do some genetic tests, in order to better diagnose and guide the next treatment.

Besides, in the past, there may be not experiments that have stated Apo lipoprotein C2 and Apo lipoprotein C3 are associated with prostate cancer. This model probably the first 1 to demonstrated that Apo lipoprotein C2 is the protective factor and Apo lipoprotein C3 is the risk factor of prostate cancer, with OR values of 0.797 and 1.086, respectively. Its internal mechanism we guess may be that tumor patients accelerate the decomposition of endogenous lipids and the transformation and
oxidation of free fatty acids and glycerol due to the invasion of tumor tissue and the increase of the level of lipid metabolic hormones and insulin tolerance in the host.

4.1.2. Triglyceride and high density lipoprotein cholesterol.
TG provides essential fatty acids in lipid metabolism, which still remains controversy among scholars. Allot et al. believe that the increase of serum triglycerides is related to the occurrence of prostate cancer. However, Asare et al. have not found any significance difference with TG between Benign prostatic hyperplasia and prostate cancer. High density lipoprotein cholesterol (HDL-C) is an anti-atherosclerotic lipoprotein that transfers cholesterol from extra hepatic tissue to liver for metabolism. A case-control study conducted by Magura et al. showed that high TC (total cholesterol), high LDL-C (low density lipoprotein cholesterol), and low HDL-C may be risk factors for prostate cancer. However, the discussion on the relationship between blood lipids and prostate cancer is still controversial. In general, No experiments based on Chinese people have been created in order to study on the internal association between TG, HDL-C and prostate cancer.

Consistent with the results of this study, triglyceride and high density lipoprotein cholesterol were protective factors for prostate cancer, with OR values of 0.288 and 0.147, respectively, reflecting that low triglyceride and low high density lipoprotein cholesterol increase the risk of prostate cancer.

4.1.3. Free calcium.
Calcium ion is an indispensable ion for maintaining normal physiological activities of the body, and it is very important for the regulation of electrical activity on both sides of the cell membrane. At the same time, calcium intake can affect the signal transduction pathway, promote the secretion of vascular endothelial factor and increase hypoxia inducible factor. In the last century, X-ray microanalysis has been performed on freeze-dried cryosections of normal, hyperplastic, and neoplastic human prostate, studies had found that calcium is the major prostate acinar cell cation. In recent years, a number of studies at home and abroad have also shown that calcium-binding proteins can activate a variety of pathways to promote the spread of invasive prostate cancer cells. In addition, our experiment results, which have shown, high iCa may avoid calcium-binding proteins creating.

In this experiment, as a protective factor of prostate cancer, the OR value of iCa is 0.002, which is of little statistical significance and has little guiding significance for clinical work, but it has a certain enlightening effect on scientific research.

4.1.4. Creatine kinase isoenzyme.
Creatine kinase isoenzyme (CK-MB) is mainly used in the diagnosis of myocardial infarction. However, in the early years, A Gries et al. accidentally found that the number of CK-MB may be related to prostate tumors. Since then, based on the continuous development of proteomics, many scholars have suggested that CK-MB as a marker of malignant tumor should be included in clinical screening. Up to 2015, there is no systematic review or clinical application report on the false increase of CK activity caused by other CK-MB isozymes in malignant tumors.

In this study, CKMB is a risk factor for prostate cancer, the OR value is 1.086, the increase of CKMB will increase the risk of prostate cancer. This suggests that researchers should study and develop new indicators about CK-MB, and provide evidence for previous experiments.

4.2. Implications
Clinically, researchers should also pay attention to patients’ cardiovascular disease and make a timely distinction from prostate cancer. Furthermore, latest research indicates that hasmiR-940 act as a diagnostic and prognostic tool for prostate cancer. Besides, ix co-expressed miRNAs (hsa-miR-17-3p, −377-3p, −410-3p and −495) and p2 miRNA panel (hsa-miR −377-3p, −410-3p, −27a-3p, 149-5p and 940) mainly associated with prostate cancer. In other aspects, respect is a non-invasive, label-free, laser-based technique that identifies molecular composition of tissues and cells, which experiments have demonstrated that such technique could provide insight into different pathways leading to pre-cancerous anal squamous intraepithelial lesions. It is believed that can also be extended in several carcinoma, including prostate cancer. Meanwhile, previous studies had indicated that bone scan-negative patients with a relatively high PSA level and velocity, the risk of distant disease is much greater, and PET imaging may serve as a useful whole-body staging method. Now more tracers for PET/CT are shown to be more accurate in the detection of recurrent disease as compared with radiolabelled choline PET/CT. It is exciting for the clinical doctors to improve the efficiency of diagnostic tools in the future. In the level of genomics, researchers suggested that variations in tumor epigenetic landscape of individuals are partly mediated by genetic differences, which may affect prostate cancer progression. It inspires us that these results could be applied in clinical practice that is helpful to distinguish indolent prostate cancer from advanced disease.

