Abstract

Purpose

To develop a novel Taiwanese prostate cancer (PCa) risk model for predicting PCa, comparing its predictive performance with that of two well-established PCa risk calculator apps.

Methods

1545 men undergoing prostate biopsies in a Taiwanese tertiary medical center between 2012 and 2019 were identified retrospectively. A **5-fold cross-validated** logistic regression risk model was created to calculate the probabilities of PCa and high-grade PCa (Gleason score ≧7), to compare those of the Rotterdam and Coral apps. Discrimination was analyzed using the area under the receiver operator characteristic curve (AUC). Calibration was graphically evaluated with the goodness-of-fit test. Decision-curve analysis was performed for clinical utility. At different risk thresholds to biopsy, the proportion of biopsies saved versus low- and high-grade PCa missed were presented.

Results

Overall, 278/1309 (21.2%) patients were diagnosed with PCa, and 181 out of 278 (65.1%) patients had high-grade PCa. Both our model and the Rotterdam app demonstrated better discriminative ability than the Coral app for detection of PCa (AUC: **0.795** vs 0.792 vs 0.697, DeLong's method: P <0.001) and high-grade PCa (AUC: 0.869 vs 0.873 vs 0.767, P <0.001). Using a ≥10% risk threshold for high-grade PCa to biopsy, our model could save **67.2%** of total biopsies; among these saved biopsies, only **3.4%** high-grade PCa would be missed.

Conclusion

Our new logistic regression model, similar to the Rotterdam app, outperformed the Coral app in the prediction of PCa and high-grade PCa. Additionally, our model could save unnecessary biopsies and avoid missing clinically significant PCa in the Taiwanese population.

Keywords
Diagnosis; mHealth; mobile apps; prostate cancer; prostate-specific antigen; risk calculator.

Introduction

Mobile health (mHealth) has become a growing trend with around 350,000 apps available in the health and fitness and medical categories in 2017 [1]. The World Health Organization referred to mHealth as "mobile wireless technologies for public health" [2], delivering healthcare services via portable devices like basic phones, tablets and wearables. Taiwan has a population of greater than 23 million with an estimated 18.8 million smartphone users in 2020 [3]. Medical professionals and patients might have more ready access to mHealth apps than web-based risk calculators. Encouraging the use of prostate cancer (PCa) risk calculator apps might assist in shared decision-making regarding prostate biopsy between urologists and patients, especially in clinics.

PCa is the second most frequent cancer and the fifth leading cause of cancer deaths globally among males [4], though the age-standardized rate (ASR is lowest for Asian men [5]. However, the ASR of PCa in Taiwan has been rising from 8.58 per 100,000 males in 1996 to 30.5 in 2016, and PCa incidence rates have risen by around 2.5-fold in Taiwanese men over the last 20 years [6]. Although there is no governmental prostate-specific antigen (PSA)-based screening program in Taiwan, increased awareness of cancer screening, westernized diets and the ageing population are all likely contributors to this increase [7].

Nevertheless, PSA is far from being an ideal tumor marker. A PSA >4ng/mL has a sensitivity of 21%, and a specificity of 91% for identifying prostate cancer [8]. The current indications for a biopsy, based on such an abnormal PSA level and/or an abnormal DRE, lead to a myriad of unnecessary biopsies and associated complications, including hematuria, hematochezia, acute urinary retention, urinary tract infection, and sepsis. In order to reduce post-biopsy morbidities, professionals have formulated a number of PCa risk calculators using meaningful predictors to improve predictive accuracy; such multivariable risk approaches have performed better than PSA/DRE alone [9].

The Rotterdam Prostate Risk Calculator and Coral—Prostate Cancer Nomogram Calculator are mHealth apps derived from the two most-studied web-based PCa risk calculators, the European Randomised Study for Screening of Prostate Cancer risk calculator (ERSPC-RC) [10] and the Prostate Cancer Prevention Trial risk calculator version 2.0 (PCPT-RC 2.0) [11], respectively. Medical professionals and patients can easily enter pre-biopsy information into the apps to calculate the risks of PCa and high-grade PCa. **Although ERSPC-RC and PCPT-RC have been well**

validated, most of them have only been validated in independent cohorts; neither superiority nor global applicability has been shown [12]. Additionally, because of ethnic/racial disparity, Eastern Asian urologists have endeavored to develop better approaches for PCa risk prediction in Asian populations [13]. Calibration adjustments might be needed for different hospital settings in Western countries. The applicability of these apps to the Taiwanese population needs to be clarified before widespread implementation.

