СРПСКИ АРХИВ
ЗА ЦЕЛОКУПНО ЛЕКАРСТВО
SERBIAN ARCHIVES
OF MEDICINE

Address: 1 Kraljice Natalije Street, Belgrade 11000, Serbia
☏ +381 11 4092 776, Fax: +381 11 3348 653
E-mail: office@srpskiarhiv.rs, Web address: www.srpskiarhiv.rs

Paper Accepted*

ISSN Online 2406-0895

Original Article / Оригинални рад

Milorad M. Stojadinović[1,†], Damnjan N. Pantić[2], Miroslav M. Stojadinović[1,2]

External validation of prostate health index-based nomogram for predicting prostate cancer at extended biopsy

Екстерна валидација номограма заснованих на простата здравственом индексу у предикцији карцинома простате при проширеној биопсији

[1]University of Kragujevac, Faculty of Medical Sciences, Kragujevac, Serbia;
[2]Kragujevac Clinical Centre, Clinic of Urology and Nephrology, Department of Urology, Kragujevac, Serbia;

Received: December 26, 2018
Revised: August 2, 2019
Accepted: September 18, 2019
Online First: September 26, 2019
DOI: https://doi.org/10.2298/SARH181226107S

*__Accepted papers__ are articles in press that have gone through due peer review process and have been accepted for publication by the Editorial Board of the *Serbian Archives of Medicine*. They have not yet been copy-edited and/or formatted in the publication house style, and the text may be changed before the final publication.

Although accepted papers do not yet have all the accompanying bibliographic details available, they can already be cited using the year of online publication and the DOI, as follows: the author's last name and initial of the first name, article title, journal title, online first publication month and year, and the DOI; e.g.: Petrović P, Jovanović J. The title of the article. Srp Arh Celok Lek. Online First, February 2017.

When the final article is assigned to volumes/issues of the journal, the Article in Press version will be removed and the final version will appear in the associated published volumes/issues of the journal. The date the article was made available online first will be carried over.

†**Correspondence to:**
Miroslav M. STOJADINOVIĆ
Kragujevac Clinical Centre, Clinic of Urology and Nephrology, Department of Urology, Zmaj Jovina 30, 34000 Kragujevac, Serbia
E-mail: **midinac@gmail.com**

External validation of prostate health index-based nomogram for predicting prostate cancer at extended biopsy

Екстерна валидација номограма заснованих на простата здравственом индексу у предикцији карцинома простате при проширеној биопсији

SUMMARY

Introduction/Objective Prostate Health Index (PHI)-based nomograms were created by the Lughezzani et al. and Zhu et al. for predicting prostate cancer (PCa) at extended biopsy. The aim of the study was to externally validate two nomograms in Serbian population.

Methods This retrospective study comprised 71 patients irrespective of digital rectal examination (DRE) findings, with prostate-specific antigen level <10 ng/ml, who had undergone prostate biopsies, and PHI testing. Data were collected in accordance with previous nomograms predictors. Independent predictors were identified by using logistic regression. The predictive accuracy was measured by the area under the receiver operating characteristic curve (AUC). The calibration belt was used to assess model calibration. The clinical utility was measured by using decision curve analysis (DCA).

Results There were numerous differences in underlying risk factors between validation dataset and the previously available data. Analysis demonstrated that the DRE and PHI were independent predictors. AUCs for both nomograms, in patients with normal DRE had shown to have a good discriminatory ability (77.2–86.2%). In the entire population AUC of nomogram had exceptional discrimination (92.9%). Zhu et al. nomogram is associated with lower false positive predictions. The calibration belt for Zhu et al. nomogram was acceptable. Our DCA suggested that both nomograms are likely to be clinically useful.

Conclusion We performed external validation of two PHI-based nomograms predicting the presence of PCa in both the initial and the repeat biopsy setting. The PHI-based nomograms displayed adequate accuracy and justifies its use in Serbian men.

Keywords: prostate cancer; prostate biopsy; external validation; nomogram; Prostate Health Index

САЖЕТАК

Увод/Циљ Лугецани са сарадницима и Жу са сарадницима креирали су номограме засноване на Простата здравственом индексу (*PHI*) у предвиђању карцинома простате (КП) при проширеној биопсији. Циљ студије је да екстерно валидира ове номограме у Српској популацији.

Методе Ова ретроспективна студија укључила је 71 болесника, независно од дигиторектално налаза, са серумским нивоом специфичног антигена простате (ПСА) мањим од 10 нг/мл, код којих је учињена биопсија простате, и *PHI* тестирање. Прикупљани су подаци о претходно дефинисаним предикторима у номограмима. Коришћена је логистичка регресија за идентификацију независних предиктора. Предиктивна тачност процењена је пољем испод *ROC* криве (*AUC*). Калибрација номограма процењена је калибрационим појасом. Клиничка корисност је процењена анализом криве одлучивања (*DCA*).

