REVIEW ARTICLE

Prostate health index vs percent free prostate-specific antigen for prostate cancer detection in men with ``gray'' prostate-specific antigen levels at first biopsy: systematic review and meta-analysis

DARIO BRUZZESE, CLAUDIA MAZZARELLA, MATTEO FERRO, SISTO PERDONÀ, PAOLO CHIODINI, GIUSEPPE PERRUOLO, and DANIELA TERRACCIANO

NAPLES AND MILAN, ITALY

The most promising approach to improve the specificity of prostate-specific antigen (PSA) test relies on the measurement of different molecular isoforms of PSA in serum. Currently, in men with a total PSA (tPSA) level between 2 and 10 ng/mL, measurement of %fPSA (free to total PSA ratio ×100) is used as reflex testing to better distinguish between malignant and benign prostate disease. Recently, Beckman Coulter developed the prostate health index (PHI) and several studies suggested that this test may improve the diagnostic ability of %fPSA.We performed a meta-analysis to evaluate the usefulness of PHI compared with %fPSA in the detection of prostate cancer (PCa) at first biopsy in men with tPSA "gray" levels of 2–10 ng/mL. Data on sensitivity and specificity were extracted from 8 eligible studies. Only observational studies comparing the diagnostic ability of PHI and %fPSA in tPSA range of 2–10 ng/mL were included. A total of 8 studies involving 2969 patients with a tPSA range of 2–10 ng/mL undergoing first biopsy were included in this meta-analysis. Biopsy-confirmed PCa was detected in 1287 (43.3%) men. Selected studies determined both PHI and %fPSA as a reflex test. The areas under curve (AUCs) of PHI and %fPSA were 0.74 (95% confidence interval (CI), 0.70–0.77) and 0.63 (95% CI, 0.58–0.67), respectively. Meta-regression analysis confirmed the superiority of PHI which showed, compared with %fPSA, a relative diagnostic odds ratio of 2.81 (95% CI, 2.19–3.6; $P < 0.0001$). In conclusion, PHI instead of %fPSA as a reflex test in men with tPSA "gray" levels is a better predictor of positive first biopsy and can offer a reduction in unnecessary biopsies. (Translational Research 2014;164:444–451)

Abbreviations: AUC = area under curve; CIs = confidence intervals; DOR = diagnostic odds ratio; DRE = digital rectal examination; FDA = food and drug administration; fPSA = freePSA; HSROC = hierarchical summary receiver-operator curves; PCa = prostate cancer; PHI = prostate health index; PSA = prostate specific antigen; pPSA = p roPSA; p2PSA = (-2)proPSA; USPSTF = US preventive services task force

From the Department of Public Health, University "Federico II", Naples, Italy; Department of Translational Medical Sciences, University "Federico II", Naples, Italy; Division of Urology, European Institute of Oncology, Milan, Italy; Urology Unit, IRCCS Fondazione G. Pascale, Naples, Italy; Medical Statistics Unit, Second University of Naples, Naples, Italy.

Submitted for publication March 31, 2014; revision submitted June 17, 2014; accepted for publication June 18, 2014.

Reprint requests: Daniela Terracciano, Department of Translational Medical Sciences, University "Federico II," Via Sergio Pansini 5, 80131 Naples, Italy; e-mail: daniela.terracciano@unina.it.

1931-5244/$ - see front matter

© 2014 Elsevier Inc. All rights reserved.