5. Conclusion
In this study, the innovative use of propensity score matching method reduces the differences between groups in the data, makes a better comparison between groups. In addition, this experiment also introduces the neural network model to improve the adaptability of the model to the nonlinear relationship between different complex variables. Among them The logistic regression model performed very well (ROC, 0.963; 95% confidence interval, 0.951–0.978) and artificial neural network model (ROC, 0.983; 95% confidence interval, 0.964–0.997). The most important was that Apo lipoprotein E, Apo lipoprotein C2, Apo lipoprotein C3, Triglyceride, High density lipoprotein cholesterol iCa and CKMB are risk factors related to prostate cancer that have never been discovered or disputed, increase the trust of the known evidence or point out the direction for future research. What is more, increasing the apo test in the physical routine examination is a better way to improve the accuracy of the prostate cancer screening.

5.1. Limitation
This study also has many shortcomings: the final model does not involve pathological diagnosis, MRI imaging, Gleason score, digital rectal examination and other strong pathological factors as risk factors, since all data were collected through routine physical examinations; the short periods between risk measure and incidence of prostate cancer identification; and the exclusion of additional unmeasured or unexamined variables. Besides, the number of patients need to be increased and expected to do conduct a multi-center study. In addition, there are not many related research reports in China, and the experiments based on a
certain factor are not convincing, and more experts and scholars
are needed to provide external medical record data to verify the
advantages and disadvantages of the model. For propensity score
matching, Propensity score matching cannot assess and balance
all possible outcome-influencing factors, such as the LDL,[50] such as the LDL, several researches have indicated the pathways that are activated
by LDL.[51] But the LDL has been excluded in this model. Furthermore, the availability of clinical practices based on a large
number of validations to test.[52] Despite these limitations, this is
the first significant study of clinical prediction modeling assessing
the incidence risk of prostate cancer by biomedical parameters in
China. Besides, these new parameters that were digged in this
study also inspire us to explore the inner connection and
molecular functions between the biomedical index and prostate
cancer. In order to better apply this model and related research
to China’s domestic clinical work. Up to now, there are a lot of risk
calculators have been set up,[53] we hope more and more scholars
to work in this area.

Acknowledgments

Thanks to the Population and Health Science Data Sharing
Platform, National Clinical Medical Science Data Center 301
Hospital provides data support and related help for this article.
Thanks for M.D. Xue Sheng’s careful guidance.

Author contributions

Conceptualization: Hanxu Guo, Xianjie Jia.

Data curation: Hanxu Guo.

Formal analysis: Hanxu Guo.

Funding acquisition: Hanxu Guo.

Investigation: Hanxu Guo.

Methodology: Hanxu Guo.

Project administration: Hanxu Guo.

Resources: Hanxu Guo.

Software: Hanxu Guo.

Supervision: Hanxu Guo.

Validation: Hanxu Guo.

Visualization: Hanxu Guo.

Writing – original draft: Hanxu Guo.

Writing – review & editing: Hao Liu, Xianjie Jia.

References

[1] RL S, KD M, A J. Cancer statistics, 2019. CA 2019;69:7–34.
[2] Al P, BA M, PC A, et al. The role of increasing detection in the rising
incidence of prostate cancer. JAMA 1995;273:548–52.
[3] DC G, SJ C, DO K, et al. Screening for prostate cancer: US preventive
services task force recommendation statement. JAMA 2018;319:
1901–13.
[4] Wilt T, Vo T, Langetmo L, et al. Radical prostatectomy or observation
for clinically localized prostate cancer: extended follow-up of the Prostate
Cancer Intervention Versus Observation Trial (PIVOT). Eur Urol 2020;
77:713–24.
[5] Evans S, Millar J, Davis I, et al. Patterns of care for men diagnosed with
prostate cancer in Victoria from 2008 to 2011. Med J Australia
2013;198:540–5.
[6] Wilt T, Dahm P. PSA screening for prostate cancer: why saying No is a
high-value health care choice. J Natl Compr Canc Netw 2015;13:
1566–74.
[7] Y Z, JY W, YJ S, et al. External validation of the prostate cancer
prevention trial and the European randomized study of screening
for prostate cancer risk calculators in a Chinese cohort. Asian J Androl
2012;14:738–44.
[35] L M, R B, B H, et al. Hypercholesterolemia and prostate cancer: a hospital-based case-control study. Cancer Causes Control 2008;19:1259–66.