Резултати Постојале су бројне разлике у предиспонирајућим факторима ризика наше валидационе базе података са претходно публикованим подацима из којих су изведени номограми. Анализа је показала да су дигиторектални налаз и *PHI* независни предиктори. Код болесника са нормалним дигиторектални налазом *AUC* за оба номограма су показала добру дискриминациону способност (77.2–86.2%).У целој популацији AUC номограма показао је изузетну дискриминацију (92.9%). Номограм Жуа и сарадника је повезан са мање лажно позитивних предикција. Калибрациони појас за номограм Жуа и сарадника био је прихватив. Наша *DCA* указује да оба номограма могу бити клинички корисна.

Закључак Спроведена је екстерна валидација два номограма заснованих на *PHI* који предвиђају присуство КП при иницијалној или поновљеној биопсији. Номограми засновани на *PHI* показали су добру тачност и оправдавају употребу код српских мушкараца.

Кључне речи:карцином простате; биопсија простате; екстерна валидација; номограм; Простата здравствени индекс;

INTRODUCTION

Prostate cancer (PCa) is the most prevalent cancer among male population in Europe and the sixth main cause of mortality due to cancer in men worldwide [1]. Contemporary guidelines recommend ten to twelve core systematic transrectal-ultrasounds (TRUS)–guided prostate needle biopsy for early discovery of PCa [2]. Due to lack of specifity common risk factors and treatment complications with prostate biopsy, several prediction tools were introduced to assist with the identification of those at highest risk of detecting PCa on prostate needle biopsy and avoid unnecessary biopsies.

Several nomograms have been developed to predict individual PCa outcomes that range from biopsy outcome prediction in men at risk of PCa through prediction of increase in Gleason score grade between biopsy and radical prostatectomy pathology to prediction of specific direction and location of extracapsular invasion at RP and mortality rate from hormone-refractory PCa [3]. The predictive accuracy of the nomogram extended from 73% to 76% in prediction of PCa detection. Furthermore, compared to extended biopsy schemes, earlier predictive nomograms (sextant biopsy) are less accurate in predicting the chance of PCa [4]. Discrepancies in disease risk factors may influence the performance of nomogram. Hence, they have to be approved before using in a specific geographic region and in contemporary patients. If a predictive tool is used for a population that differs from the one used for its development, it should be externally validated so that it can provide general and clinical appropriateness. In addition, nomograms should be reassessed regularly [5].

Recent studies have shown that Prostate Health Index (PHI), precursor PSA isoform [−2]proPSA (p2PSA) derivative, may increase our capability to discriminate patients with and without PCa independently or in models [6-9]. Recently developed PHI-based nomorams [10, 11] incorporated several traditional PCa factors, along with PHI.

Based on these considerations, the aim of the study was to externally validate two published PHI-based nomograms for predicting individual risk for PCa at extended biopsy within a Serbian population and compare their predictive accuracy.

METHODS

Patient population

We validated two published PHI-based nomograms using patients who had undertaken TRUS-guided prostate biopsies and p2PSA testing, between May 2017 and December 2017 at Clinical Centre Kragujevac in accord with standards of the institutional committee on ethics. Inclusion criteria were PSA level <10 ng/ml and at least 10 core biopsies undergone. This retrospective study comprised 71 patients irrespective of DRE findings. The study was permitted by the institutional review boards (01/17/2608). Patients with incomplete data, acute bacterial prostatitis and patients who had undergone previous endoscopic surgery of the prostate were excluded as well as those being treated with dutasteride or finasteride. Patients with chronic kidney disease, hemophilia or previous polytransfusion were also excluded, as these conditions may change the concentration of p2PSA. Data were collected regarding the candidate predictors in accordance with previous nomograms. The Zhu et al nomogram [11] is based on three criteria: age, prostate volume (PV) and PHI; the Lughezzani et al. nomogram [10] was constructed using the following predictors: age, DRE, PV, biopsy history and PHI.

At presentation, blood samples were drawn prior to biopsy and any prostate manipulation using regular methods and were processed and frozen at −70°C within 8 hours for future analysis. Samples were defrosted and analysed for tPSA and [–2]proPSA simultaneously using UniCelDxI 600 Access Immunoassay System, Beckman Coulter, USA. The equation (p2PSA/fPSA)*√PSA was used to calculate PHI.

DRE were done by an urologist on all patients. The DRE was assigned as normal, or suspicious/positive. In order to gain ultrasound data and prostate biopsy, Toshiba (Aplio 300) ultrasound device with 5-10-MHz probe was used. Prostate volumes were calculated by measuring the gland in three dimensions, and using the following formula: 0.52 [length (cm) × width (cm) × height (cm)]. TRUS-guided prostate biopsies were performed according to a standardized extended scheme. After obtaining a median of twelve core biopsies (range, 10 to 12 cores), it was assessed by local pathologists.

