Review

Xavier Filella* and Nuria Giménez

Evaluation of [–2] proPSA and Prostate Health Index (phi) for the detection of prostate cancer: a systematic review and meta-analysis

Abstract

The usefulness of %[–2] proPSA and Prostate Health Index (phi) in the detection of prostate cancer are currently unknown. It has been suggested that these tests can distinguish prostate cancer from benign prostatic diseases better than PSA or %fPSA. We performed a systematic review and meta-analysis of the available scientific evidence to evaluate the clinical usefulness of %[–2] proPSA and phi. Relevant published papers were identified by searching computerized bibliographic systems. Data on sensitivity and specificity were extracted from 12 studies: 10 studies about %[–2] proPSA (3928 patients in total, including 1762 with confirmed prostate cancer) and eight studies about phi (2919 patients in total, including 1515 with confirmed prostate cancer). The sensitivity for the detection of prostate cancer was 90% for %[–2] proPSA and phi, while the pooled specificity was 32.5% (95% CI 30.6–34.5) and 31.6% (95% CI 29.2–34.0) for %[–2] proPSA and phi, respectively. The measurement of %[–2] proPSA improves the accuracy of prostate cancer detection in comparison with PSA or %fPSA, particularly in the group of patients with PSA between 2 μg/L and 10 μg/L. Similar results were obtained measuring phi. Using these tests, it is possible to reduce the number of unnecessary biopsies, maintaining a high cancer detection rate. Published results also showed that %[–2] proPSA and phi are related to the aggressiveness of the tumor.

Keywords: evidence-based laboratory medicine; meta-analysis; prostate cancer; Prostate Health Index (phi); prostate specific antigen (PSA); ProPSA; systematic review.

***Corresponding author: Xavier Filella**, Department of Biochemistry and Molecular Genetics, Hospital Clinic, Villarroel 170, 08036 Barcelona, Catalonia, Spain, E-mail: xfilella@clinic.ub.es
Xavier Filella: Department of Biochemistry and Molecular Genetics (CDB), Hospital Clinic, IDIBAPS, Barcelona, Spain
Nuria Giménez: Research Unit, Research Foundation Mútua Terrassa, University of Barcelona, Barcelona, Spain; Laboratory of Toxicology, Autonomous University of Barcelona, Barcelona, Spain; and Evidence Based Laboratory Medicine Commission of the Spanish Society of Clinical Biochemistry and Molecular Pathology (SEQC)

Introduction

Prostate specific antigen (PSA) is a serum tumor marker that is widely used in the early detection of prostate cancer. However, since the specificity (Sp) of PSA is limited, biopsy is positive in approximately 25% of patients with PSA in the range between 2 μg/L and 10 μg/L [1]. Furthermore, prostate cancer is detected on repeated biopsy in 10%–35% of patients with a negative first biopsy. So, according to the guidelines of the European Association of Urology, it is necessary to repeat the biopsy in these patients [2].

The measurement of the several fractions of PSA (free PSA, complexed PSA) has been proposed with the aim to improve the Sp of total PSA. A meta-analysis, published in 2005, showed that the use of the percentage of free PSA (%fPSA) is useful to improve the detection of prostate cancer [3]. More recently, fPSA has been found to include the subforms BPSA, iPSA and proPSA [4, 5]. BPSA and iPSA are associated with benign tissue, but proPSA is associated with cancer. It is possible to detect three truncated forms of proPSA in serum, [–2], [–4] and [–5,–7], with [–2] proPSA being the most stable form. Several studies suggested the clinical usefulness of proPSA in the detection of prostate cancer using different non-commercial assays, including the measurement of the cumulative sum of all truncated forms [6, 7] and the measurement of [–5,–7] proPSA [8, 9]. However, these tests have not been shown to be as useful as the new assay for the measurement of [–2] proPSA. Also, the use of a panel of four kallikrein markers – total PSA, free PSA, intact PSA and hK2 – in the detection of prostate cancer has been proposed by recent studies [10, 11].

