Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Eur Urol. Author manuscript; available in PMC 2016 January 01.

Published in final edited form as:

Eur Urol. 2015 July ; 68(1): 139–146. doi:10.1016/j.eururo.2014.08.010.

Comparison between the four-kallikrein panel and Prostate Health Index (PHI) for predicting prostate cancer

Tobias Nordström, M.D.[a,b,*], **Andrew Vickers, Ph. D.**[c], **Melissa Assel, Ph. D.**[c], **Hans Lilja, M.D. Ph.D.**[d,e,f] **[Professor]**, **Henrik Grönberg, M.D.**[b] **[Professor]**, and **Martin Eklund, Ph. D.**[b,g]

[a]Department of Clinical Sciences at Danderyd's Hospital, Karolinska Institutet S-182 88 Stockholm, Sweden

[b]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet S-171 77 Stockholm, Sweden

[c]Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY, USA

[d]Departments of Laboratory Medicine, Surgery (Urology), and Medicine (Genitourinary Oncology), Memorial Sloan-Kettering Cancer Center, New York, NY, USA

[e]Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK

[f]Department of Laboratory Medicine, Lund University, Skåne University Hospital, Malmö, Sweden

[g]Department of Surgery, University of California at San Francisco, San Francisco CA 94115, USA

Abstract

BACKGROUND—The four-kallikrein panel and the Prostate Health Index (PHI) have been shown to improve prediction of prostate cancer compared to prostate-specific antigen (PSA). No comparison between the four-kallikrein panel and PHI has been presented.

OBJECTIVE—To compare the four-kallikrein panel to PHI for predicting prostate cancer in an independent cohort.

DESIGN, SETTING AND PARTICIPANTS—Participants were from a population-based cohort of PSA-tested men in Stockholm County. We included men (n=531) with PSA 3–15 ng/ml undergoing first-time prostate biopsy during 2010–2012.

OUTCOME MEASUREMENTS AND STATISTICAL ANALYSIS—Models were fitted to case status. We computed calibration curves, the area under the receiver-operating characteristics curve (AUC), decision curves, and percent saved biopsies.

RESULTS AND LIMITATIONS—The four-kallikrein panel showed AUCs of 69.0 when predicting any-grade prostate cancer and 71.8 when predicting high-grade cancer (≥Gleason 7).

[*]Corresponding author: Tobias Nordström, tobias.nordstrom@ki.se, Telephone: +46 (0) 70 5391791; Fax: +46 (8) 314975.

CONFLICT OF INTEREST

Hans Lilja holds patents for free PSA, Human Kallikrein 2, and intact PSA assays, and is named, along with Andrew Vickers, on a patent application for a statistical method to detect prostate cancer. Andrew Vickers is a paid consultant to Opko, which plans to commercialize the test based on the four kallikreins.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Similar values were found for PHI: 70.4 and 71.1, respectively. Both models had higher AUC than a base model with PSA and age ($p<0.0001$ for both); differences between models were not significant. Sensitivity analyses including men with any PSA or a previous biopsy did not materially affect our findings. Using 10% predicted risk of high-grade prostate cancer by the four-kallikrein panel or PHI=39 as cutoff for biopsy saves 29% of performed biopsies to the cost of delayed diagnosis for 10% of the men with high-grade cancers. Both models showed limited net benefit in decision analysis. The main study limitation was lack of digital rectal examination data and biopsy decision being based on PSA information.

CONCLUSIONS—The four-kallikrein panel and PHI similarly improved discrimination when predicting prostate cancer and high-grade prostate cancer. Both are simple blood tests that can reduce the number of unnecessary biopsies compared to screening with total PSA, representing an important new option to reduce harm.

Keywords

Prostatic neoplasms; Biomarkers; Prostate-specific antigen; Kallikrein-related peptidases

INTRODUCTION

Numerous studies have addressed the limited diagnostic accuracy of PSA and yet it remains the only widely adopted biomarker for prostate cancer[1–4].

Although sensitivity to detect increased risk of metastasis or death from prostate cancer many years later may be unique[5], insufficient test specificity drives frequent prostate biopsying and a large proportion of benign biopsies. Several additional biomarkers have been suggested, none of which have reached widespread clinical use.

However, a few blood-based biomarkers have proved promising. Prostate Health Index is an algorithm including the PSA isoform [-2]proPSA with total and free PSA. It has been demonstrated to increase predictive performance in several ethnically diverse cohorts for predicting prostate cancer at biopsy or radical prostatectomy specimens[6–11].

Several other human kallikrein-related peptidases have been explored and a four-kallikrein panel including kallikrein-related peptidase 2 (hK2), intact PSA, free and total PSA have repeatedly been shown to predict prostate biopsy outcome in primarily European men with an elevated PSA and to save a substantial number of biopsies[12–16].

The four-kallikrein panel and PHI represent improved tests for prostate cancer that potentially can be of widespread clinical use. However, no study has compared the performances of these tests.

MATERIAL AND METHODS

Study design

This observational, prospectively collected study included men with blood samples drawn before a prostate biopsy resulting in cancer diagnosis (cases) or benign findings (controls). The study was designed to compare the diagnostic performance for predicting prostate

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

cancer using a base model containing total PSA and age, PHI, and the four-kallikrein panel, respectively.

Study population

Men referred to PSA-testing in laboratories in Stockholm County between 2010 and 2012 were invited to the population-based cohort STHLM2 at the blood sampling visit. A total of 26,712 men were included during the 22-month study period.