Statistical analyses

Descriptive statistics was used for predictor variables. Univariate and multivariate logistic regression analyses with Backward–Wald stepwise were used in order to identify and quantify the independent predictors of PCa. The results were expressed in odds ratios (ORs) with 95% confidential interval (CI).

For patients with a normal DRE the probability of PCa was calculated according to Lughezzani et al. and Zhu et al. PHI-based nomogram and compared with their outcome and for the entire population with a suspected and nonsuspected DRE, only the Lughezzani et al. nomogram was aplied. We assigned the points of each attribute of the patient by drawing a vertical line from that variable to the points' scale, then, sum all the points, and draw a vertical line from the total points scale to obtain the probability of PCa. The predictive accuracy (c-index) was measured by the area under the receiver operating characteristic curve (AUC). We calculated AUC analysis and the Brier score for each nomogram, and compared AUCs by the DeLong test. The Hosmer–Lemeshow (HL) goodness-of-fit statistics was used to assess model calibration and we also plotted a calibration belt [12]. The calibration belt is a fitted polynomial logistic function curve between the logit transformation of the predicted likelihood and result with surrounding 80% and 95% CI [13]. We also compared the specificities of PHI-based nomograms at 90% sensitivities using a bootstrap based method [14]. By using decision curve analyses (DCA), clinical usefulness was assessed [15]. All analyses were performed using SPSS version 23.0 (SPSS Inc., Chicago, IL) or STATA version 13.0 (STATA Corp., TX, USA). Statistical significance was set at $p < 0.05$.

RESULTS

Table 1 represents the features of the patients used for each PHI-based nomogram and our validation cohort. Comparison between our validation dataset and the previously published data has shown numerous differences in underlying risk variables. The mean age was similar in all cohorts. Except disparity in study period, the proportion of men manifested with suspicious findings on DRE was also different (17.7% vs. 28.2%, P= 0.044), while Zhu et al. included only patients with normal DRE. Chinese men had significantly smaller prostate glands ($p<0.001$), the lowest p2PSA value and the lowest detection rate. Similar to our validation cohort Lughezzani et al. included both initial and repeat biopsy, while Zhu et

al. nomogram was confined to initial biopsy. There was a notable difference between the original cohort and the validation cohort in regard to repeated biopsies (p=0.01). Our patients had significantly lower tPSA comparing to Chinese men (p<0.001). The highest median value of PHI was established in the European cohort.

The univariate logistic regression has shown that all of variables with the exception of biopsy history were significant predictors of PCa. However, only DRE and PHI sustained their prognostic significance during multivariable analyses (Table 2).

AUC for both nomograms, in patients with normal DRE showed to have a good discriminatory ability (77.2–86.2%) (Figure 1, Table 3), and in pairwise comparison of ROC curves the difference between areas of Zhu et al. and Lughezzani et al. nomogram (9%) was nonsignificant (p = 0.229). In the entire population, AUC of nomogram had exceptional discrimination (92.9%), and their predictive accuracy was not significantly lower (p = 0.312) comparing to patients with normal DRE. All HL tests had p value higher than 0.05, indicating that there are no significant differences between the observed and expected outcomes and consequently all models suggest good overall calibration. The better (lower) value of Brier score was for nomogram by Zhu et al.

Figure 2 presents both nomograms calibration belt as related to the external validation dataset, in patiens with normal DRE (Figures 2a and 2b), and in the entire population (Figure 2c). The predicted probability of the previously reported nomograms is represented on the x-axis, and the actual proportion of biopsy-proven PCa is represented on the y-axis. The calibration belt for Zhu et al. nomogram was acceptable only, and showed deviations irrelevant from ideal calibration (Figure 2b). Conversely, for Lughezzani et al. nomogram the calibration curve calibrates poorly in all risk range, in the entire cohort (Figure 2c), and overestimated PCa in the first three risk deciles, in patients with normal DRE (Figure 2a).

In patients with normal DRE, at a 90% sensitivity, the specificity of the Zhu et al. nomogram (88.4%) was significantly higher (p = 0.011) than the specificity of the Lughezzani et al. nomogram (66.5%). This phenomenon indicates that Zhu et al. nomogram is associated with lower false positive predictions.

Figure 3 shows the results of the DCA. All biopsy strategies suggest that if all patients are biopsied, all will avoid an unfavorable outcome. If the risk is higher than 8% and if patients agree to undergo further intervention, our DCA suggested that both nomograms have

a chance to be suitable for that. However, Zhu et al. nomogram (green line) lead to the higher net benefit compared with Lughezzani et al. nomogram (purpure line) in various threshold probabilities above approximately 18% (Figure 3a). However, their curves are partly overlapping. The reduction in the number of avoidable biopsies per 100 patients is net of false negatives, without a decrease in the number of patients with PCa who duly have PCa. Also, in this case, Zhu et al. nomogram(green line) outperformed Lughezzani et al. nomogram (purpure line) above approximately 18% (Figure 3c). For example, at a probability threshold of 20%, the use of the Lughezzoni et al and Zhu et al. nomogram decreases the number of avoidable biopsies by about 45 and 55 per 100 patients, respectively, without missing any of PCa.