The development of the [–2] proPSA assay by Beckman Coulter opens a new field of study in the detection of prostate

cancer. Currently, several studies have suggested that in men with a total PSA between 2.5 µg/L and 10 µg/L, the percentage of [–2] proPSA to fPSA (%[–2] proPSA) can distinguish between malignant and benign prostate diseases better than total PSA or %fPSA. Also, several studies underlined the usefulness of the Prostate Health Index (phi), a mathematical combination of total PSA, fPSA and [–2] proPSA according to the formula [–2] proPSA/fPSA)×√tPSA.

The objective of this systematic review was to assess the usefulness of %[–2] proPSA and phi in the detection of prostate cancer. A critical analysis of results referring to the relationship between these tests and the aggressiveness of prostate cancer was also performed.

Methods

Meta-analysis was performed in accordance with the preferred reporting items from systematic reviews and meta-analysis (consensus PRISMA) adapted to studies of diagnostic tests [12]. In short, the PRISMA statement is a consensus that intends to inform by evidence whenever possible and consists of a 27-item checklist and a four-phase flow diagram that are available for researchers on internet for free (http://www.prisma-statement.org/).

Search strategy and study selection

A systematic search of several electronic databases was performed: MedLine, Embase, Cancerlit, Cochrane Library, Web of Science and Scopus. A strategy search in title, abstract or keyword lists was done looking for combinations of the following search terms: as medical subject headings MeSH ("Prostatic Neoplasms", "Sensitivity and Specificity", "Diagnosis", "Evidence-Based Medicine") and as free search terms ("proPSA", "p2PSA", "[–2]proPSA", "[–2]proenzyme prostate specific antigen", "Prostate Health Index", "phi", "Prostate tumor", "Prostate tumour"). This literature search was complemented with the review of three specialized journals in Urology (European Urology, Journal of Urology and Prostate) from January 1990 to December 2011. Furthermore, the authors checked the cited bibliographies of selected studies and contacted experts.

To avoid duplication of information, when the same population was reported in several publications, priority was given to scientific articles over meeting abstracts or in case there was more than a scientific article, the most complete study was chosen.

Eligibility criteria

All the studies about diagnostic tests and systematic review about %[–2] proPSA and phi were considered eligible for inclusion if they met the following criteria: original data and confirmation of prostate cancer on biopsy. There were no language restrictions.

Data extraction

All the studies were assessed independently by both researchers to determine study inclusion. Both reviewers, separately, screened all titles and excluded studies if obviously irrelevant and removed duplicate citations. When there was any doubt concerning the eligibility of a study, the abstract was examined and, if necessary, the full text. After selecting relevant studies, data extraction was carried out using a standardized form. The analysis of the concordance between both researchers about the eligibility of a study and the values of true positive (TP), false-positive (FP), false negative (FN) and true negative (TN) was done by calculating the kappa index. Disagreements about eligibility and data extraction were resolved by consensus.

Assessment of risk of bias

The quality of the selected studies was assessed by using quality assessment of diagnostic accuracy studies (QUADAS) [13]. The QUADAS tool consists of a set of 14 items, phrased as questions, each of which should be scored as yes, no or unclear. Possible sources of heterogeneity between studies were examined. Methodological heterogeneity or differences in design or quality were assessed during the selection of relevant studies and statistical heterogeneity was measured using I^2 scores and the χ^2-test.

The protocol was prepared a priori and this study was done in accordance with the Research Ethics Committee of Mútua Terrassa Hospital, Barcelona, Spain.

Data analysis

For each study, 2×2 tables for each test with TP, FP, FN and TN results using data extraction from the original referred scientific articles were performed. Pooled estimates of sensitivity (Se) and Sp as the main outcome measures were calculated as well as the limits of the 95% confidence intervals for such values. Forest plot was represented

as figures. Methodological heterogeneity was assessed during selection.

The threshold effect is a characteristic source of heterogeneity in the meta-analysis of diagnostic tests and arises when the included studies uses different cut-off points to define what is considered as a positive result of a diagnostic test. The analysis of diagnostic threshold was assessed through receiver operating characteristic (ROC) plane and correlation coefficient Spearman. The ROC plane is the graphic representation of the pairs of Se and Sp and, characteristically its points show a curvilinear pattern if the threshold effect exists. Statistical heterogeneity was measured using the χ^2-test and I^2 scores. I^2 score was used as a measure of the inconsistency between studies in the meta-analysis and was interpreted as low (25%–50%), moderate (51%–75%) and high (>75%).