http://dx.doi.org/10.1016/j.trsl.2014.06.006

INTRODUCTION

Prostate-specific antigen (PSA) has been widely used in the management of patients with prostate cancer (PCa). However, serum PSA levels showed limited specificity particularly in a PSA range of 2–10 ng/mL.[1] Several approaches have been proposed to address these limitations, including the measurement of PSA molecular forms especially %fPSA. Anyway, the ability to detect PCa at initial biopsy remains limited.[2] Free PSA (fPSA) comprises proPSAs (pPSAs), benign PSA (BPSA), and intact PSA. Mikolajczyk et al[3] reported that pPSA is associated with cancer and BPSA with benign diseases, whereas the association of intact PSA is currently unknown. pPSA includes several truncated forms that can be measured in serum by immunoassays.[4] The [-2]proPSA (p2PSA) is the most cancer-specific form of all, being preferentially expressed in cancerous prostatic epithelium and significantly increased in serum of men with PCa.[2] During the past 2 years, 2 biomarkers have been approved by the US Food and Drug Administration. These include p2PSA as part of the prostate health index (PHI) by Beckman Coulter, Inc, calculated by a mathematical formula ([-2][pPSA/fPSA) × sqrt (PSA)] combining PSA molecular forms.[5] Currently, several studies suggested that increased PHI levels seem to preferentially detect patients with PCa.[5-9]

The hypothesis regarding the improvement in predicting biopsy outcome of PHI compared with %fPSA is inspiring, but till now, no consensus has been reached on which is the best test recommended in clinical practice. At present, only one meta-analysis has been performed[10] to assess the usefulness of %[-2]proPSA and PHI in PCa detection in the overall PSA range and in the initial and subsequent biopsies. Our meta-analysis is undertaken to evaluate the diagnostic value of PHI compared with %fPSA in the "gray" PSA range of 2–10 ng/mL in patients undergoing first biopsy.

METHODS

Meta-analysis was performed in accordance with the preferred reporting items from systematic reviews and meta-analysis (PRISMA) adapted to the study of diagnostic test.[11]

Relevant published articles were identified by searching computerized bibliographic systems (Pubmed, Web of Science, Scopus, Cochrane Library, and Cancerlit) from January 2009 to December 2013. A search strategy was used that contained the following text words and medical subject headings in their titles, abstracts, or keyword lists: PSA testing (prostate health index, PHI, tumor markers, p2PSA, sensitivity, specificity and performance) and PCa detection (PCa diagnosis or biopsy outcome). This literature search was complemented with the review of 5 specialized journals in Urology (European Urology, Journal of Urology, British Journal of Urology, International Journal of Urology and Prostate). The summary of literature search and selection of studies included is shown in Fig 1.

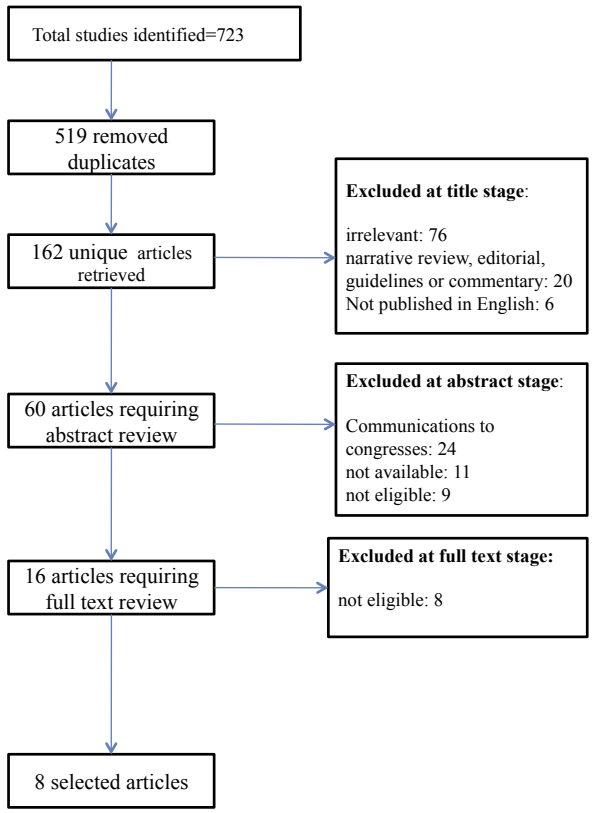

Fig 1. Summary of literature search and selection of studies included.