We selected all new prostate cancer cases in STHLM2 reported to the National Prostate Cancer Register and all men having a biopsy with benign findings reported after inclusion and before June 20, 2012. In the main analysis, only previously unbiopsied men with PSA 3–15 ng/ml were included. Biopsies were decided on according to clinical practice, including information on PSA levels, DRE findings, prostate volume and family history. Biopsies were 10–12 core ultrasound-guided biopsies.

Laboratory analysis

Whole blood for plasma analysis was collected in separate ethylenediaminetetraacetic acid (EDTA) tubes without gel. Study samples were drawn at the local laboratory and transported to KI Biobank, Karolinska Institutet within 24 hours. After centrifugation plasma was aliquotted and stored at −80°C.

PSA and free PSA were analyzed using Roche Modular E170 and [-2]proPSA was analyzed using UniCel DxI800 Immunoassay System analyzer (Beckman & Coulter, Brea, CA, USA), all at Karolinska University Hospital. Another aliquot of cryopreserved plasma was used to measure total and free PSA with the dual-label DELFIA Prostatus assay (Perkin-Elmer, Finland), calibrated against WHO 96/670 and WHO 68/668 standards, hK2 and intact PSA at Wallenberg Research Laboratories, Department of Laboratory Medicine, Skåne University Hospital as previously described[17]. All analyses were performed blinded to biopsy result.

Register data

Historical PSA-test data, biopsy records and prostate cancer records were retrieved from the continuously updated STHLM0 database[18]. Briefly, this database consists of all men in Stockholm County who had at least one PSA analyzed since 2003 together with their biopsy records and prostate cancer history as recorded by the National Prostate Cancer Register. Pathology reports on all prostate biopsy results were retrieved. Participant data were matched against the National Cancer Registry and National Prostate Cancer Registry, both with coverage >93%, to obtain cancer status and clinical information[19].

Statistical methods

The base model contained age and total PSA and was modeled using logistic regression. The PHI score is calculated as [-2]proPSA/freePSA*√totalPSA[6]. The four-kallikrein panel was modeled using restricted cubic splines for total and free PSA to model possible non-linear relationships, as described previously[13]. The base- and 4K models were fitted to data from the UK ProtecT study. Separate models were created for the endpoint of prostate cancer and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

high-grade cancer. These models were then used to predict the risks in the independent cohort collected within this study, thus providing entirely external assessments of all models.

Digital rectal exam (DRE) data was not available for the study cohort. High-grade cancer was defined as Gleason score ≥7. When assessing high-grade cancer, low-grade cancers were classified as controls (i.e. men with low-grade cancers were grouped together with men with benign biopsies).

Since men with total PSA levels exceeding 15 ng/ml will almost certainly be biopsied regardless of the blood concentrations of the other markers, we restricted the dataset to include only men with total PSA in the range 3–15 ng/ml for our main analyses. We performed sensitivity analyses on data restricted to PSA 2–10 as well as unrestricted for PSA (Appendix).

Since both the 4K and PHI models were developed for predicting results of a first-time biopsy, we included only biopsy naïve men in the main analysis. We present results where previously biopsied men were included in the Appendix.

No imputation on missing data was performed; we used a missing completely at random assumption. No significant deviations from this assumption were observed when testing for associations between missing values in the biomarkers and having a positive biopsy.

Predicted discrimination was visualized using receiver operating characteristics (ROC) curves and summarized using the area under the ROC curve (AUC). Confidence intervals and p-values when comparing AUC were calculated using the method of DeLong et al[20]. All p-values were based on two-sided hypotheses.

We used decision curve analysis to evaluate if the models improved accuracy and subsequent clinical management of patients[21]. Briefly, decision curve analysis graphically illustrates the net benefit obtained by using the predictive models in a patient by assuming that the threshold probability for having all prostate cancer or highgrade prostate cancer at which a patient would opt for biopsy is informative of how the patient weighs the relative harms of a false-positive and a false-negative prediction. This relationship is used to derive the net benefit of the model across different threshold probabilities. Plotting net benefit against threshold probability yields the “decision curve.”

Percent saved biopsies for one model in comparison to another was computed according to $(TP_1+FP_1-TP_2-FP_2)/(TP_1+FP_1)$, where TP_1, FP_1, TP_2, and FP_2 denote the true and false positives for the two models.

We calculated calibration plots for 4K by plotting predicted risk versus observed proportions of cases for strata of the predicted risk. The PHI equation represents a simplified logistic regression model developed at Beckman Coulter[22] and is not formulated as a risk estimation. To assess the PHI score’s relation to risk in this particular dataset, we plotted PHI againstobserved proportions of cases in 10 strata of the PHI score.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Statistical analysis was performed using Stata 11.2 (Stata Corp, College Station, TX) and R[23].

RESULTS

Clinical characteristics of the study population are listed in Table 1. 531 previously unbiopsied participants with PSA 3–15ng/ml were identified (271 cases and 260 controls). 134 (49%) of the cases were characterized as Gleason grade ≥7. Cases were older than controls (mean age 66.0 vs. 64.6, p=0.04). Concentrations of free PSA and [-2]proPSA differed significantly between groups (p<0.001; p=0.038).