Data were analyzed using a free statistical software package Metadisc version 1.4 [14], with the only exception of the analysis of the concordance between reviewers and kappa index which was performed using SPSS 17.0 (SPSS Inc., Chicago, IL, USA).

Assays used in the references evaluated in this study

In the studies corresponding to references [15–27] the concentrations of [–2] proPSA were measured in a Beckman Coulter ACCESS® immunoassay system, using dual monoclonal antibodies. [–2] proPSA was measured in references [28, 29] using a dual monoclonal sandwich assay in a microtiter plate. PSA and fPSA were measured using a Beckman Coulter ACCESS® immunoassay system in references [15–24] or Hybritech Tandem PSA and Tandem free PSA assays in reference [28]. The measurement of PSA and fPSA in reference [29] was determined with Hybritech Tandem PSA and Tandem free PSA assays (Beckman Coulter, Inc.) in site 2 (Washington University) and with the Abbott total and free PSA assays (Abbott Laboratories, Chicago, IL, USA) in site 1 (Innsbruck University).

Phi was calculated in studies corresponding to references [16–21, 25, 27] using the formula [–2] proPSA/fPSA)×√tPSA.

Results

Two hundred and thirteen potentially relevant references were obtained by electronic databases and supplementary sources in our systematic search. The results of the search and study selection process are shown in Figure 1. There were 31 articles requiring full-text review, and 12 studies were finally included in the meta-analysis. Data on Se and Sp were pooled from 10 studies for %[–2] proPSA (3928 patients in total, including 1762 with confirmed prostate

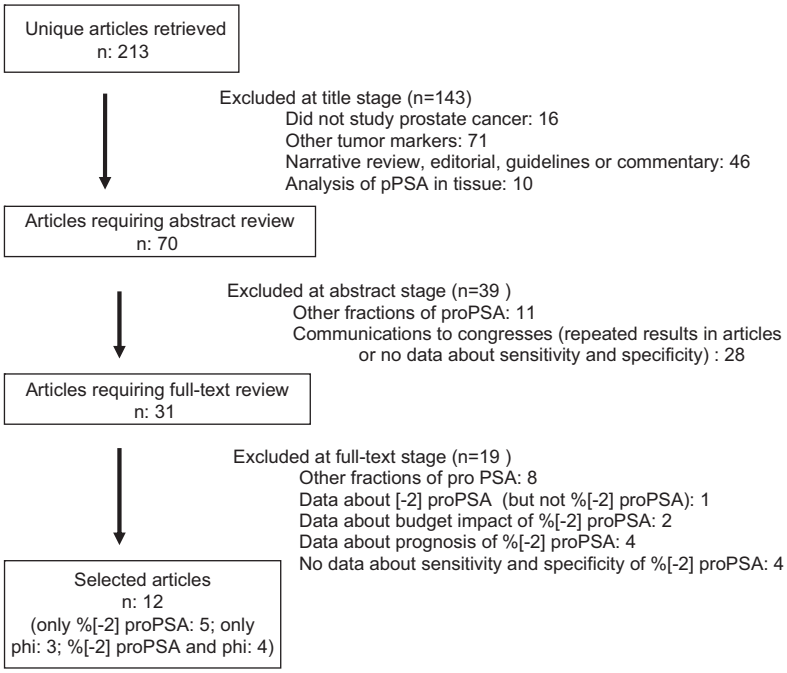

Figure 1 Summary of literature search and selection of studies included.

cancer) and eight studies about phi (2919 patients in total, including 1515 with confirmed prostate cancer).

The study by Jansen et al. [15] contained two different populations (Rotterdam and Innsbruck), and was treated as two separate studies.

The results about concordance between both reviewers had a coincidence of 94% and a kappa index of 0.812 (95% CI 0.635–0.990).