All the studies were considered eligible for inclusion if they met the following criteria:

1. original data;
2. study including at least 20 patients with PCa;
3. confirmation of PCa on transrectal ultrasound-guided needle biopsy (minimum ≥6 cores)
4. serum levels of fPSA, tPSA, and p2PSA evaluated by commercially available kits of Beckman Coulter using Hybritech calibration;
5. tPSA included between 2 and 10 ng/mL;
6. sufficient data to allow us to calculate true positive (TP), false negative (FN), false positive (FP) and true negative (TN) values for PCa diagnosis;
7. blood was sampled before prostate manipulation or biopsy and antiandrogen therapy;
8. the indication for biopsy was independent of the PHI test result;
9. results were based on first biopsy;

10. comparison of PHI and %fPSA diagnostic ability among men with tPSA levels between 2 and 10 ng/mL;
11. English language.

Studies with no usable data, lacking control groups, bad in quality, receiving therapy, or underwent to digital rectal examination (DRE) before samples were taken were excluded.

Two of the authors (D.T. and C.M.), who were blinded to the journal, author, institution, and date of publication, independently reviewed each publication. Both reviewers, separately, screened all titles and excluded studies if obviously irrelevant and removed duplicate citations.

To assess inter-rater consistency, an independent reviewer also extracted data using the same criteria for a random subset of articles (25%). The analysis of the concordance between both researchers about the eligibility of a study was done by calculating the kappa index.

We assessed the quality of studies using the quality assessment of studies of diagnostic accuracy (QUADAS) included in systematic review checklist,[12,13] which contains 14 items specifically developed to assess the quality of primary studies of diagnostic tests. Each item was scored as "yes," "no," or "unclear."

Data synthesis and analysis. Sensitivity and specificity with their 95% confidence intervals (CIs) were computed for both diagnostic markers in each study and reported graphically on forest plots.

Data extraction showed a large heterogeneity in the reporting of diagnostic accuracy measures: some studies reported the diagnostic performance constraining the sensitivity at a high value, whereas others fixed a priori the value of specificity; in some cases the pairs of sensitivity and specificity were determined by looking at the best combination of detection rate (sensitivity) and FP rate. Occasionally, multiple pairs of sensitivity and specificity were reported. This led to a large variability in the positivity thresholds that, additionally, in some studies were not specified.

For such reasons, the average operating points (ie, summary sensitivity and specificity) with the corresponding 95% confidence regions were not computed as recommended in the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.[14] On the contrary, the summary Receiver operating characteristics (ROC) curves for PHI and % fPSA were constructed using the hierarchical model proposed by Rutter and Gatsonis.[15] This approach, that accounts for both within and between studies variation, models the relationship between sensitivity and specificity by using three parameters: (1) a proxy for positivity threshold, measured on an underlying latent scale; (2) a measure of accuracy defined as the logarithm of the diagnostic odds ratio (DOR), which measures the ratio between the odds of a positive result between diseased and nondiseased subjects; (3) a scale parameter that models the symmetry of the curve allowing test accuracy to vary with "threshold." The first 2 are treated as random effects (ie, study dependent), whereas the scale parameter is considered fixed. The Hierarchical summary receiver-operator curves (HSROC) parameterization is considered the most appropriate approach when the variability in threshold is expected to be larger than the variability in the accuracy.[16]

Diagnostic performance of each marker was thus summarized using the area under curve (AUC) of the summary ROC curves, restricted to observed FP rates and normalized; the partial AUC was computed using the estimates of the HSROC parameters. In case of symmetric curves, the DOR was also computed, as for symmetric curve the DOR does not depend on the specific threshold.

The comparison between the 2 markers was undertaken by using a meta-regression approach, treating the type of test (PHI and %fPSA) as a dummy variable and exploring its effect on the shape (symmetry) and position (accuracy) of the summary ROC curve.

The SAS statistical package, the Stata software (StataCorp, College Station, Texas), and the R language were used for all the statistical analysis. A P value <0.05 was used to denote statistical significance.

RESULTS

Study characteristics and quality. In the literature search, 723 articles were identified. Among these, 162 unique articles were retrieved. As shown in Fig 1, 16 articles required full-text review and 8 studies, meeting the inclusion criteria, were finally included in this meta-analysis. All these studies contain data on both PHI and %fPSA. Of note, the study by Jansen was performed on 2 distinct populations (Rotterdam and Innsbruck) and was analyzed separately as 2 studies.

The results about the concordance between both reviewers had a kappa index of 0.88 (95% CI, 0.64–1.00; $P < 0.001$).