To the best of our knowledge, this is the first study to develop a novel Taiwanese PCa risk model for predicting PCa and high-grade PCa. We compare its predictive performance with that of the two well-established PCa risk calculator apps.

Methods

Ethics approval was granted by the Internal Review Board of a Taiwanese tertiary medical center (IRB No.: VGHKS19-CT3-13). 1545 men undergoing transrectal ultrasound (TRUS) prostate biopsies from January 2012 to September 2019 inclusive were enrolled. The indication for prostate biopsy included an abnormal PSA level (>4 ng/mL), an abnormal DRE, or a hypoechoic lesion identified on TRUS. All patient data were retrospectively recorded via electronic medical records, including age, family history of PCa and previous biopsy history. **In this study, 315 patients had a history of previous negative biopsy.** PSA was collected as the latest total serum PSA level prior to prostate biopsy. Each patient underwent DRE and TRUS before biopsy was undertaken; prostate volume (PV) was calculated by the ellipsoid formula (length x width x height x $\pi/6$).

A 12-core systematic biopsy strategy has been implemented in our hospital for more than a decade. 50 patients had multi-parametric magnetic resonance imaging (mpMRI) scans on a self-pay basis because this is not approved pre-biopsy for reimbursement by the Taiwan National Health Insurance. All these mpMRI scans were reviewed by dedicated uro-radiologists, and scored using the Prostate Imaging Reporting and Data System version 2 (PI-RADS v2) [14]. **With regard to hypoechoic lesions on TRUS or PI-RADS 4-5 lesions on mpMRI, additional 3 cores would be obtained for each lesion. No MR fusion biopsy was performed among our enrolled patients.** The histopathology of TRUS biopsy specimens was evaluated by consultant uro-pathologists. High-grade PCa was defined as Gleason score ≥7.

The risks of PCa and high-grade PCa were calculated by inputting data into the Rotterdam and Coral apps. Based on the pre-specified variable ranges for both apps, 223 patients were excluded for: PV <10 mL or >110 mL **(n = 60)**, PSA <0.4 or >50 ng/mL **(n = 151)**, or history of pre-biopsy MRI plus age <50 or >75 years **(n = 12)**. **Another 13 patients were excluded because of:** previous positive biopsy (n = 8), different pathology (gastrointestinal stromal tumors or urothelial carcinoma, n = 3), and incomplete data (n = 2). The theoretical number of high-grade PCa missed, number of biopsies saved, and number of diagnoses of low-grade PCa spared, at different risk thresholds to biopsy, were assessed.

Following the above exclusion criteria **(n = 236)**, 1309 patients were available for developing the new Taiwanese PCa risk model. Statistical analyses were performed using R and SPSS 18 software. The Kolmogorov-Smirnov test was utilized to assess normality of distribution of parameters. Continuous variables were presented as mean and standard deviation or median and interquartile range, based on their normal or non-normal distribution, respectively (assessed using Student *t*-tests and Mann-Whitney *U* tests). Categorical variables were assessed with the chi-square test. Univariate and multivariate analyses were conducted to determine the independent predictors of PCa and high-grade PCa. Logistic regression analyses of independent variables were performed to evaluate the probabilities of predicting PCa and high-grade PCa. **5-fold cross validation was employed to assess the stability and reliability of our new prediction model.**

The predictive performance of both PCa risk calculator apps and the new model in a Taiwanese population was statistically analyzed on the basis of its calibration, discrimination, and clinical usefulness [15]. Calibration was assessed graphically with a calibration plot, in which the predicted probabilities were plotted against the observed probabilities, enabling evaluation of the extent of risk over- or under-estimation [9]. Statistical significance of mis-calibration was determined by the Hosmer-Lemeshow goodness-of-fit test; the observed and predicted risks of PCa were compared across deciles of the predicted risk using a chi-square test, and a *P*-value >0.05 indicates good calibration. Discrimination was quantified using the area under the receiver operator characteristic curve (ROC). The areas under ROC curves (AUCs) of the PCa risk models were compared using DeLong's method [16]. For clinical usefulness, decision-curve analysis was implemented to determine which risk prediction model would lead to better decisions, by calculating the net benefit [17]. Moreover, based on different risk thresholds to biopsy, our model was compared with both apps to demonstrate the number and percentage of biopsies

saved versus low- and high- grade PCa missed in those who did not undergo a biopsy.