DISCUSSION

Various methods have been suggested to determine the likelihood of PCa, which may decrease the amount of avoidable prostate biopsies in near future. We assessed the performance of an earlier developed PHI-based nomogram by studying three aspects of validity: discrimination, calibration and clinical usefulness. In the present population, our external validation results validated a proper precision of the previously developed nomograms for predicting the likelihood of PCa in the initial and repeat biopsy setting. The superior diagnostic value of Zhu et al. nomogram over Lughezzani et al. nomogram was evidenced in patients with normal DRE. The clinical benefit of the PHI-based nomograms was additionally confirmed by DCA. These results suggest that peviously developed nomograms may help clinicians and patients to make evidence-based choices for prostate biopsy based on patients' individual conditions.

Previous existing nomograms have established criteria associated with higher risk of PCa in the initial and repeat biopsy setting. They included age [4, 10, 11, 16-23], race [22], digital rectal examination [4, 10, 16-22], total PSA [4, 16-23], percent free PSA [4, 16, 18-21], prostate volume [10, 11, 17, 20-22], PSAD [19, 23],hypoechoic lesions on ultrasound [19, 21], biopsy history [10, 23], family history [22], PHI [6-8, 10, 11], PHI density [9], Prostate Cancer gene-3(PCA3) [22] and magnetic resonance imaging (MRI) [23]. Despite several variables having shown statistically significant prediction value in the univariate analysis, only few sustained their independent value in the multivariate analysis. According

to the analysis, encouraging prediction of PCa is possible on the basis of DRE and PHI. Our findings were in accordance with earlier studies that PHI, as part of a multivariable approach, was the most accurate in predicting PCa at initial and repeat biopsy [6, 8].

Earlier developed predictive models or nomograms (sextant biopsy) are less precise in predicting the likelihood of PCa on initial biopsy [4]. Extended biopsy schemes changed the rate of PCa detection as well as the capability of typical risk factors, such as percent free PSA, to predict the likelihood of PCa on needle biopsy. Futhermore, concept of sampling density supported the idea to increase the number of core biopsies in order to improve the diagnostic yield [4].

The earlier developed PHI-based nomograms verified their capability to determine the presence of PCa at biopsy in their original cohort [10, 11]. Validation on diverse external data sets allows for assessment of the generalizability of the prediction tool to wider population than originally stated. Additionally, it is generally believed that external validation is more reliable then internal validation for prediction models, since it is insisting on transportability rather than reproducibility [24]. We are not the first researchers to carry out a validation between different PHI-based nomograms. When the nomogram applied to five external validation populations from European tertiary care centers, its yielded moderate predictive accuracies of 75.2% [5]. In our study we found that the accuracy was better (77.2-92.9) than the accuracy of many earlier ones (70-77%) which externally validated different nomograms [4, 16, 20, 22].

Calibration is one of the crucial features of every predictive model. Unfortunately, using the traditional approach of calibration (HL test, calibraton plot), still shows several limitations. The traditional plot is not supplemented by any data on the statistical significance of deviations from the bisector [12]. On the other hand, the calibration belt is providing information on the direction, extent, and risk classes affected by divergences between the observed and predicted PCa [13]. In the analysis, only Zhu et al. nomogram had acceptable calibration. This is probably due to varieties between populations. Exept disparity in study period, there were significant dissimilarities between the original and the validation cohort which include inclusion critera (variety of PSA ranges, DRE findings), the incidence of PCa, proportion of men presenting with doubtful findings on DRE, prostate volume, tPSA, p2PSA, PHI and biopsy history. It indicates that certain patient characteristics are the difference in distribution between the validation sample and the development sample. It is questionable

whether perfect calibration could be achieved in practice by any model. Also, time variation may be a potential explanation why the previous models are not considered to be better than the recent ones. Although these differences most likely affect our calibration of PCa, they allow validity, and maybe generality, of a model to a more diverse and various populations. We also consider models originated from a specific country more convenient for local utilization [21].

In our DCA we confirmed clinical uselessness of these PHI-based nomograms. We also identified range of threshold probabilities (<10%) in which nomograms were of value. In patients with normal DRE, Zhu et al. nomogramlead to the higher net benefit compared with Lughezzani et al. nomogramin various threshold probabilities above approximately 18%. Futhermore, Zhu et al. nomogram is associated with lower false positive predictions, when specificity is observed at fixed sensitivity. Superiority of Zhu et al. nomogram could be partly explained by its derivation from men with normal DRE.