Performance comparison

The predictive performance of the different models applied to unbiopsied men with PSA 3–15ng/ml is shown in Table 2. For predicting high-grade prostate cancer, the four-kallikrein panel and PHI had AUCs of 71.8 (95% CI: 66.8–76.7) and 71.1 (66.0–76.2) respectively, both significantly higher than the base model (p<0.0001). For predicting all prostate cancer, the base model containing age and total PSA had an AUC of 54.5 (95% CI: 49.6–59.4). The AUC increased significantly to 69.0 (95% CI: 64.5–73.4) for the four-kallikrein panel, the corresponding AUC for PHI was 70.4 (95% CI: 66.1–74.8). Figure 2 shows the ROC-curves for predicting all prostate cancer and high-grade disease.

AUC of the PHI model did not differ significantly from the AUC of the four- kallikrein panel model when predicting high-grade cancer (p=0.77) or all cancer (p=0.52).

Relevant biomarkers in prostate cancer diagnostics should perform well in high sensitivity ranges. Thus, we explored the partial AUC in the sensitivity range 75–100% and there was no significant difference between the models when sensitivity exceeds 75% (Figure 2, shaded area; Table 2). .

The performances of the respective models in the datasets including previously biopsied men and men with any PSA were similar to the performances in the main analysis (Appendix Table A1–2). There was no significant difference when predicting high-grade disease. However, in these datasets, the AUC of PHI for predicting all cancer was higher than the four- kallikrein panel AUC (p=0.04 and p=0.02, respectively).

Saved biopsies, decision curve analysis, and calibration

Figure 3 shows the decision curves when predicting all and highgrade disease. The four-kallikrein panel showed net benefit when the cutoff for biopsy exceeds 18% risk for all cancer and 8% risk for high-grade disease. The clinical utility of PHI was strongly dependent on the cut-off used; in this dataset we found that PHI cutoffs 25–30 had poor properties. However, cutoffs ranging from 30–40 did have clinical utility.

The four-kallikrein panel was well calibrated for predicting all prostate cancer and highgrade cancer (Appendix Figures A1–2). PHI was poorly calibrated to previously reported associations between PHI and observed risk for all cancer[24] and PHI cutoffs of 25–30 corresponded to very low risks of high-grade disease (Appendix Figures A1–2).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Both the four-kallikrein panel and PHI save biopsies, but to a cost of missing high-grade cancers (Table 3). The models save almost 30% of the biopsies to the cost of missing 10% highgrade cancers if using 10% risk of highgrade cancer as predicted by the four-kallikrein panel or PHI=39 as cutoff for biopsy.

DISCUSSION

We have compared the performance of two predictive models – Prostate Health Index (PHI) and a four kallikrein-panel – to each other and a base model when applied to previously unbiopsied men with intermediately increased PSA. We find that both PHI and the four-kallikrein panel increase predictive accuracy compared to the base model and that the PHI-model and the four-kallikrein panel perform similarly for predicting both high-grade disease and all cancer.

Lilja and colleagues have repeatedly demonstrated that the four-kallikrein panel improves on PSA, potentially saving approximately half of the biopsies to the cost of missing about every tenth to twentieth of the high-grade cancers if using a fixed risk-cutoff for biopsy at 20% all-cancer risk. The AUC of the four-kallikrein panel when predicting outcome after first biopsy has been shown to be 0.71–0.76[12–16]. We validate these results in an independent cohort, where a four-kallikrein panel had an AUC for all prostate cancer of 0.69 in unbiopsied men with PSA 3–15ng/ml and were well calibrated for predicting prostate cancer. Using a 15% risk for high-grade cancer as cutoff, the model saves 44% of the biopsies to the cost of missing approximately 20% (29/134) high grade cancers. The fraction of missed high-grade cancers among the saved biopsies was 13.3% and 18.1% for the 15% and 20% risk cutoffs, respectively. It is unclear whether the somewhat poorer results of the 4K panel than previously reported are an effect of the higher prevalence of high grade cancers in the current dataset or the lack of DRE information.

The FDA-approved PHI-index is the commercially available application of [-2]proPSA measurements. There is evidence that PHI-models increases discriminative accuracy between patients with and without prostate cancer, especially in the PSA-range 2–10ng/ml[6]. Lazzeri and colleges recently showed that a PHI-based model yielded an AUC of 0.71 (0.67–0.75) for predicting cancer in first biopsy, compared to a base model's AUC of 0.65 (0.61–0.69)[10,11]. We find that the PHI-model adds predictive performance to the base model, showing an AUC of 0.70 and 0.71 when predicting all and high-grade cancer. The model saves biopsies and misses high-grade cancers in the same range as the four-kallikrein panel. In high sensitivity ranges, PHI seems to discriminate somewhat better than the four-kallikrein panel for predicting all prostate cancer (Figure 2A). In this dataset, commonly used cutoffs (PHI 25–30) correspond to low risk of high-grade disease and were of limited clinical value. Again, this may be because PHI is intended for use only in patients with a negative DRE. PHI-values between 35–40 had better properties as illustrated by given example of PHI=39 (Figure 3, Table 3), also being in line with recent guidelines from the National Comprehensive Cancer Network[25].

We find that the four-kallikrein panel and the PHI-model perform comparably well for predicting high-grade prostate cancer. Including previously biopsied men or men with PSA

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

outside the 3–15ng/ml range did not alter performance materially for either 4K or PHI, indicating robustness of the models. From a clinical point of view this is reassuring since previous biopsy status may not always be known and since the models potentially are of value also for deciding on prostate biopsy in men with PSA outside the 3–15 range.