Results

Overall, 278/1309 (21.2%) patients were diagnosed with PCa, and 181 out of 278 (65.1%) patients had high-grade PCa (Gleason≥7). Comparing non-cancer, low-grade and high-grade PCa groups, patients with high-grade PCa were significantly older, having higher PSA levels, smaller PVs, more abnormal findings on TRUS and DRE, and higher PI-RADS scores on mpMRI **(Supplementary Table 1)**. In univariate and multivariate analysis, PSA, PV, TRUS and DRE abnormalities were independent significant predictors of any **and high-grade** PCa **(Supplementary Table 2)**. Our **5-fold cross-validated** logistic regression model was established based on the aforementioned predictive factors for PCa and high-grade PCa.

Both our model and the Rotterdam app demonstrated better discriminative ability than the Coral app for detection of PCa (AUC: **0.795** vs 0.792 vs 0.697; *P* <0.001) and high-grade PCa (AUC: 0.869 vs 0.873 vs 767; *P* <0.001) (Fig. 1 **and Supplementary Table 3**). No significant difference was shown between our model and the Rotterdam app for predicting PCa (*P* = **0.928**) and high-grade PCa (*P* = **0.683**). Calibration was plotted graphically and only the Rotterdam app revealed a good calibration (*P* >0.05) for detecting high-grade PCa. Our new model was well calibrated in the prediction of PCa and high-grade PCa **(Supplementary Fig. 1)**.

In decision-curve analysis, our model and the Rotterdam app provided clinical net benefits in the threshold probability range from **10-90% and** 10-85%, **respectively**; the Coral app demonstrated benefit from 15-95% for detecting any PCa. In the detection of high-grade PCa, our model demonstrated net benefits in the threshold probability range of 5-80%, the Rotterdam app 5-65%, and the Coral app 10-95%. When comparing these three models, the net benefit was greater for our model and the Rotterdam app in the prediction of any PCa, across the range of threshold probabilities from **10-85%**; and high-grade PCa, across the range from 5-65% **(Supplementary Fig. 2)**. Using different high-grade PCa risk thresholds to biopsy, the number and percentage of biopsies saved versus low- and high- grade PCa missed in those who did not undergo a biopsy are displayed in Table 1. **Our predictive nomograms and closed-form solutions for PCa and high-grade PCa are demonstrated in Supplementary Fig. 3.**

Discussion

Based upon the 2016 Cancer Registry Annual Report, PCa ranked as the fifth most common cancer and the seventh leading cause of cancer death in Taiwanese males. Although no PSA screening policy is currently implemented in Taiwan, ad-hoc screening is common in self-funded health examinations. PSA and DRE are traditionally used for predicting risk of PCa, but risk calculators have been recommended by the American Urological Association and EAU guidelines to further improve prediction and help determine what the likely risk group of PCa might be [10, 11, 18-20]. Moreover, PCa risk calculator apps have the advantages of worldwide accessibility, cost effectiveness and saving time.

In the present study, both mHealth apps were applicable to the Taiwanese cohort of patients, albeit they were developed from western populations. In comparison with our **cross-validated** model on discriminative performance, ours was similar to the Rotterdam app; both outperformed the Coral app for predicting PCa and high-grade PCa. A similar result was presented by Nunzio et al., who conducted the first external validation study for the PCa risk calculator apps [21]. They enrolled a two-center cohort of 1682 patients across Italy and Spain to validate the Rotterdam and Coral apps. In the prediction of any PCa on ROC analysis, the AUCs of the Rotterdam and Coral apps were 0.7 and 0.631; for high-grade PCa, the AUCs of both apps were 0.75 and 0.69, respectively. Although the Italian and Spanish patients might be ethnically similar to the population of whom data were utilized to develop these apps, the AUCs in our cohort from Asia were shown as good or even slightly better than those in the southern European cohort. Accordingly, comparable to the Rotterdam app, our model provided a fairly accurate discriminative capacity for predicting PCa and high-grade PCa.