The quality assessment of the eligible studies was moderate-high according to QUADAS scale (Table 1) [15–24, 28, 29]. The main characteristics about the selected studies are shown in Table 2 including the description of the population of each study, the sampling frame and the criteria and characteristics of prostate biopsy. Table 3 shows the performance of %[–2] proPSA and phi and compares the area under the curve (AUC) corresponding to these tests with the AUC for PSA and %fPSA. The accuracy of %[–2] proPSA and phi in the detection of prostate cancer is reported in Table 4. Data presented in this table were extracted from the included studies. Of the 12 studies included, only three specified the cut-off value. The cut-off level for %[–2] proPSA at a Se of 90% was 2.5% for Mikolajczyk et al. [28] and 1.06% for Miyakubo et al. [19]. The cut-off reported for phi at a Se of 90% was 24.9% for Miyakubo et al. [19] and 21.1% for Catalona et al. [16].

Methodological heterogeneity was assessed before analyses and no studies were excluded due to this reason. The existence of a threshold effect was ruled out after examining the ROC plane and Spearman's correlation coefficient (r=0.636 and p-value=0.048 for %[–2] proPSA and r=0.262 and p-value=0.531 for phi).

When revising the studies, it was found that they had in common the results for sensibility of 90% and therefore it was decided to extract the data and perform calculations to this Se. There was a high degree of statistical heterogeneity (I^2 score ≥75%) in Sp of %[–2] proPSA (χ^2=84.24; p<0.0001) and phi (χ^2=36.07; p<0.0001). Results are shown in Figure 2. For this selected Se of 90%, the pooled Sp of %[–2] proPSA was 32.5% (95% CI 30.6–34.5%, I^2 score=89.3%, p<0.001, Figure 2A) and the pooled Sp of phi was 31.6% (95% CI 29.2–34.0%, I^2 score=80.6%, p<0.001, Figure 2B).

Discussion

A low %fPSA has been shown to be associated with prostate cancer and several studies have indicated that this test is useful in reducing the number of negative biopsies [3]. However, currently, we know that fPSA is composed of three distinct molecular forms, which are associated differently with cancer. Initial clinical studies showed that proPSA may be a useful marker for the detection of prostate cancer, and more recently Beckman Coulter introduced a new immunoassay for the measurement of the [–2] proPSA, a stable form of proPSA [30].

This meta-analysis is the first study that shows the available information on the clinical usefulness of this tumor marker in the detection of prostate cancer. Data on Se and Sp about %[–2] proPSA and the derivative test phi were extracted from 12 eligible studies. At Se of 90%, which is clinically acceptable, the Sp was 32% for %[–2] proPSA, ranging between 21% and 49%, and 32% for phi, ranging between 26% and 43%. The AUCs obtained by ROC analysis were also clinically acceptable, with results between 0.635 and 0.780 for %[–2] proPSA and between 0.703 and 0.77 for phi.

This study has some limitations. For one, information about the cut-offs used was showed only in three studies [16, 19, 28]; therefore, there was heterogeneity in primary studies. The high level of inconsistency in the global Sp for %[–2] proPSA (89%) and for phi (81%) shows the heterogeneity of the studies included in this meta-analysis. Differences in recruitment strategy, in population characteristics, and in the number of cores obtained in biopsies may contribute to these variations. We must underline that the same assay was used in the majority of studies, with only two exceptions, corresponding to the earlier references [28, 29] that uses a non-commercial assay for the measurement of [–2] proPSA. This factor may influence in part in the heterogeneity of results. PSA and fPSA were measured using an equivalent assay (Beckman Coulter ACCESS® immunoassay or Hybritech Tandem assays) in all studies, only with a partial exception in reference [29], that used the Abbott total and free PSA assays in part of the measurements.

%[–2] proPSA and phi have a similar performance for patients with PSA between 2 µg/L and 4 µg/L and for patients with PSA between 4 µg/L and 10 µg/L according to different studies [17, 22, 24, 29]. So, Guazzoni et al. [17] showed that the AUC for %[–2] proPSA is 0.76 for patients with PSA between 2 µg/L and 4 µg/L and 0.78 for patients with PSA between 4 µg/L and 10 µg/L. For both groups of patients the AUC for phi was 0.76. Similar results were indicated for %[–2] proPSA in other studies [22, 24, 29].