The baseline characteristics of the selected studies are shown in Table I including the type of population studied, age range of participants, size of each study, and the sampling frame. In addition, tPSA range, percentage of positive DRE, and biopsies were indicated. The criterion used to establish the diagnostic accuracy of each marker (high sensitivity, high specificity, and best combination) is also shown together with the reported positivity thresholds. The existence of a large variability in

Table I. Characteristics of the studies included in the meta-analysis

Reference	Sampling frame	Years of recruitment of patients	Population	Age of patients	tPSA range (ng/mL)	Digital rectal examination	Number of biopsies	Patients with cancer	Criterion for reporting diagnostic accuracy	PHI positivity threshold (greater than)	%fPSA positivity threshold (less than)
Ferro et al 2013[25]	Single centre and prospective	Not reported	Referral	60.4 ± 6 (Mean ± SD)	2–10	19% Positive	300	108 (36%)	Best combination	42.8	12%
Guazzoni et al 2011[7]	Single centre and prospective	2010	Referral	63.3 ± 8.2 (Mean ± SD)	2–10	Negative	268	107 (40%)	High specificity	48.5	29%
Jansen et al 2010 Site 1[6]	Retrospective	1994–1997	Screening and not serial	66 (55–75) Median (range)	2–10	Positive	405	226 (56%)	High specificity	Not reported	Not reported
Jansen et al 2010 Site 2[6]	Retrospective	Started in 1993	Screening and not serial	69 (50–77) Median (range)	2–10	Not reported	351	174 (50%)	High specificity	Not reported	Not reported
Lazzeri et al 2013[26]	Multicentre and prospective	2011–2012	Referral	64.2 ± 7.5 (Mean ± SD)	2–10	18% Positive	646	264 (40%)	Best combination	41.5	15%
Le et al 2010[5]	Single centre and prospective	2007	Screening and consecutive	65 (Median)	2.5–10	Negative	63	26 (41%)	High sensitivity	Not reported	Not reported
Loeb et al 2013[27]	Multi-centre and prospective	2003–2009	Selected	≥50	2–10	Negative	706	430 (48%)	High sensitivity	35.0	Not reported
Ng et al 2013[28]	Single centre and retrospective	2008–2013	Referral	65.9 (50–79) Mean (range)	4–10	Negative	230	21 (9%)	High sensitivity	26.5	28%

Abbreviations: %fPSA, free to total PSA ratio ×100; *PHI,* prostate health index; *tPSA,* total prostate-specific antigen; *SD,* standard deviation.

Table II. Quality of the studies included in the meta-analysis according to the quality assessment of studies of diagnostic accuracy included in systematic review questionnaire

Reference	Patients are representative of the question	Selection criteria clearly described	Number of cores per biopsy ≥10	Assays for the measurement of PHI are described	Blinded
Ferro et al 2013[25]	Yes	Yes	Yes	Yes	Yes
Guazzoni et al 2011[7]	Yes	Yes	Yes	Yes	Yes
Jansen et al 2010 Site 1[6]	Yes	Yes	No	Yes	No
Jansen et al 2010 Site 2[6]	Yes	Yes	No	Yes	No
Lazzeri et al 2013[26]	Yes	Yes	Yes	Yes	Yes
Le et al 2010[5]	Yes	Yes	Not reported	Yes	Yes
Loeb et al 2013[27]	Yes	Yes	Yes*	Yes	Yes
Ng et al 2013[28]	Yes	Yes	Yes	Yes	Yes

*Median (range): 12 (6–34)

Table III. Quality assessment of studies of diagnostic accuracy included in systematic review questionnaire