Our study has several strengths. First, we use a contemporary, well-characterized cohort with men giving blood for PSA-analysis. From this cohort we included all men who had a first-time prostate biopsy with cancer or benign findings. This means that our results reflect a relevant, contemporary clinical setting. Second, we use EDTA-plasma throughout the laboratory analyses, providing stable and easy-to-handle samples. Third, we validate the results of previous studies of biomarker models in an entirely independent cohort, giving our study external validity. Fourth, we explore biomarker models readily incorporated in clinical practice, lending clinical applicability. Fifth, we demonstrate improvement of clinical decision-making by saved biopsies when introducing the biomarker models. From a pragmatic viewpoint, all explored models are potential alternatives in a biopsy decision situation. Being simple blood-tests, both models offer logistical advantages compared to urine tests demanding DRE or tissue based tests.

However, our study is not devoid of limitations, the most prominent being that (i), all biopsies were based on PSA and clinical information (this limitation is in fact shared with virtually all studies of prostate cancer, being ethically problematic to biopsy a large group of men with PSA < 3); (ii), we had no access to DRE data, prostate volume, or family history, which often are considered in a biopsy decision. PHI was developed primarily for men with benign DRE and the commercialized four-kallikrein model includes DRE information. The effect of this could be that this study underestimates the discrimination of the four-kallikrein panel and the calibration of PHI. (iii) Since free-to-total PSA ratio is quite commonly used in Sweden, we would expect biomarker models including this variable (i.e. both PHI and the four-kallikrein panel) to do relatively less well. (iv) the study population is ethnically homogenous, and (v), Gleason scores were missing for some of the participants, possibly diluting any differences between the models in predicting high-grade disease.

CONCLUSIONS

The four-kallikrein panel and PHI similarly improved discrimination when predicting prostate cancer and high-grade prostate cancer. Both are simple blood tests that can reduce the number of unnecessary biopsies compared to screening with total PSA, representing an important new option to reduce harm.

Acknowledgments

FUNDING

This study was supported by grants from The Strategic Research Programme on Cancer (StratCan), Karolinska Institutet; the Linné Centre for Breast and Prostate Cancer (CRISP, number 70867901), Karolinska Institutet; The Swedish Research Council (number K2010-70X-20430-04-3) and The Swedish Cancer Society (numbers 11-0287 and 11-0624); Stiftelsen Johanna Hagstrand och Sigfrid Linnérs minne; Swedish Council for Working Life and Social Research (FAS), number 2012-0073; National Cancer Institute (R33CA127768-03, R01CA160816 and P50-CA92629); the Sidney Kimmel Center for Prostate and Urologic Cancers; David H. Koch through the Prostate Cancer Foundation; the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Program; and Fundacion Federico SA. The funding source had no role in the study design; collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. The researchers were all independent from the funding source.

Karolinska University Laboratory, Aleris Medilab, Unilabs and the Regional Prostate Cancer Registry for performing analyses and help to retrieve data. Carin Cavalli–Björkman and Britt-Marie Hune for their enthusiastic work as research nurses. Astrid Björklund, our data administrator, who always in spite of short notice, helps us producing correct datasets. We wish to thank the BBMRI.se biobank facility at Karolinska Institutet for biobank services. We also thank Gun-Britt Eriksson, Mona Hassan Al-Battat, and AnnaPeri Erlandsson for expert assistance with immunoassay measurements of the blood samples at the Wallenberg Research Laboratories at Lund University in Malmö, Sweden.