With reference to calibration, our model demonstrated good calibration for both PCa and high-grade PCa. As for risk calculator apps, only the Rotterdam app showed good calibration in the prediction of high-grade PCa; the others resulted in poor calibration. He et al. [13] maintained that the ERSPC-RC tended to have better calibration and discriminative capacity than the PCPT-RC in the Chinese and Korean cohorts. They explained that this might be due to similarities between the East Asian population and the cohorts used to develop the ERSPC-RC. Similarly, one meta-analysis assessing the older versions of these two RCs demonstrated better performance of the ERSPC-RC [12]. Another independent validation study by Poyet et al. evaluated the updated versions of both RCs with a European Caucasian cohort and reported a slightly superior performance of the ERSPC-RC

[9].

The reason for the lower predictive performance of the PCPT-RC or the Coral app might be explained by the relatively low prevalence of PCa in the PCPT cohort. Since the PCPT-RC was generated from the placebo arm of a large PSA screening cohort, the substantial differences (e.g. only 3.8% diagnoses of high-grade PCa, or 13.9% of PSA levels before biopsy >4 ng/mL) might always exist when compared with contemporary referral populations [11]. Additionally, ethnic/racial variability might be one of the most leading causes of predictive inaccuracy and mis-calibration. Differences in clinical settings (screening or referral) and number of biopsy cores (sextant or 12-core systematic) should also be considered. On the other hand, clinical referral settings, systematic prostate biopsy strategy (10 or 12 cores), PV, TRUS findings, and MRI had been incorporated into the ERSPC-RCs, indicating their wider applicability in different ethnic/racial groups, including Asians. For instance, the ERSPC study group utilized the Hong Kong development and validation cohorts to develop a re-calibrated version of the ERSPC-RC3 for the Chinese population [22].

How to avoid missing clinically significant or high-grade PCa is pivotal to clinicians and patients. Using a ≥10% risk threshold for high-grade PCa to biopsy, our model could therefore save **67.2%** of total biopsies at the expense of missing 5.9% low-grade PCa and **3.4%** high-grade PCa. In comparison, the Rotterdam app could save 66.0% of total biopsies, missing 6.9% low-grade PCa and 3.0% high-grade PCa. In review of the original web-based versions of the Rotterdam app, Alberts et al. [23] utilized a two-country cohort of 961 patients to augment the predictive accuracy of the ERSPC-RC3 and ERSPC-RC4, by incorporating the MRI PI-RADS score and a larger age range. At a risk threshold to biopsy of ≥10% for high-grade PCa of the MRI-ERSPC-RC3 (for biopsy-naïve males) and MRI-ERSPC-RC4 (for previously biopsied males), 14% and 36% biopsies could be saved. However, they would miss 13% and 15% diagnoses of low-grade PCa, and 10% and 4% diagnoses of high-grade PCa, respectively. At the same ≥10% risk threshold, our model could save **67.2%** of biopsies and merely miss **3.4%** high-grade PCa. Hence, our new nomogram specifically developed for the Taiwanese population provided solid net benefits for reducing unnecessary biopsies and minimizing missed diagnoses of high-grade PCa.

There are several limitations in the present study. Firstly, it is a single-institution retrospective study. We need more cohorts from other Taiwanese hospitals to

validate our new model. Secondly, no data of prostate health index (PHI), which is one of the predictors listed in the Rotterdam app, is available in our institution. However, the additional merit of PHI to ERSPC RCs remains limited [24]. Thirdly, only 50 patients received mpMRI before biopsy in our cohort. More data on pre-biopsy mpMRI will be to refine our model further. Nevertheless, the predictor of the mpMRI used in the Rotterdam app was based on the PI-RADS v1 guidelines; our MRI images were graded according to the PI-RADS v2 scheme. Furthermore, the issues of inter-observer variability and heterogeneous definitions of abnormality in mpMRI interpretation remain to be tackled.