The majority of studies reported in this meta-analysis showed that the AUC for %[–2] proPSA (ranging between 0.635 and 0.78) was higher than the AUC for %fPSA. Sokoll et al. [22] communicated an exception to this criteria, but in this study, too, the AUC for %[–2] proPSA was higher to %fPSA in the group of patients with PSA between 2 µg/L

Study	Patients		Test			Results	
Author	Patients are representative of the question	Selection criteria according DRA and PSA serum levels	Biopsy is performed in all patients	Number of cores per biopsy ≥ 10	Assays for the measurement of [−2] proPSA and phi are described	Blinded	Cut-off reported
Catalona et al., 2011 [16]	Yes	Yes (Normal DRE, PSA 2–10 μg/L)	Yes	Yes	Yes	Yes	Yes (only for phi)
Guazzoni et al., 2011 [17]	Yes	Yes (Normal DRE, PSA 2–10 μg/L)	Yes	Yes	Yes	Yes	No
Houlgatte et al., 2011 [18]	Yes	Yes (Normal DRE, PSA 2–10 μg/L)	Yes	Yes	Yes	Yes	No
Miyakubo et al., 2011 [19]	Yes	Only according to PSA 2–10 μg/L; No information about DRE	Yes	Age- and prostate volume-adjusted multiple-core biopsies	Yes	No	Yes
Vincendeau et al., 2011 [20]	Yes	Yes (Normal DRE, PSA 2–10 μg/L)	Yes	Yes	Yes	No	No
Jansen et al., 2010 Site 1 (Rotterdam) [15]	Yes	Yes (PSA >4 μg/L or abnormal DRE or abnormal TRUS)[a]	Yes	No	Yes	No	No
Jansen et al., 2010 Site 2 (Innsbruck) [15]	Yes	Yes (according to estimation of prostate cancer risk using ANN)[b]	Yes	No	Yes	No	No
Le et al., 2010 [21]	Yes	Yes (Normal DRE, PSA 2–10 μg/L)	No[c]	Not reported	Yes	Yes	No
Sokoll et al., 2010 [22]	Yes	Yes (Normal and abnormal DRE, PSA 0.29–310.6 μg/L)	Yes	Yes	Yes	Yes	No
Stephan et al., 2009 [23]	Yes	Normal DRE, PSA 0.26–28.4 μg/L	Yes	8–12 cores	Yes	No	No
Sokoll et al., 2008 [24]	Yes	PSA 0.48–33.18 μg/L; No information about DRE	Yes	Yes	Yes	Yes	No
Mikolajczyk et al., 2004 [28]	Yes	PSA 4–10 μg/L; No information about DRE	Yes	Not reported	Yes[d]	Yes	Yes (only for %[−2] proPSA)
Catalona et al., 2003 [29]	Yes	PSA 2–10 μg/L; No information about DRE	Yes	Yes in Innsbruck site; No in Washington site	Yes[d]	Yes	No

Table 1 Quality of 12 studies included in the meta-analysis according to the questionnaire QUADAS.

[a]In 1997, this combination was replaced by PSA testing only; [b]Indication for biopsy based on the estimation of prostate cancer by an artificial neural network (ANN) including PSA, fPSA, age, DRE, and TRUS. In addition, PSA velocity was incorporated in 2005; [c]Biopsy was performed in all patients included for the calculation of the sensitivity and specificity of the tests; [d]Non-commercial assay. DRE, digital rectal examination; TRUS, transrectal ultrasound.