	Generalizability	Clarity					Validity							
Reference	Q1	Q2	Q8	Q9	Q13	Q14	Q3	Q4	Q5	Q6	Q7	Q10	Q11	Q12
Ferro et al 2013[25]	+	+	+	+	+	+	+	?	+	+	+	+	+	Does not apply
Guazzoni et al 2011[7]	+	+	+	+	+	+	+	?	+	+	+	+	+	Does not apply
Jansen et al 2010 Site 1[6]	−	+	+	+	+	+	+	?	+	+	+	−	−	Does not apply
Jansen et al 2010 Site 2[6]	−	+	+	+	+	+	+	?	+	+	+	−	−	Does not apply
Lazzeri et al 2013[26]	+	+	+	+	+	+	+	?	+	+	+	+	+	Does not apply
Le et al 2010[5]	−	+	+	−	+	−	+	?	−	+	+	+	+	Does not apply
Loeb et al 2013[27]	+	+	+	+	+	+	+	?	+	+	+	+	+	Does not apply
Ng et al 2013[28]	+	+	+	+	+	+	+	?	+	+	+	+	+	Does not apply

Q1, spectrum of patients; Q2, selection criteria; Q8, index test; Q9, reference test; Q13, uninterpretable and/or intermediate test results; Q14, withdrawals; Q3, reference test; Q4, time between reference test and index test; Q5, verification using reference test; Q6, reference standard regardless of index test results; Q7, reference standard independent of the index test; Q10, blinding to reference test; Q11, blinding to index test; Q12, same data available before interpretation of both index and reference tests.
+ Indicates "yes", − "no", and ?, "unclear."

the threshold ranging from 26.5 to 48.5 for PHI and from 12% to 29% for %fPSA is evident.

The quality evaluation of the selected studies was assessed according to the quality assessment of studies of diagnostic accuracy included in systematic review scale[12,13] (Tables II and III).

Before meta-analysis was performed, methodological heterogeneity was evaluated. However, no studies were excluded because of this cause.

Diagnostic performance of PHI compared with %fPSA as a reflex test in the tPSA range of 2–10 ng/mL. Fig 2 shows the sensitivity and the specificity along with their 95% CIs for each individual study for both PHI and %fPSA in the tPSA range of 2–10 ng/mL. Their funnel-like shape, which reflects the inverse relation between sensitivity and specificity, is sharpened by the heterogeneity in the criterion used to report the diagnostic performance. As expected, the I^2 statistics was significant ($P < 0.001$) for both the tests and the proportion of heterogeneity because of a threshold effect was 0.94 for PHI and 1.00 for %fPSA.

Fig 3 displays the pairs of sensitivity and specificity in the ROC space. For each individual study, a straight line joins the results associated with the 2 tests. The corresponding partial AUCs were 0.63 (95% CI, 0.58; 0.67) for %fPSA and 0.74 (95% CI, 0.70–0.77) for PHI. Both diagnostic tests resulted in a symmetric curve (test for the shape coefficient not significant: $P = 0.57$ and $P = 0.78$, respectively). The DOR was estimated to be 5.57 (95% CI, 4.20–7.40) for PHI and 2.01 (95% CI, 1.56–2.60) for %fPSA.

When comparing the 2 tests, no difference in shape emerged. In particular, the increase in the −2log likelihood observed when removing the covariate for shape was not significant (chi square = 2.82; 2 df; $P > 0.05$). On the contrary, PHI showed a significant increase in the accuracy parameter leading to a relative DOR of 2.81 (95% CI, 2.19–3.6; $P < 0.0001$). This means that the ratio between the odds of a positive result between diseased and nondiseased subjects is at least doubled when using PHI instead of %fPSA.