Conclusions

In conclusion, our new logistic regression model, similar to the Rotterdam app, outperformed the Coral app in the prediction of PCa and high-grade PCa. Additionally, our model could save unnecessary biopsies and avoid missing clinically significant PCa in the Taiwanese population.

Authors' contributions

IA Chen: Project development, Data Collection, Manuscript writing
C Chu: Data analysis
J Lin: Manuscript editing
J Tsai: Manuscript editing
C Yu: Manuscript editing
AN Sridhar: Manuscript editing
M Chand: Project development, Manuscript editing
P Sooriakumaran: Project development, Manuscript editing

Disclosure of potential conflicts of interest

The authors declared no potential conflicts of interest.

Research involving Human Participants and/or Animals

Our study was performed in accordance with the principles of the 1964 Declaration of Helsinki. Ethics approval was granted by the Internal Review Board of a Taiwanese tertiary medical center (IRB No.: VGHKS19-CT3-13).

Informed Consent

Owing to the retrospective nature of our study, the informed consent was waived.

References

1. Research2Guidance. mHealth App Economics: Current statuts and future trends in mobile health 2017 [Available from: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/.
2. World Health Organization. eHealth at WHO 2016 [Available from: https://www.who.int/ehealth/about/en/.
3. Statista Research Department. Number of smartphone users in Taiwan from 2015 to 2020 (in millions) [updated Dec 16, 2016. Available from: https://www.statista.com/statistics/755288/taiwan-number-of-smartphone-users/.
4. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
5. Ferlay J EM, Lam F, Colombet M, Mery L, Pineros M, Znaor A, et al. Global cancer observatory: cancer today. Lyon, France: International Agency for Research on Cancer. 2018 [updated 2019/07/28. Available from: https://gco.iarc.fr/today.
6. Health Promotion Administration. Taiwan Cancer Registry Annual Report, 2016 2018 [Available from: https://www.hpa.gov.tw/Pages/TopicList.aspx?nodeid=269&idx=0.
7. Chiang CJ, Lo WC, Yang YW, You SL, Chen CJ, Lai MS. Incidence and survival of adult cancer patients in Taiwan, 2002-2012. J Formos Med Assoc. 2016;115(12):1076-88.
8. Wolf AM, Wender RC, Etzioni RB, Thompson IM, D'Amico AV, Volk RJ, et al. American Cancer Society guideline for the early detection of prostate cancer: update 2010. CA Cancer J Clin. 2010;60(2):70-98.
9. Poyet C, Nieboer D, Bhindi B, Kulkarni GS, Wiederkehr C, Wettstein MS, et al. Prostate cancer risk prediction using the novel versions of the European Randomised Study for Screening of Prostate Cancer (ERSPC) and Prostate Cancer Prevention Trial (PCPT) risk calculators: independent validation and comparison in a contemporary European cohort. BJU Int. 2016;117(3):401-8.
10. Roobol MJ, Verbeek JFM, van der Kwast T, Kummerlin IP, Kweldam CF, van Leenders G. Improving the Rotterdam European Randomized Study of Screening for Prostate Cancer Risk Calculator for Initial Prostate Biopsy by Incorporating the 2014 International Society of Urological Pathology Gleason Grading and Cribriform growth. Eur Urol. 2017;72(1):45-51.
11. Ankerst DP, Hoefler J, Bock S, Goodman PJ, Vickers A, Hernandez J, et al. Prostate Cancer Prevention Trial risk calculator 2.0 for the prediction of low- vs high-grade prostate cancer. Urology. 2014;83(6):1362-7.