	Sampling frame	Years of recruitment of patients	Population	Age of Patients	Inclusion criteria	Indication for biopsy	Number of cores in biopsy	Patients with biopsy	Patients with cancer	%[−2] proPSA Assay	Algorithms
Catalona et al., 2011 [16]	Multi-center: Prospective and retrospective[a]	2003–2009	Selected	62.8±7.0 (mean±S.D.)	≥50 year, PSA 2–10 μg/L & biopsy	All patients included in the study	89.8% had ≥12 cores; 98% had ≥10 cores	892	430	Beckman Coulter	Phi
Guazzoni et al., 2011 [17]	Prospective	2010	Referral patients/ consecutive	63.3±8.2 (mean±S.D.)	PSA 2–10 μg/L & DRE -	All patients included in the study	18–22 biopsy cores	268	107	Beckman Coulter	Phi
Houlgatte et al., 2011 [18]	Retrospective	Not reported	Selected	Not reported	PSA 2–10 μg/L	All patients included in the study	12 or more cores	452	243	Beckman Coulter	Phi
Miyakubo et al., 2011 [19]	Retrospective	2004–2007	Consecutive	Not reported	PSA 4–10 μg/L	All patients included in the study	Age- and prostate volume-adjusted multiple-core biopsies	239	53	Beckman Coulter	Phi
Vincedeau et al., 2011 [20]	Retrospective	Not reported	Early detection/ selected	Not reported	PSA 2–10 μg/L & DRE -	All patients included in the study	≥10 cores	250	143	Beckman Coulter	Phi
Jansen et al., 2010 Site 1 (Rotterdam) [15]	Retrospective	1994–1997	Screening/non serial	55–75 (66) range (median)	≥50 year, PSA 2–10 μg/L & biopsy[a]	PSA >4, DRE + or TRUS + (In 1997 replaced by PSA only)	6 or more cores	405	226	Beckman Coulter	Phi
Jansen et al., 2010 Site 2 (Innsbruck) [15]	Retrospective	Started in 1993	Screening/non serial	50–77 (69) range (median)	≥50 year, PSA 2–10 μg/L & biopsy[a]	ANN including PSA, fPSA, age, DRE and TRUS (PSA velocity was incorporated in 2005)	6 or more cores	351	174	Beckman Coulter	Phi
Le et al., 2010 [21]	Prospective	2007	Screening/ consecutive	65 (median)	PSA 2.5–10 μg/L & DRE –	PSA ≥2.5 μg/L & DRE +	Not reported	63	26	Beckman Coulter	Phi
Sokoll et al., 2010 [22]	Prospective multicenter	Not reported	Early detection/ consecutive	61.7±8.6 (mean ± S.D.)	>40 year, no prior prostate surgery, biopsy or history of PCa	All patients included in the study	≥10 cores	566 With PSA between 2 and 10 μg/L: 429	245 With PSA between 2 and 10 μg/L: 195	Beckman Coulter	LR including age, race, DRE, prostate cancer family history, log PSA, log %fPSA and log %[−2] proPSA
Stephan et al., 2009 [23]	Retrospective	2002–2006	Referral patients	62.1±5.63 (PCa) 67.2±7.01 (subjects with negative biopsy) (mean±S.D.)	Referred to department of Urology for suspected PCa	All patients included in the study	8–12 cores	586 With PSA between 2 and 10 μg/L: 475	311 With PSA between 2 and 10 μg/L: 264	Beckman Coulter	ANN and LR models including [−2] proPSA, %fPSA, tPSA and age
Sokoll et al., 2008 [24]	Retrospective, multicenter	Not reported	Early detection/ selected	62.2±8.2 (mean±S.D.)	Indication for prostate biopsy	All patients included in the study	≥10 cores	123 With PSA between 2 and 10 μg/L: 89	63 With PSA between 2 and 10 μg/L: 50	Beckman Coulter	LR including PSA, BPSA, %fPSA, %[−2] proPSA, [−2] proPSA/ BPSA, testosterone
Mikolajczyk et al., 2004 [28]	Retrospective	1995–2001	Screening/non serial	66 (median)	PSA 4–10 μg/L	All patients included in the study	Not reported	380	238	Research assay	No
Catalona et al., 2003 [29]	Retrospective, 2 institutions (Innsbruck & Washington)	Innsbruck: 1999–2002 Washington: 1995–2001	Screening/non serial	Not reported	PSA 2–10 μg/L	All patients included in the study	Innsbruck: 10 core biopsy Washington: 6 core biopsy	1091	456	Research assay	No

Table 2 Characteristics of the studies included in the review.

ANN, artificial neural network; CaP, prostate cancer; DRE, digital rectal examination; LR, logistic regression; TRUS, transrectal ultrasound. [a]Only 3.1% were retrospective samples.