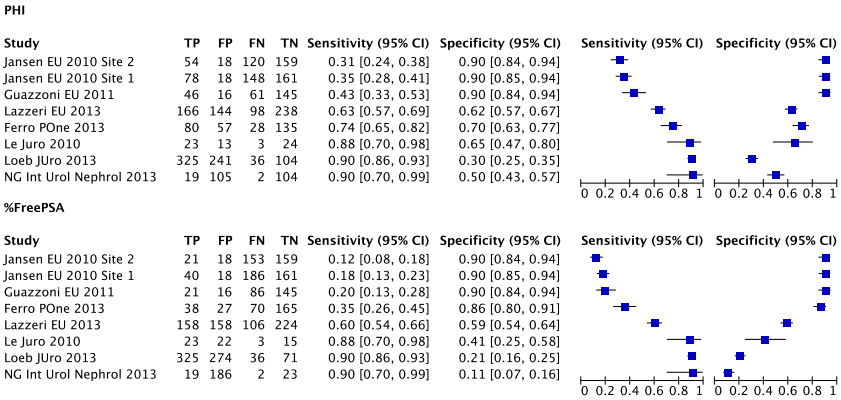

Fig 2. Forest plots of study results (sensitivity and specificity) for PHI and %fPSA as a predictor of prostate cancer. Squares represent study estimates, whereas lines represent the corresponding 95% CI obtained with the exact binomial method. CI, confidence interval; TP, true positive; FP, false positive; FN, false negative; TN, true negative; PHI, prostate health index; %fPSA, free to total PSA ratio ×100.

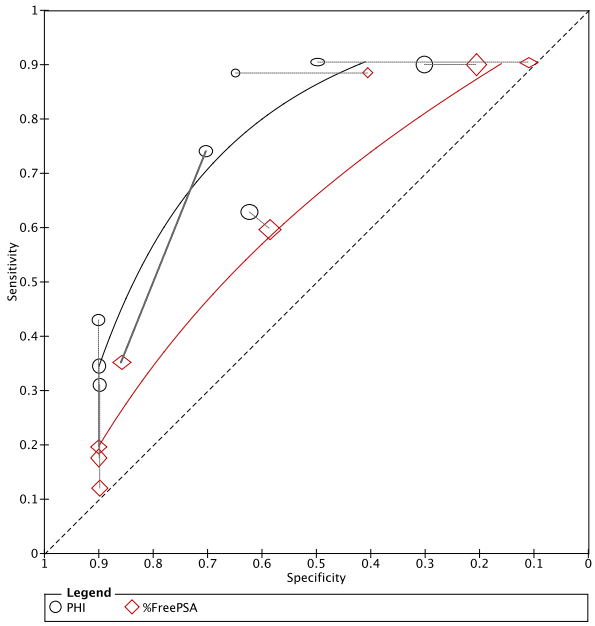

Fig 3. ROC space showing the relationship between sensitivity and specificity for PHI (black circles) and %fPSA (red diamonds). For each study, the results of the 2 tests are joined by a straight line. The curves represent the summary ROC curve associated with PHI (black) and %fPSA (red). PHI, prostate health index; %fPSA, free to total PSA ratio ×100. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

DISCUSSION

Current clinical practice has been driven largely by the widespread use of PSA, allowing for unnecessary biopsies and detection of many cases of indolent PCa. Recent U.S. preventive services task force[17] recommends against PSA-based screening for PCa in all age groups, highlighting the need for novel clinical useful biomarkers.

The interest in more specific biomarkers encouraged the development of new test improving clinical outcome. Several studies have indicated that reflex tests based on PSA isoforms can improve cancer detection in men with tPSA levels between 2 and 10 ng/mL. In particular, it has been widely accepted that a low %fPSA is a useful test to reduce the number of unnecessary biopsies.[18] However, it is now known that fPSA fraction is composed of at least 3 different types of enzymatically inactive PSA: BPSA, intact inactive PSA, and pPSA, of which BPSA and pPSA are the best characterized. In patients with PCa, serum pPSA comprised primarily a truncated form of pPSA that contains a proleader peptide consisting of only 2 ([-2]pPSA) rather than the usual 7 amino acids ([-7]pPSA).[19] Thus, serum pPSAs gained attention as a potentially specific form of fPSA that may help overcome the current limitations of %fPSA, reducing the highest number of unnecessary biopsies. On the basis of substantial evidence for the role of [-2]pPSA in early PCa detection, Beckman Coulter Inc developed a mathematical algorithm incorporating [-2]pPSA, tPSA, and fPSA for use in patients with PSA levels of 2–10 ng/mL with a nonsuspicious prostate on DRE.