The most significant limitation of this study is small validation cohort from a single institution. The differences in population characteristics for both nomograms development and the validation cohort were the next difficulty. Furthermore, regardless of the use of a standardized comprehensive biopsy scheme, the PCa discovery rate may have been dissatisfactory in some of these patients. Lastly, diagnostic imaging is turning into an essential element of prostate cancer diagnosis. Multiparametric MRI is helping clinicians with new information to better guide prostate biopsies [23]. However, we have shown that the nomogram remains highly predictive even in the different population and may be a significant tool to help clinicians in discriminating between patients with and without PCa. Nevertheless, when making decision about carrying out prostate biopsy we should consider multiple factors, including the patient's life expectancy, co-morbidity, and preference apart from risk of PCa. Secondary, it is also important to notice that clinicians could have lack of enthusiasm to use predictive tools. A United States survey has shown that only 35.5% of radiation oncologists and urologists currently use a decision aid in clinical practice. [25]. We believes that there was no similar nomogram that has been developed or validated in Serbian population.

CONCLUSION

In our study, we performed external validation of two PHI-based nomograms predicting the probability of PCa in both the initial and the repeat biopsy setting. The PHI-based nomogram displayed adequate accuracy and calibration properties. The satisfying performance of the nomograms in the validation cohort justifies its use in Serbian men.

ACKNOWLEDGEMENTS

The authors were financially supported through a research grant N0175014 of the Ministry of Science and Technological Development of the Republic of Serbia. The authors thank the Ministryfor this support.

Conflict of interest: None declared.

REFERENCES

1. Arnold M, Karim-Kos HE, Coebergh JW, Byrnes G, Antilla A, Ferlay J, et al. Recent trends in incidence of five common cancers in 26 European countries since 1988: Analysis of the European Cancer Observatory. Eur J Cancer. 2015;51(9):1164-87. doi: 10.1016/j.ejca.2013.09.002. PMID: 24120180.
2. Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. Eur Urol. 2017; 71(4):618-29. doi: 10.1016/j.eururo.2016.08.003. PMID: 27568654.
3. Chun FK, Karakiewicz PI, Briganti A, Gallina A, Kattan MW, Montorsi F, et al. Prostate cancer nomograms: an update. Eur Urol. 2006;50(5):914-26; PMID: 16935414.
4. Chun FK, Briganti A, Graefen M, Montorsi F, Porter C, Scattoni V, et al. Development and external validation of an extended 10-core biopsy nomogram. Eur Urol. 2007;52(2):436-44. PMID: 17010505.
5. Lughezzani G, Lazzeri M, Haese A, McNicholas T, de la Taille A, Buffi NM, et al. Multicenter European external validation of a prostate health index-based nomogram for predicting prostate cancer at extended biopsy. Eur Urol. 2014;66(5):906-12. doi:10.1016/j.eururo.2013.12.005. PMID: 24361258.
6. Loeb S, Sanda MG, Broyles DL, Shin SS, Bangma CH, Wei JT, et al. The prostate health index selectively identifies clinically significant prostate cancer. J Urol. 2015; 193(4): 1163-9. doi: 10.1016/j.juro.2014.10.121. PMID: 25463993.
7. de la Calle C, Patil D, Wei JT, Scherr DS, Sokoll L, Chan DW, et al. Multicenter Evaluation of the Prostate Health Index to Detect Aggressive Prostate Cancer in Biopsy Naïve Men. J Urol. 2015; 194(1):65-72. doi: 10.1016/j.juro.2015.01.091. PMID: 25636659.
8. Roobol MJ, Vedder MM, Nieboer D, Houlgatte A, Vincendeau S, Lazzeri M, et al. Comparison of Two Prostate Cancer Risk Calculators that Include the Prostate Health Index. Eur Urol Focus. 2015; 1(2):185-190. doi: 10.1016/j.euf.2015.06.004. PMID: 28723432.
9. Tosoian JJ, Druskin SC, Andreas D, Mullane P, Chappidi M, Joo S, et al. Prostate Health Index density improves detection of clinically significant prostate cancer. BJU Int. 2017; 120(6):793-8. doi: 10.1111/bju.13762. PMID: 28058757.
10. Lughezzani G, Lazzeri M, Larcher A, Lista G, Scattoni V, Cestari A, et al. Development and internal validation of a Prostate Health Index based nomogram for predicting prostate cancer at extended biopsy. J Urol. 2012;188(4):1144-50. doi: 10.1016/j.juro.2012.06.025. PMID: 22901589.
11. Zhu Y, Han CT, Zhang GM, Liu F, Ding Q, Xu JF, et al. Development and external validation of a prostate health index-based nomogram for predicting prostate cancer. Sci Rep. 2015;5:15341. doi: 10.1038/srep15341. PMID: 26471350.
12. Finazzi S, Poole D, Luciani D, Cogo PE, Bertolini G. Calibration belt for quality-of-care assessment based on dichotomous outcomes. PLoS One. 2011;6(2):e16110. doi: 10.1371/journal.pone.0016110. PMID: 21373178.
13. Nattino G, Finazzi S, Bertolini G. A new calibration test and a reappraisal of the calibration belt for the assessment of prediction models based on dichotomous outcomes. Stat Med. 2014; 33(14):2390-407. doi: 10.1002/sim.6100. PMID: 24497413.
14. Qin G, Hsu YS, Zhou XH. New confidence intervals for the difference between two sensitivities at a fixed level of specificity. Stat Med. 2006; 25:3487-502. PMID: 16345124.
15. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. BMJ. 2016; 25;352:i6. doi: 10.1136/bmj.i6.PMID: 26810254.
16. Karakiewicz PI, Benayoun S, Kattan MW, Perrotte P, Valiquette L, Scardino PT, et al Development and validation of a nomogram predicting the outcome of prostate biopsy based on patient age, digital rectal examination and serum prostate specific antigen. J Urol. 2005;173(6):1930-4. PMID: 15879784.