	AUC PSA (95% CI)	AUC %fPSA (95% CI)	AUC %[−2] proPSA (95% CI)	AUC phi (95% CI)	Relationship of %[−2] proPSA and Gleason score	Relationship of phi and Gleason score
Catalona et al., 2011 [16]	0.525	0.648	Not reported	0.703	Not reported	Yes The probability of Gleason score ≥7 was 26.1% when phi <25, and 42.1% when phi ≥55.
Guazzoni et al., 2011 [17]	0.53 (0.47–0.59)	0.58 (0.52–0.64)	0.76 (0.71–0.81)	0.76 (0.70–0.81)	%[−2] proPSA was significantly associated with Gleason score (Spearman r: 0.303; p≤0.002), but it did not improve the prediction of Gleason score ≥7 PCa in multivariable accuracy analyses	Phi was significantly associated with Gleason score (Spearman r: 0.387; p<0.002), but it did not improve the prediction of Gleason score ≥7 PCa in multivariable accuracy analyses
Houlgatte et al., 2011 [18]	0.56 (0.51–0.64)	0.59 (not reported)	0.72 (not reported)	0.73 (0.67–0.77)	Not reported	Not reported
Miyakubo et al., 2011 [19]	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported
Vincedeau et al., 2011 [20]	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported
Jansen et al., 2010, Site 1 (Rotterdam) [15]	0.585 (0.535–0.634)	0.675 (0.627–0.721)	0.716 (0.669–0.759)	0.750 (0.704–0.791)	%[−2] proPSA discriminates Gleason score ≥7 (with biopsy Gleason score, p:0.002; with pathologic Gleason score, p:0.09)	Phi discriminates Gleason score ≥7 (with biopsy Gleason score, p:<0.0001; with pathologic Gleason score, p:0.02)
Jansen et al., 2010, Site 2 (Innsbruck) [15]	0.534 (0.473–0.594)	0.576 (0.523–0.629)	0.695 (0.644–0.743)	0.709 (0.658–0.756)	No (neither with biopsy or pathologic Gleason score)	No (neither with biopsy or pathologic Gleason score)
Le et al., 2010 [21]	0.50	0.68	0.76	0.77	Not reported	Not reported
Sokoll et al., 2010 [22]	0.66 (0.62–0.71) For PSA 2–10 μg/L: 0.58 (0.53–0.64)	0.70 (0.65–0.74) For PSA 2–10 μg/L: 0.66 (0.61–0.71)	0.67 (0.62–0.71) For PSA 2–10 μg/L: 0.70 (0.65–0.75)	Not reported LRM[1]: 0.79 (0.75–0.82) For PSA 2–10 μg/L: 0.76 (0.72–0.81)	Yes %[-2] proPSA increased with increasing Gleason score (p<0.001 for all patients and 0.02 for patients with PSA between 2 μg/L and 10 μg/L	Not reported
Stephan et al., 2009 [23]	0.56 (0.51–0.61)	0.77 (0.73–0.81)	0.78 (0.74–0.82)	Not reported (ANN[2]: 0.85; 0.81–0.88) (LR[2]: 0.84; 0.80–0.87)	Yes: %[−2] proPSA is significantly elevated in PCa (p<0.0001)	Not reported
Sokoll et al., 2008 [24]	0.52 (0.42–0.63) For PSA 2 10 μg/L 0.52 (0.40–0.64)	0.61 (0.51–0.71) For PSA 2–10 μg/L 0.53 (0.41–0.65)	0.69 (0.60–0.79) For PSA 2–10 μg/L 0.73 (0.63–0.84)	Not reported LRM[3]: 0.73; 0.64–0.83 For PSA 2–10 μg/L: 0.73 (0.62–0.84)	Not reported	Not reported
Mikolajczyk et al., 2004 [28]	0.526	0.627	0.635	Not reported	Not reported	Not reported
Catalona et al., 2003 [29]	Not reported	0.602	0.638		Not reported	Not reported

Table 3 AUCs for PSA, %fPSA, %[−2] proPSA and phi, and relationship of %[−2] proPSA and phi with Gleason score.