This meta-analysis is the first study that compares the published data on the clinical usefulness of PHI compared with %fPSA in subjects undergoing first prostate biopsy with PSA levels in the 2–10 ng/mL. The results of this meta-analysis showed that the use of PHI compared with %fPSA can improve the detection of PCa in men who have PSA levels of 2–10 ng/mL. The AUCs obtained by the HSROC analysis were 0.74 for PHI and 0.63 for %fPSA, suggesting a better ability of PHI in PCa detection. Although both tests showed a

large variability in the thresholds, mainly because of the different criteria used to derive the diagnostic performance, the symmetry of the corresponding summary ROC curves allowed to compare the 2 tests also in terms of DORs. DOR values showed that the odd ratio of positive results between patients with cancer and without cancer is about 3-fold greater for PHI than for %fPSA in patients with tPSA levels included between 2 and 10 ng/mL. Of note, in all 8 selected studies in the same study population PHI and %fPSA were equally assayed, so recruitment strategy or population characteristics were not likely causes of differences in diagnostic performance of the two tests.

An increasing number of patients are being diagnosed with potentially low-risk, clinically insignificant cancers. In these patients, active surveillance has been proposed as an alternative treatment strategy with the aim of reducing the risk of radical prostatectomy side effects. However, we currently lack an ideal definition of indolent PCa, and circulating biomarkers could be a promising tool to identify patients harboring aggressive disease.

Unfortunately, at present only few studies evaluated the usefulness of PHI as a predictor of aggressive PCa,[8,20] impeding to perform meta-analysis. However, Filella and Gimenez[10] reported a critical analysis showing that PHI may distinguish low- and high-risk PCa. Further studies are required to better address this issue.

This study has some limitations because of a very high heterogeneity of the cutoffs used in the 8 eligible studies included in this meta-analysis. In particular, as shown in Table I, sensitivity and specificity are reported differently among selected studies and information about best cutoffs was not shown in about 50% of the primary studies. Therefore, our meta-analysis was not able to obtain data on pooled specificity about PHI, to recommend an optimal cutoff value, and to quantify missed cancers and avoided biopsies. To these aims a more homogeneous criterion in reporting diagnostic accuracy may be recommended in future reports.

Within specific PCa detection plan, a cost-benefit analysis needs to be carried out to quantify whether the extra costs of a PHI test with its corresponding reduced numbers of biopsies will offer a net saving to the health care provider. Such an analysis was recently performed,[21] suggesting that PSA plus PHI was a less costly diagnostic strategy to detect PCa compared with the PSA test alone. Taking into account that the cost of %fPSA is clearly higher than PSA alone and PHI can be expected to reduce false negative tests, the use of PHI instead of %fPSA as a reflex test in tPSA range of 2–10 ng/mL may decrease the global costs and disutility related to the prostate biopsy procedure.

In conclusion, our results suggested that PHI in patients with "gray" values of tPSA may be a better predictor of positive biopsy compared with %fPSA. Filella and Gimenez[10] reported in overall PSA range and initial and subsequent biopsies that PHI increases the specificity in PCa detection.

Several differences exist between our results and those reported in Filella and Gimenez.[10] From a methodological perspective, our meta-analysis did not estimate the summary pair of sensitivity and specificity but rather we derived a summary ROC curve allowing to compare the 2 markers along the whole range of sensitivity and specificity pairs. This choice was motivated by the large inconsistency in the threshold used that could lead to a misleading interpretation of the pooled results.[22]

Moreover, differently from Filella and Gimenez,[10] our results focused on the PSA range (2–10 ng/mL) with the highest overlap of benign and malignant pathology, addressing the potential usefulness of PHI compared with %fPSA as a reflex test.

Therefore, the use of PHI can offer a reduction in unnecessary biopsies, whereas maintaining a high cancer detection rate. However, given PCa heterogeneity, in later years efforts to increase diagnostic accuracy focused on a combination of biomarkers. Promising results have been obtained by the combination of Prostate cancer antigen 3 (PCA3) and v-ets erythroblastosis virus E26 oncogene homolog (TMPRSS2-ERG) gene fusions[23] and with prostate health index (PHI).[24]

Furthermore, more studies are needed to define the optimal PHI cutoff value and to evaluate the ability of PHI in the discrimination of indolent and aggressive PCa. Finally, worthy of further attention is the potential improvement in diagnostic performance of the combination of PHI with other biomarkers.