12. 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(5):848-64.
13. He BM, Chen R, Sun TQ, Yang Y, Zhang CL, Ren SC, et al. Prostate cancer risk prediction models in Eastern Asian populations: current status, racial difference, and future directions. Asian J Androl. 2019.
14. Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ, et al. PI-RADS Prostate Imaging - Reporting and Data System: 2015, Version 2. Eur Urol. 2016;69(1):16-40.
15. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology. 2010;21(1):128-38.
16. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837-45.
17. Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vickers AJ, et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. Eur Urol. 2018;74(6):796-804.
18. 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.
19. 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.
20. 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.
21. De Nunzio C, Lombardo R, Tema G, Cancrini F, Russo GI, Chacon R, et al. Mobile phone apps for the prediction of prostate cancer: External validation of the Coral and Rotterdam apps. Eur J Surg Oncol. 2019;45(3):471-6.
22. Chiu PK, Roobol MJ, Nieboer D, Teoh JY, Yuen SK, Hou SM, et al. Adaptation and external validation of the European randomised study of screening for prostate cancer risk calculator for the Chinese population. Prostate Cancer Prostatic Dis. 2017;20(1):99-104.
23. Alberts AR, Roobol MJ, Verbeek JFM, Schoots IG, Chiu PK, Osses DF, et al. Prediction of High-grade Prostate Cancer Following Multiparametric Magnetic

Resonance Imaging: Improving the Rotterdam European Randomized Study of Screening for Prostate Cancer Risk Calculators. Eur Urol. 2019;75(2):310-8.
24. 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):185-90.

Table 1. Biopsies saved versus low- and high- grade PCa missed among three models at different risk thresholds for high-grade PCa to biopsy.

Risk threshold for high-grade PCa		No of biopsies saved (% of total Bx)	No. of low-grade PCa missed (% of Bx saved)	No. of high-grade PCa missed (% of Bx saved)
5%	Rotterdam	574 (43.9%)	32 (5.6%)	12 (2.1%)
	Coral	75 (5.7%)	6 (8%)	5 (6.7%)
	Model	**538 (41.1%)**	**35 (6.5%)**	**8 (1.5%)**
10%	Rotterdam	864 (66.0%)	60 (6.9%)	26 (3.0%)
	Coral	490 (37.4%)	38 (7.8%)	20 (4.1%)
	Model	**880 (67.2%)**	**52 (5.9%)**	**30 (3.4%)**
15%	Rotterdam	978 (74.7%)	66 (6.7%)	41 (4.2%)
	Coral	843 (64.4%)	54 (6.4%)	58 (6.9%)
	Model	**983 (75.1%)**	**67 (6.8%)**	**48 (4.9%)**
20%	Rotterdam	1051 (80.7%)	70 (6.7%)	57 (5.4%)
	Coral	1033 (78.9%)	73 (7.1%)	81 (7.8%)
	Model	**1051 (80.7%)**	**68 (6.5%)**	**64 (6.1%)**

PCa **High-grade PCa**

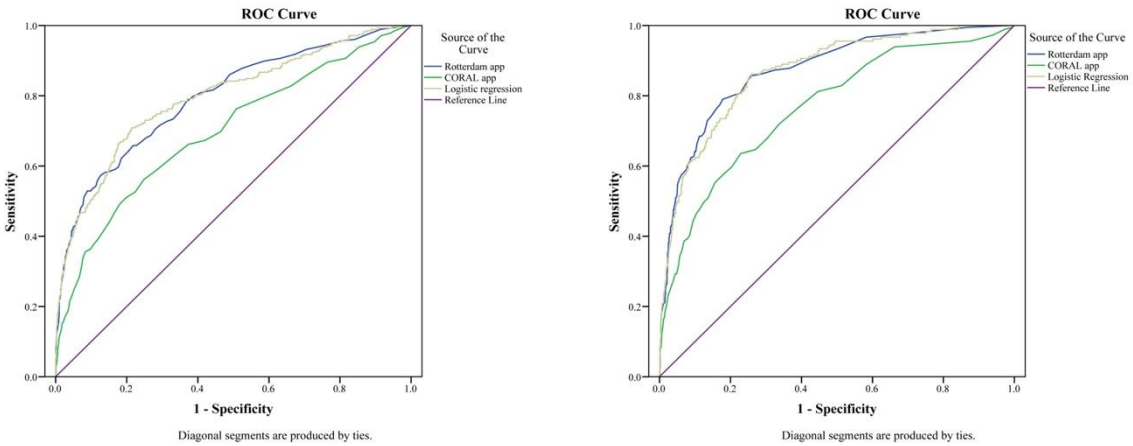

Fig. 1 ROC curves for the discriminative ability of the Rotterdam and Coral apps and our new model. (**Left**) PCa vs no PCa; (**Right**) High-grade PCa vs Low-grade plus no PCa.