[36] KE T, J H, G K, et al. Intracellular distribution of calcium and zinc in normal, hyperplastic, and neoplastic human prostate: X-ray microanalysis of freeze-dried cryosections. Prostate 1989;15:41–51.

[37] G S. S100A8 and S100A9: new insights into their roles in malignancy. J Innate Immunity 2012;4:31–40.

[38] Maly IV, Domaradzki TM, Gosy VA, et al. Myosin is oform expressed in metastatic prostate cancer stimulates cell invasion. Sci Rep 2017;7:8476.

[39] A G, F W, M W, et al. False increased CK-MB value after cryoablation of the prostate without myocardial infarct. Anaesthesiologie, Intensivmedizin, Notfallmedizin, Schmerztherapie 1997;32:580–2.

[40] HG H, KK W, WR W, et al. Clinical applications of microarray technology: creatine kinase B is an up-regulated gene in epithelial ovarian cancer and shows promise as a serum marker. Gynecologic Oncol 2005;96:77–83.

[41] M A, R G, M F, et al. Macro creatine kinase: illness marker. Practical guide for the management. Anales de Medicina Interna 2006;23:272–5.

[42] CC C, CB L, MJ S, et al. Creatine Kinase (CK)-MB-to-Total-CK Ratio: a Laboratory Indicator for Primary Cancer Screening. Asian Pacif J Cancer Prevention 2015;16:6599–603.

[43] Rajendiran S, Parwani AV, Hare RJ, et al. MicroRNA-940 suppresses prostate cancer migration and invasion by regulating MIEN1. Mol Cancer 2014;13:250.

[44] Liu Y, Wang M, Luo Y, et al. MiRNA-target network analysis identifies potential biomarkers for Traditional Chinese Medicine (TCM) syndrome development evaluation in hepatitis B caused liver cirrhosis. Sci Rep 2017;7:11054.

[45] Oda R, Agsalda-Garcia M, Loi N, et al. Raman-enhanced spectroscopy distinguishes anal squamous intraepithelial lesions in human immunodeficiency virus-serodiscordant couples. AIDS Res Hum Retroviruses 2019;35:287–94.

[46] Li R, Ravizzini G, Gorin M, et al. The use of PET/CT in prostate cancer. Prostate Cancer Prostatic Dis 2018;21:4–21.

[47] Silverman D, Hoh C, Seltzer M, et al. Evaluating tumor biology and oncological disease with positron-emission tomography. Semin Radiat Oncol 1998;8:183–96.

[48] Evangelista L, Briganti A, Fanti S, et al. New clinical indications for (18) F/(11)C-choline new tracers for positron emission tomography and a promising hybrid device for prostate cancer staging; a systematic review of the literature. Eur Urol 2016;70:161–75.

[49] Lin P, Giannopoulou E, Park K, et al. Epigenomic alterations in localized and advanced prostate cancer. Neoplasia 2013;15:373–83.

[50] Reiffel JJT Ajom. Propensity score matching: the ‘Devil is in the details where more may be hidden than you know. Am J Med 2020;133:178–81.

[51] Jung YY, Ko J-H, Um J-Y, et al. LDL cholesterol promotes the proliferation of prostate and pancreatic cancer cells by activating the STAT3 pathway. J Cell Physiol 2021;236:5253–64.

[52] Raghunathan K, Layton JB, Ohnuma T, et al. Observational research using propensity scores. Adv Chronic Kidney Dis 2016;23:367–72.

[53] YS W, NZ, SH L, et al. The Huashan risk calculators performed better in prediction of prostate cancer in Chinese population: a training study followed by a validation study. Asian J Androl 2016;18:925–9.