References

1. Oesterling JE, Martin SK, Bergstralh EJ, Lowe FC. The use of prostate-specific antigen in staging patients with newly diagnosed prostate cancer. Jama. 1993; 269:57–60. [PubMed: 7677962]
2. Anderson JR, Strickland D, Corbin D, Byrnes JA, Zweiback E. Age-specific reference ranges for serum prostate-specific antigen. Urology. 1995; 46:54–7. [PubMed: 7541586]
3. Jacobsen SJ, Bergstralh EJ, Guess HA, Katusic SK, Klee GG, Oesterling JE, et al. Predictive properties of serum-prostate-specific antigen testing in a community-based setting. Arch Intern Med. 1996; 156:2462–8. [PubMed: 8944739]
4. Harvey P, Basuita A, Endersby D, Curtis B, Iacovidou A, Walker M. A systematic review of the diagnostic accuracy of prostate specific antigen. BMC Urol. 2009; 9:14. [PubMed: 19744336]
5. Vickers AJ, Ulmert D, Sjoberg DD, Bennette CJ, Björk T, Gerdtsson A, et al. Strategy for detection of prostate cancer based on relation between prostate specific antigen at age 40–55 and long term risk of metastasis: case-control study. Bmj. 2013; 346:f2023. [PubMed: 23596126]
6. Hori S, Blanchet J-S, McLoughlin J. From prostate-specific antigen (PSA) to precursor PSA (proPSA) isoforms: a review of the emerging role of proPSAs in the detection and management of early prostate cancer. BJU Int. 2013; 112:717–28. [PubMed: 22759214]
7. Catalona WJ, Partin AW, Sanda MG, Wei JT, Klee GG, Bangma CH, et al. A multicenter study of [-2]pro-prostate specific antigen combined with prostate specific antigen and free prostate specific antigen for prostate cancer detection in the 2.0 to 10.0 ng/ml prostate specific antigen range. J Urol. 2011; 185:1650–5. [PubMed: 21419439]
8. Stephan CC, Vincendeau SS, Houlgatte AA, Cammann HH, Jung KK, Semjonow AA. Multicenter evaluation of [-2]proprostate-specific antigen and the prostate health index for detecting prostate cancer. Clin Chem. 2012; 59:306–14. [PubMed: 23213080]
9. Guazzoni G, Lazzeri M, Nava L, Lughezzani G, Larcher A, Scattoni V, et al. Preoperative Prostate-Specific Antigen Isoform p2PSA and Its Derivatives, %p2PSA and Prostate Health Index, Predict Pathologic Outcomes in Patients Undergoing Radical Prostatectomy for Prostate Cancer. Eur Urol. 2012; 61:455–66. [PubMed: 22078333]
10. Guazzoni G, Nava L, Lazzeri M, Scattoni V, Lughezzani G, Maccagnano C, et al. Prostate-specific antigen (PSA) isoform p2PSA significantly improves the prediction of prostate cancer at initial extended prostate biopsies in patients with total PSA between 2.0 and 10 ng/ml: results of a prospective study in a clinical setting. Eur Urol. 2011; 60:214–22. [PubMed: 21482022]
11. Lazzeri M, Haese A, la Taille de A, Palou Redorta J, McNicholas T, Lughezzani G, et al. Serum Isoform [-2]proPSA Derivatives Significantly Improve Prediction of Prostate Cancer at Initial Biopsy in a Total PSA Range of 2–10 ng/ml: A Multicentric European Study. Eur Urol. 2013; 63:986–94. [PubMed: 23375961]
12. Vickers A, Cronin A, Roobol M, Savage C, Peltola M, Pettersson K, et al. Reducing unnecessary biopsy during prostate cancer screening using a four-kallikrein panel: an independent replication. J Clin Oncol. 2010; 28:2493–8. [PubMed: 20421547]
13. Vickers AJ, Cronin AM, Roobol MJ, Savage CJ, Peltola M, Pettersson K, et al. A four-kallikrein panel predicts prostate cancer in men with recent screening: data from the European Randomized Study of Screening for Prostate Cancer, Rotterdam. Clin Cancer Res. 2010; 16:3232–9. [PubMed: 20400522]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

14. Vickers AJ, Gupta A, Savage CJ, Pettersson K, Dahlin A, Bjartell A, et al. A panel of kallikrein marker predicts prostate cancer in a large, population-based cohort followed for 15 years without screening. Cancer Epidemiol Biomarkers Prev. 2011; 20:255–61. [PubMed: 21148123]
15. Vickers AJ, Cronin AM, Aus G, Pihl C-G, Becker C, Pettersson K, et al. A panel of kallikrein markers can reduce unnecessary biopsy for prostate cancer: data from the European Randomized Study of Prostate Cancer Screening in Göteborg, Sweden. BMC Medicine. 2008; 6:19. [PubMed: 18611265]
16. Benchikh A, Savage C, Cronin A, Salama G, Villers A, Lilja H, et al. A panel of kallikrein markers can predict outcome of prostate biopsy following clinical work-up: an independent validation study from the European Randomized Study of Prostate Cancer screening, France. BMC Cancer. 2010; 10:635. [PubMed: 21092177]
17. Väisänen V, Peltola MT, Lilja H, Nurmi M, Pettersson K. Intact Free Prostate-Specific Antigen and Free and Total Human Glandular Kallikrein 2. Elimination of Assay Interference by Enzymatic Digestion of Antibodies to F(ab') 2Fragments. Anal Chem. 2006; 78:7809–15. [PubMed: 17105175]
18. Nordström T, Aly M, Clements MS, Weibull CE, Adolfsson J, Grönberg H. Prostate-specific Antigen (PSA) Testing Is Prevalent and Increasing in Stockholm County, Sweden, Despite No Recommendations for PSA Screening: Results from a Population-based Study, 2003–2011. Eur Urol. 2013; 63:419–25. [PubMed: 23083803]
19. Barlow L, Westergren K, Holmberg L, Talback M. The completeness of the Swedish Cancer Register: a sample survey for year 1998. Acta Oncol. 2009; 48:27–33. [PubMed: 18767000]
20. DeLong ERE, DeLong DMD, Clarke-Pearson DLD. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988; 44:837–45. [PubMed: 3203132]
21. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006; 26:565–74. [PubMed: 17099194]
22. Jansen FH, van Schaik RH, Kurstjens J, Horninger W, Klocker H, Bektic J, et al. Prostate-specific antigen (PSA) isoform p2PSA in combination with total PSA and free PSA improves diagnostic accuracy in prostate cancer detection. Eur Urol. 2010; 57:921–7. [PubMed: 20189711]
23. Team RDC. R: A language and environment for statistical computing. R Foundation for Statistical Computing; n.d. http://wwwR-projectorg/
24. BeckmanCoulter. Hybritech p2PSA Draft Directional Insert. U.S. Food and Drug Administration; 2011.
25. Carroll, P.; Kellogg Parsons, J.; Andriole, GL.; Bahnson, RR.; Barocas, DA.; Catalona, WJ., et al. Prostate Cancer Early Detection. National Comprehensive Cancer Network; 2014. p. 1-35.

APPENDIX

Table A1

Performance of models predicting prostate cancer when applied to a cohort of 897 men of which 707 were not previously biopsied (455 controls and 442 prostate cancer cases of which 221 with high-grade cancer).