REFERENCES

1. Schroder FH, Roobol MJ. Defining the optimal prostate-specific antigen threshold for the diagnosis of prostate cancer. Curr Opin Urol 2009;19:227–31.
2. Sokoll LJ, Wang Y, Feng Z, et al. [-2]Proenzyme prostate specific antigen for prostate cancer detection: a national cancer institute early detection research network validation study. J Urol 2008; 180:539–43. discussion 43.
3. Mikolajczyk SD, Marks LS, Partin AW, Rittenhouse HG. Free prostate-specific antigen in serum is becoming more complex. Urology 2002;59:797–802.
4. Mikolajczyk SD, Catalona WJ, Evans CL, et al. Proenzyme forms of prostate-specific antigen in serum improve the detection of prostate cancer. Clin Chem 2004;50:1017–25.
5. Le BV, Griffin CR, Loeb S, et al. [-2]Proenzyme prostate specific antigen is more accurate than total and free prostate specific antigen in differentiating prostate cancer from benign disease in a prospective prostate cancer screening study. J Urol 2010;183: 1355–9.

6. Jansen FH, van Schaik RH, Kurstjens 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.
7. Guazzoni G, Nava L, Lazzeri M, 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.
8. Guazzoni G, Lazzeri M, Nava L, 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.
9. Catalona WJ, Partin AW, Sanda MG, 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.
10. Filella X, Gimenez N. Evaluation of [-2] proPSA and Prostate Health Index (phi) for the detection of prostate cancer: a systematic review and meta-analysis. Clin Chem Lab Med 2013;51:729–39.
11. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339:b2700.
12. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 2003;3:25.
13. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529–36.
14. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM. Systematic reviews of diagnostic test accuracy. Ann Intern Med 2008;149:889–97.
15. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med 2001;20:2865–84.
16. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. Biostatistics 2007;8:239–51.
17. Moyer VA. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2012;157:120–34.
18. Roddam AW, Duffy MJ, Hamdy FC, et al. Use of prostate-specific antigen (PSA) isoforms for the detection of prostate cancer in men with a PSA level of 2-10 ng/ml: systematic review and meta-analysis. Eur Urol 2005;48:386–99. discussion 98–9.
19. Hori S, Blanchet JS, 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 2012;112:717–28.
20. Sottile A, Ortega C, Berruti A, et al. A pilot study evaluating serum pro-prostate-specific antigen in patients with rising PSA following radical prostatectomy. Oncol Lett 2012;3:819–24.
21. Nichol MB, Wu J, Huang J, Denham D, Frencher SK, Jacobsen SJ. Cost-effectiveness of Prostate Health Index for prostate cancer detection. BJU Int 2011;110:353–62.
22. Leeflang MM. Systematic reviews and meta-analyses of diagnostic test accuracy. Clin Microbiol Infect 2014;20:105–13.
23. Salami SS, Schmidt F, Laxman B, et al. Combining urinary detection of TMPRSS2:ERG and PCA3 with serum PSA to predict diagnosis of prostate cancer. Urol Oncol 2013;31:566–71.
24. Perdona S, Bruzzese D, Ferro M, et al. Prostate health index (phi) and prostate cancer antigen 3 (PCA3) significantly improve diagnostic accuracy in patients undergoing prostate biopsy. Prostate 2013;73:227–35.
25. Ferro M, Bruzzese D, Perdona S, et al. Prostate Health Index (Phi) and Prostate Cancer Antigen 3 (PCA3) significantly improve prostate cancer detection at initial biopsy in a total PSA range of 2-10 ng/ml. PLoS ONE 2013;8:e67687.
26. Lazzeri M, Haese A, de la Taille A, 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.
27. Loeb S, Sokoll LJ, Broyles DL, et al. Prospective multicenter evaluation of the Beckman Coulter Prostate Health Index using WHO calibration. J Urol 2013;189:1702–6.
28. Ng CF, Chiu PK, Lam NY, Lam HC, Lee KW, Hou SS. The Prostate Health Index in predicting initial prostate biopsy outcomes in Asian men with prostate-specific antigen levels of 4-10 ng/mL. Int Urol Nephrol 2014;46:711–7.