17. Lee A, Lim J, Gao X, Liu L, Chia SJ. A nomogram for prediction of prostate cancer on multi-core biopsy using age, serum prostate-specific antigen, prostate volume and digital rectal examination in Singapore. Asia Pac J Clin Oncol. 2017;13(5):e348-e355. doi: 10.1111/ajco.12596. PMID: 27641069.
18. Kawakami S, Numao N, Okubo Y, Koga F, Yamamoto S, Saito K, et al. Development, validation, and head-to-head comparison of logistic regression-based nomograms and artificial neural network models predicting prostate cancer on initial extended biopsy. Eur Urol. 2008;54(3):601-11. doi: 10.1016/j.eururo.2008.01.017. PMID: 18207312.
19. Silva TB, Oliveira CZ, Faria EF, Mauad EC, Syrjänen KJ, Carvalho AL. Development and validation of a nomogram to estimate the risk of prostate cancer in Brazil. Anticancer Res. 2015;35(5):2881-6. PMID: 25964571.
20. Kara O, Elshafei A, Nyame YA, Akdogan B, Malkoc E, Gao T, et al. The nomogram conundrum: a demonstration of why a prostate cancer risk model in Turkish men underestimates prostate cancer risk in the USA. Int Urol Nephrol. 2016;48(10):1623-9. doi: 10.1007/s11255-016-1328-6. PMID: 27236298.
21. Nowroozi MR, Amini S, Kasaeian A, Zavarehei MJ, Eshraghian MR, Ayati M. Development, validation and comparison of two nomograms predicting prostate cancer at initial 12-core biopsy. Asia Pac J Clin Oncol. 2016;12(2):e289-97. doi: 10.1111/ajco.12186. PMID: 24684767.
22. Greene DJ, Elshafei A, Nyame YA, Kara O, Malkoc E, Gao T, et al. External validation of a PCA-3-based nomogram for predicting prostate cancer and high-grade cancer on initial prostate biopsy. Prostate. 2016;76(11):1019-23. doi: 10.1002/pros.23197. PMID: 27197726.
23. Lee SM, Liyanage SH, Wulaningsih W, Wolfe K, Carr T, Younis C, et al. Toward an MRI-based nomogram for the prediction of transperineal prostate biopsy outcome: A physician and patient decision tool. Urol Oncol. 2017;35(11):664.e11-664.e18. doi: 10.1016/j.urolonc.2017.07.018. PMID: 28801025.
24. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J. 2014;35(29):1925-31. doi: 10.1093/eurheartj/ehu207. PMID: 24898551.
25. Wang EH, Gross CP, Tilburt JC, Yu JB, Nguyen PL, Smaldone MC, et al. Shared decision making and use of decision AIDS for localized prostate cancer : perceptions from radiation oncologists and urologists. JAMA Intern Med. 2015;175(5):792-9. doi: 10.1001/jamainternmed.2015.63. PMID: 25751604.

Table 1. Descriptive characteristics of the study population used for previous PHI-based nomograms and our external validation cohort

Characteristics	Lughezzani et al. nomogram	Zhu et al. nomogram	Validation cohort
Study period	July 2010 – July 2011	April 2012 – August 2014	May 2017 – December 2017
Patients, n	729	347	71
DRE, suspicious n (%)	129 (17.7)	0 (0)	20 (28.2)
PCa, n (%)	280 (38.4)	52 (15%)	23 (32.4)
Age, yr mean ± SD/median (range)	64.3 ± 7.8	64 (21)	64.3 ± 5.4
Total PSA, ng/ml median (range)	6.39 (0.5–19.9)	6.89 (3.09)	5.06 (2.03–9.85)
Prostate volume, ml median (range/IQR)	58 (9–230)	40 (23.4)	50 (18–128)
p2PSA, pg/ml, median (range/IQR)	16.4 (0.1–137)	13 (10)	14.3 (3.2–34.2)
PHI, median (range/IQR)	41.2 (6.5–192.8)	32.7 (19.9)	33.3 (14.2–135.4)
Previous biopsy, n (%)	244 (33.5)	0 (0)	10 (14.1)
Number of biopsies, n	≥ 12	≥ 10	≥ 10