	All cancer			High-grade cancer		
Model	**AUC**	**CI (95%)**	**p (vs base)**	**AUC**	**CI (95%)**	**p (vs base)**
Base model	62.8	(59.2–66.4)	-	68.4	(64.3–72.5)	-
4K-panel*	74.9	(7.17–78.1)	<0.0001	78.5	(75.1–82.0)	<0.0001
PHI	77.5	(74.5–80.5)	<0.0001	79.0	(75.5–82.5)	<0.0001
AUC comparison: PHI vs 4K-panel			0.22			0.81

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The base model includes age and total PSA. 4K-panel includes age, total PSA, free PSA, human Kallikrein 2 and intact PSA. PHI includes total PSA, free PSA and [-2]proPSA. Cancers with Gleason ≥7 defined high-grade cancer. AUC = area under the curve; PSA = prostate-specific antigen

* The 4K-panel risks for patients who underwent a previous biopsy were recalibrated using a 10-fold cross validated Bayes factor.

Table A2

Performance of models predicting prostate cancer when applied to a cohort of 691 men with PSA 3–15 ng/ml of which 531 were not previously biopsied (370 controls and 321 prostate cancer cases of which 156 with high-grade cancer).

	All cancer			High-grade cancer		
Model	**AUC**	**CI (95%)**	**p (vs base)**	**AUC**	**CI (95%)**	**p (vs base)**
Base model	52.0	(47.7–56.3)	-	56.7	(51.6–61.8)	-
4K-panel*	71.1	(67.3–75.0)	<0.0001	73.2	(68.8–77.5)	<0.0001
PHI	72.2	(68.5–76.0)	<0.0001	71.8	(67.2–76.5)	<0.0001
AUC comparison: PHI vs 4K-panel			0.59			0.52

* The 4K-panel risks for patients who underwent a previous biopsy were recalibrated using a 10-fold cross validated Bayes factor.

Table A3

Performance of models predicting prostate cancer when applied to a cohort of 521 previously unbiopsied men with PSA 2–10 ng/ml (269 controls and 252 prostate cancer cases of which 117 with high-grade cancer). The most commonly used PSA cut-off for biopsy in Sweden is 3ng/ml. It should therefore be noted that the men with PSA between 2 and 3 may represent an atypical population.

	All cancer			High-grade cancer		
Model	**AUC**	**CI (95%)**	**p (vs base)**	**AUC**	**CI (95%)**	**p (vs base)**
Base model	55.2	(50.3–60.2)	-	59.9	(54.4–65.5)	-
Four-kallikrein panel	69.1	(64.6–73.6)	<0.0001	70.5	(65.3–75.7)	<0.0001
PHI	69.6	(65.1–74.1)	<0.0001	69.9	(64.5–75.3)	<0.0001
AUC comparison: PHI vs four-kallikrein panel			0.84			0.81

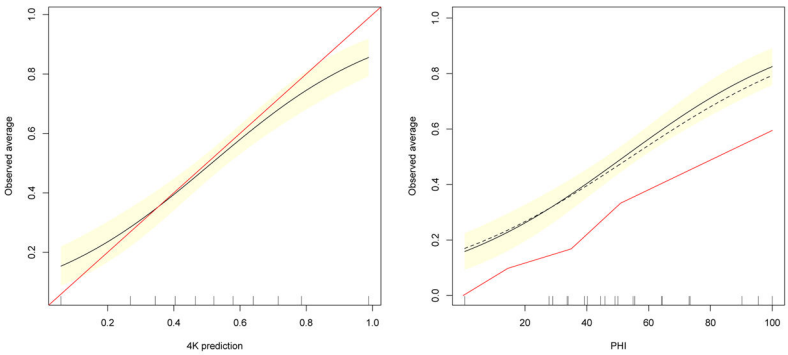

Figure A1.

Left panel. Calibration plot when using 4K to predict risk for all cancer in the STHLM2 cohort of 531 previously unbiopsied men with PSA 3–15 ng/ml. Yellow region shows 95%

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

confidence interval. The red line illustrates perfect calibration. Right panel. Solid black line: Association between PHI and observed risk for all cancer in the STHLM2 cohort of 531 previously unbiopsied men with PSA 3–15 ng/ml. Yellow region shows 95% confidence interval. Dashed black line: Association between PHI and observed risk for all cancer in the STHLM2 cohort of 397 previously unbiopsied men with PSA 4–10 ng/ml. Red line: Previously reported association between PHI and observed risk in a cohort of 445 men with PSA 4–10 ng/ml[27].

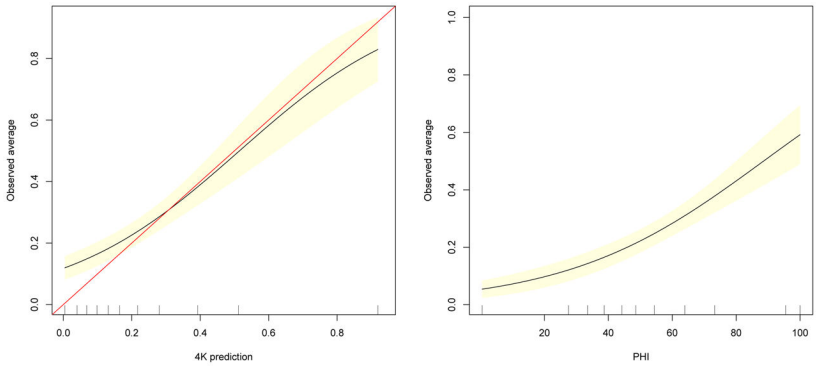

Figure A2.