DRE – digital rectal examination; PCa – prostate cancer; PHI – prostate health index; PSA – prostate-specific antigen; p2PSA – precursor PSA isoform; SD – standard deviation;

Table 2. Logistic regression analyses of previous nomogram predictors for prostate cancer detection in our validation cohort

Characteristics	Univariate analysis OR (95% CI)	p	Multivariable analysis OR (95% CI)	p
Age	1.105 (1.001–1.220)	0.048		
DRE	16.125 (4.562–56.990)	0.000	7.859 (1.193–51.786)	0.008
tPSA	1.409 (1.084–1.832)	0.010		
Prostate volume	0.963 (0.934–0.994)	0.018		
Biopsy history	0.258 (0.065–1.027)	0.055		
p2PSA	1.132 (1.052–1.218)	0.001		
PHI	1.130 (1.068–1.195)	0.000	1.126 (1.052–1.206)	0.001

DRE – digital rectal examination; CI – confidential interval; OR – odds ratio; PHI – prostate health index; p2PSA – precursor PSA isoform; tPSA – total prostate-specific antigen

Table 3. Predictive accuracy of different nomograms

Predictive accuracy	Lughezzani et al. nomogram	Zhu et al. nomogram
DRE Unsuspicious		
AUC (95% CI)	86.2 (73.6–94.2)	77.2 (63.3–87.8)
HL test χ^2, p value	11.62, 0.169	1.29, 0.257
Calibration belt, test statistic, p value	5.91, 0.015	1.10, 0.294
Brier score	0.111	0.094
DRE Unsuspicious/suspicious		
AUC (95% CI)	92.9 (86.9–98.8)	
HL test χ^2, p value	7.39, 0.495	
Calibration belt, test statistic, p value	9.27, 0.002	
Brier score	0.116	

AUC – area under the receiver operating characteristic curve; CI – confidential interval; DRE – digital rectal examination; HL – Hosmer–Lemeshow test; χ^2 – chi squared test

Table 4. Estimated specificity at fixed sensitivity of 90% for different nomograms and number of avoided biopsies versus the proportion of missed prostate cancer

Characteristics	Lughezzani et al. nomogram			Zhu et al. nomogram		
Sensitivity (90%)	Specificity (95% CI)[a]	Biopsy spread (%)	Missed (%)	Specificity (95% CI)[a]	Biopsy spread (%)	Missed (%)
DRE Unsuspicious	66.5 (49.3–85.8)	58	10	88.4 (76.7–95.4)	76	10
DRE Unsuspicious/suspicious	81.9 (54.2–97.9)	59	10			

DRE – digital rectal examination;

[a]BC_a bootstrap interval (1,000 iterations)

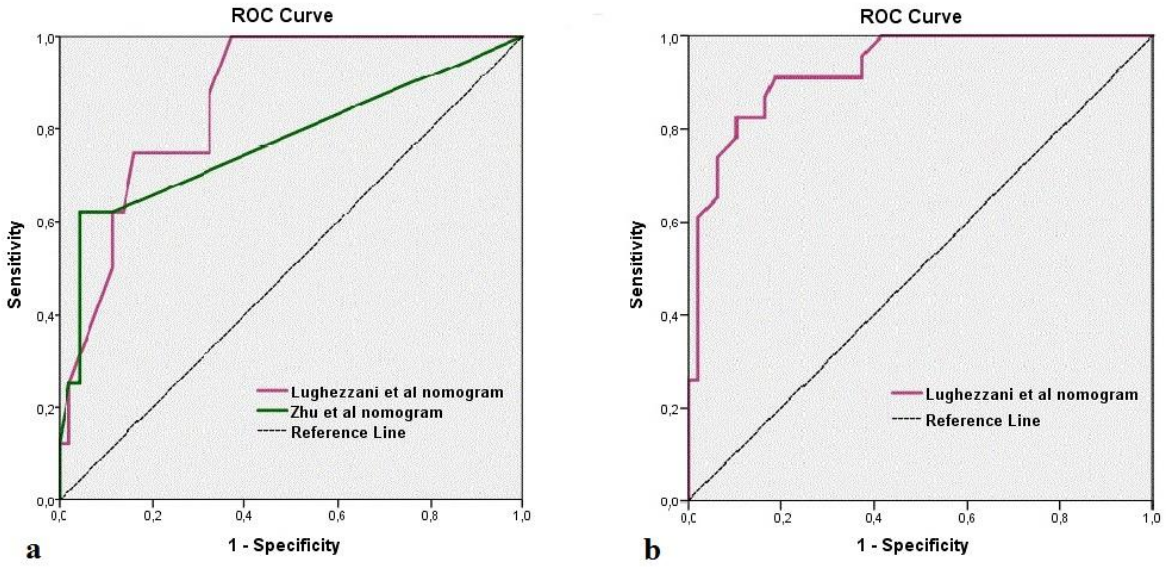

Figure 1. Receiver operating characteristic curve analyses of PHI-based nomograms in: a) patients with normal digital rectal examination; b) the entire validation cohort