Left panel: Calibration plot when using 4K to predict risk for highgrade cancer in the STHLM2 cohort of 531 previously unbiopsied men with PSA 3–15 ng/ml. Yellow region shows 95% confidence interval. The red line illustrates perfect calibration. Right panel: Association between PHI and observed risk for highgrade cancer in the STHLM2 cohort of 531 previously unbiopsied men with PSA 3–15 ng/ml. Yellow region shows 95% confidence interval.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

PATIENT SUMMARY

PSA screening is controversial due to limitations of the test. We found that two blood tests, PHI and the four-kallikrein panel, performed similarly and could both aid in decision-making in Swedish men undergoing a prostate biopsy.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

TAKE HOME MESSAGE

Two new blood tests called the four-kallikrein panel and PHI similarly improved discrimination in predicting any-grade and high-grade prostate cancer.

Compared to screening with PSA, these new markers may help reduce harms like unnecessary biopsies and overdetection.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

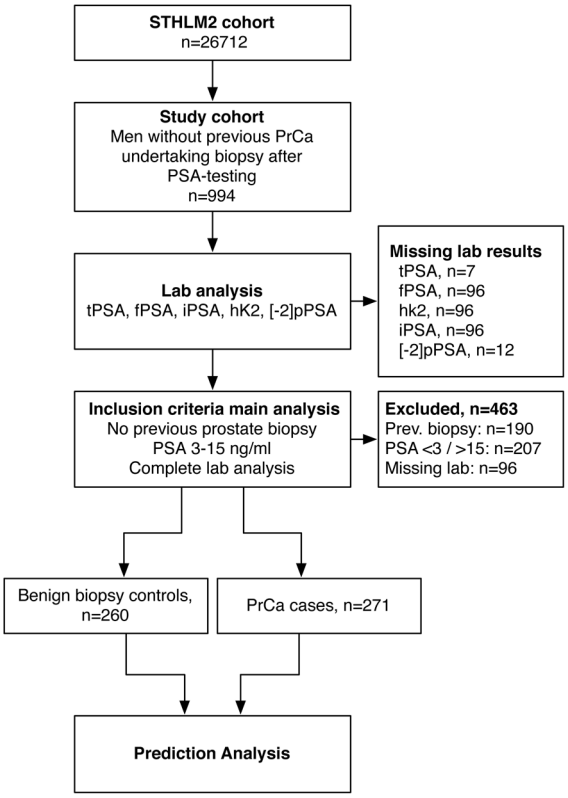

Figure 1.
Flowchart of study design

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

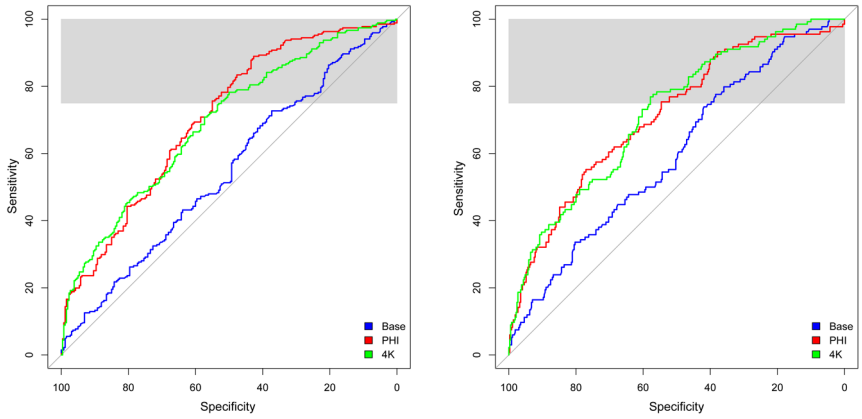

Figure 2.

Receiver operating characteristics (ROC) curves depicting the accuracy of five models in predicting prostate cancer (left) and high-grade prostate cancer (Gleason ≥7; right) in 531 previously unbiopsied men with PSA 3–15 ng/ml (260 controls and 271 prostate cancer cases of which 134 with high-grade cancer).

The base model includes age and total PSA. The four-kallikrein panel model includes age, total PSA, free PSA, human Kallikrein 2 and intact PSA. PHI includes total PSA, free PSA and [-2]proPSA. Shaded area marks sensitivity above 75%. PSA = prostate-specific antigen.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

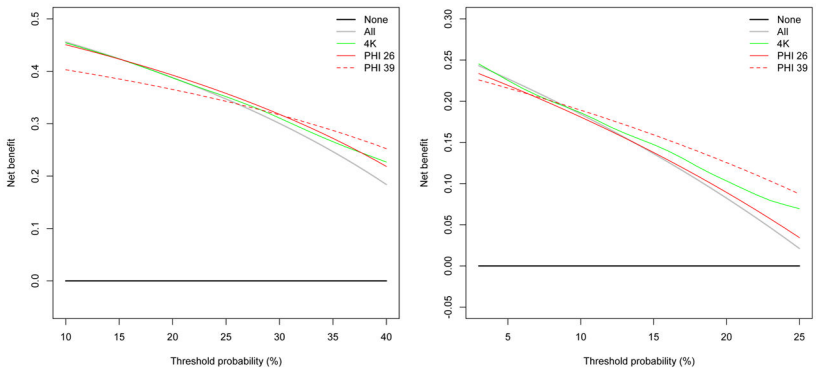

Figure 3.

Decision curve analysis showing the effect of the four-kallikrein panel and PHI on the detection of prostate cancer and high-grade prostate cancer. Clinical net benefit for the models is plotted against the risk threshold at which a patient or clinician would opt for biopsy. As comparison, the gray line represents the strategy of biopsying all men, and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1

Characteristics among a cohort of previously unbiopsied men coming for PSA-testing. All blood samples were retrieved before biopsy and diagnosis. Results are shown with and without restriction to participants with total PSA 3–15 ng/ml.

	Men with PSA 3–15 ng/ml			**Complete dataset**		
	Prostate Cancer	**Benign biopsy**	**p**	**Prostate Cancer**	**Benign biopsy**	**p**
n	271	260		375	332	
Age, yrs	66.0 (9.0)	64.6 (9.0)	0.041	67.8 (8.6)	64.8 (8.5)	<0.001
Total PSA, ng/ml	6.0 (4.1)	5.8 (3.5)	0.14	6.7 (8.4)	5.2 (4.2)	<0.001
Free PSA, ng/ml	0.86 (0.59)	1.0 (0.90)	<0.001	0.96 (1.1)	0.89 (0.96)	0.36
[-2]proPSA, ng/ml	0.019 (0.016)	0.018 (0.015)	0.038	0.022 (0.031)	0.016 (0.016)	<0.001
Prostate Health Index, PHI	55.8 (32.2)	41.6 (27.3)	<0.001	63.7 (65.0)	36.2 (26.1)	<0.001
Human Kallikrein 2, ng/ml	0.058 (0.048)	0.055 (0.051)	0.060	0.065 (0.083)	0.048 (0.049)	<0.001
Intact PSA, ng/ml	0.32 (0.30)	0.36 (0.35)	0.53	0.36 (0.48)	0.32 (0.36)	<0.001
Gleason score <=6, n (%)	105 (39)	N/A		129 (34)	N/A	
Gleason score 7, n (%)	117 (43)	N/A		142 (38)	N/A	
Gleason score >=8, n (%)	17 (6)	N/A		51 (14)	N/A	
Gleason unknown,n (%)	32 (12)	N/A		54 (14)	N/A	

Data are expressed as medians and interquartile range except age, which is expressed as mean and standard deviation.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2

Performance of models predicting prostate cancer when applied to a cohort of 531 previously unbiopsied men with PSA 3–15 ng/ml (260 controls and 271 prostate cancer cases of which 134 with high-grade cancer) including p-values for pairwise comparisons.

	All cancer			High-grade cancer		
Model	**AUC**	**CI (95%)**	**p**	**AUC**	**CI (95%)**	**p**
Base model	54.5	(49.6–59.4)	-	59.6	(54.1–65.8)	-
Four-kallikrein panel	69.0	(64.5–73.4)	<0.0001*	71.8	(66.8–76.7)	<0.0001*
PHI	70.4	(66.1–74.8)	<0.0001*	71.1	(66.0–76.2)	<0.0001*
AUC comparison: PHI vs four-kallikrein panel						
No sensitivity restriction			0.52			0.77
Restricted to sensitivity >75%			0.06			0.46

* versus base model.

The base model includes age and total PSA. The four-kallikrein panel model includes age, total PSA, free PSA, human kallikrein 2 and intact PSA. PHI includes total PSA, free PSA and [-2]proPSA. Cancers with Gleason 7 defined high-grade cancer. AUC = area under the curve; PSA = prostate-specific antigen

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3

Head-to-head evaluation of the four-kallikrein and PHI models in terms of saved biopsies and missed cancers compared to a strategy where the entire cohort is biopsied. The table shows strategies where previously unbiopsied men with PSA 3–15ng/ml would be biopsied if having 10%, 15% or 20% risk of finding high-grade cancer as predicted by the 4K model or by using the PHI values 26, 39, and 47 as cutoffs for biopsy.

	Biopsies			Cancers				High-grade cancers			
	Performed	Not performed		Caught	Missed		Risk of cancer in not performed biopsies	Caught	Missed		Risk of highgrade cancer in not performed biopsies
Cutoff	n	n	%	n	n	%	%	n	n	%	
Biopsy all	1000	0	0.0%	510	0	0.0%	NA	266	0	0.0%	NA
4K											
10%	704	296	29.6%	427	83	16.3%	28.0%	238	28	10.5%	9.5%
15%	550	450	45.0%	358	153	30.0%	34.0%	208	58	21.8%	12.9%
20%	437	563	56.3%	282	228	44.7%	40.5%	169	97	36.5%	17.2%
PHI											
26	915	85	8.5%	497	13	2.5%	15.3%	254	12	4.5%	14.1%
39	704	296	29.6%	433	77	15.1%	26.0%	240	26	9.8%	8.8%
47	550	450	45.0%	354	156	30.6%	34.7%	200	65	24.4%	14.4%

Numbers are given per 1,000 biopsied men. The four-kallikrein panel model includes age, total PSA, free PSA, human Kallikrein 2 and intact PSA. PHI includes total PSA, free PSA and [-2]proPSA. Cancers with Gleason ≥7 defined high-grade cancer. PSA = prostate-specific antigen.