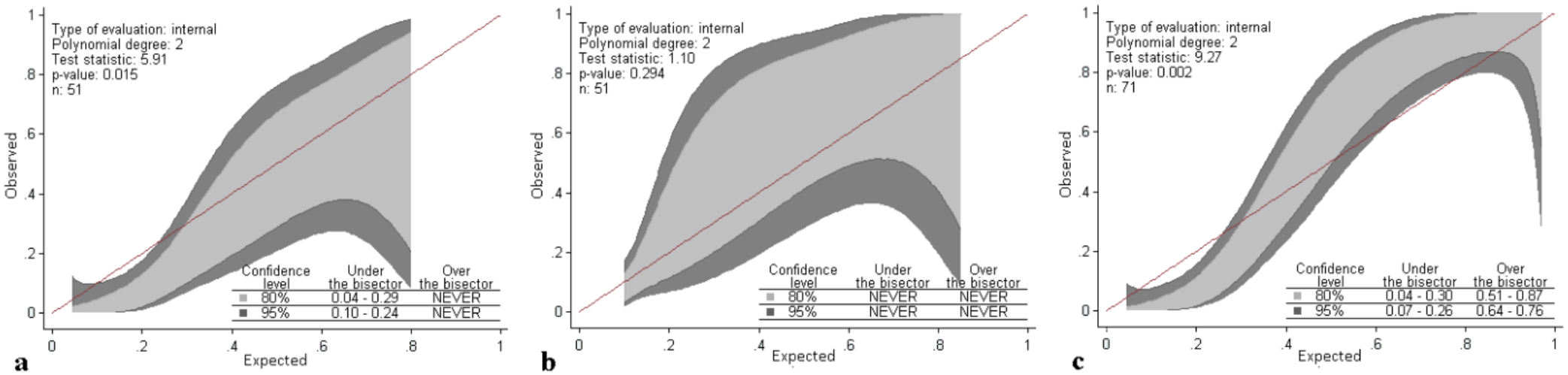

Figure 2. Calibration belt for the PHI-based nomogramsat two confidence level: a) Lughezzani et al. nomogram in patients with normal digital rectal examination; b) Zhu et al. nomogram in patients with normal digital rectal examination; c) Lughezzani et al. nomogram in the entire validation cohort; the degree of the polynomial, the Wald statistics results and the number of patients are given in the upper-left quadrant; confidence intervals: 80% (light gray area) and 95% (dark gray boundaries)

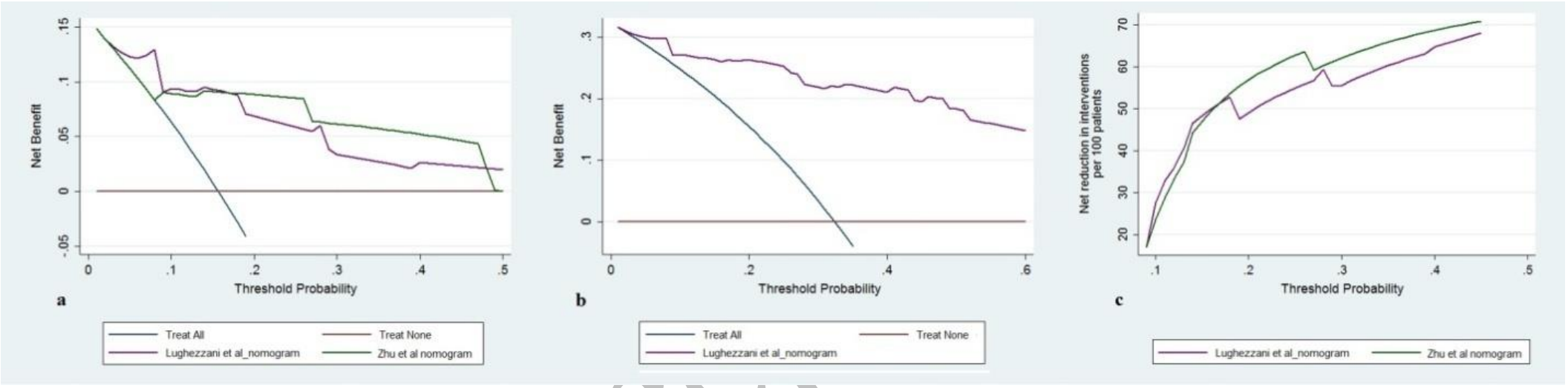

Figure 3. Decision curve analysis of the effect of PHI-based nomogram on the detection of prostate cancer: a) in patients with normal digital rectal examination; b) the entire population; c) net reduction in interventions per 100 patients is plotted against various threshold probabilities; net benefit is compared with 'Biopsied for all' strategy and 'Biopsied for none'