[a]Logistic regression model (LRM) including PSA, BPSA, %fPSA, %[−2] proPSA, [−2] proPSA/BPSA, testosterone; [b]Artificial Neural Network (ANN) and logistic regression (LR) models including %[−2] proPSA, %fPSA, tPSA and age; [c]Logistic regression model (LRM) including age, race, DRE, prostate cancer family history, log PSA, log%fPSA and log %[−2] proPSA. CI, confidence interval.

Table 4A %[–2] proPSA

Studies %[–2] proPSA	TP	FP	FN	TN	Se	Sp
Guazzoni et al., 2011 [17]	96	99	11	62	90%	39%
Miyakubo et al., 2011 [19]	48	139	5	47	90%	25%
Jansen et al., 2010, Site 1 (Rotterdam) [15]	204	122	22	57	90%	32%
Jansen et al., 2010, Site 2 (Innsbruck) [15]	154	117	17	60	90%	34%
Le et al., 2010 [21]	23	19	3	18	88.5%	48.6%
Sokoll et al., 2010 [22]	196	177	49	144	80%	44.9%
Stephan et al., 2009 [23][a]	238	123	26	88	90%	41.7%
Sokoll et al., 2008 [24]	56	38	7	22	90%	37%
Mikolajczyk et al., 2004 [28]	128	152	14	86	90%	36%
Catalona et al., 2003 [29]	410	502	46	133	90%	21%

Table 4B Phi

Studies phi	TP	FP	FN	TN	Se	Sp
Catalona et al., 2011 [16]	387	341	43	121	90%	26.2%
Guazzoni et al., 2011 [17]	96	92	11	69	90%	43%
Houlgatte et al., 2011 [18]	219	149	24	59	90%	28.2%
Miyakubo et al., 2011 [19]	48	125	5	61	90%	33%
Vincendeau et al., 2011 [20]	129	79	14	28	90%	26%
Jansen et al., 2010, Site 1 (Rotterdam) [15]	204	117	22	62	90%	35%
Jansen et al., 2010, Site 2 (Innsbruck) [15]	157	122	17	55	90%	31%
Le et al., 2010 [21]	23	13	3	24	88.5%	64.9%

Table 4 Diagnostic accuracy: sensitivity and specificity. Data were extracted from included studies.
[a]Results for patients with PSA between 2 µg/L and 10 µg/L. FN, false negative; FP, false positive; Se, sensitivity; Sp, specificity; TN, true negative; TP, true positive.

and 10 µg/L. These results underline that %[–2] proPSA may be a useful test in the detection of prostate cancer in men with PSA between 2 µg/L and 10 µg/L.

The derivative test phi showed similar or slightly better results than %[–2] proPSA, with AUCs between 0.703 and 0.77. The performance of other derivative tests obtained by artificial neural network (ANN) or logistic regression (LR) analysis was better than %[–2] proPSA. The best results were reported by Stephan et al. [23] using ANN and logistic regression models with AUCs of 0.85 and 0.84, respectively. According to this author, the ANN model, including %[–2] proPSA, %fPSA, tPSA and age, performs significantly better than %fPSA or %[–2] proPSA, enhancing the Sp of 17%–28% at sensitivities of 90% and 95%.

These results show that the measurement of %[–2] proPSA and phi increases the specificity of the detection of prostate cancer hence reducing the number of unnecessary biopsies. However, information about the recommended cut-offs for these tests were not shown in the majority of papers included in our review. The cut-off level for %[–2] proPSA at Se of 90% was 2.5% for Mikolajczyk et al. [28] and 1.06% for Miyakubo et al. [19]. More similar are the cut-offs suggested for phi by Miyakubo et al. [19] and Catalona et al. [16] showing, respectively that 24.9% and 21.1% of phi corresponds to Se of 90%. Published results showed that while the accuracy of PSA declines with age, the %fPSA increases the predictive value of PSA in older patients [31]. Results communicated by Catalona et al. [16] indicated that phi does not differ by age, and this test may be applicable to young and older men in the detection of prostate cancer.

However, although the unit cost of [–2] proPSA is two to three times higher than both PSA or fPSA, the use of %[–2] proPSA and phi for the detection of prostate cancer decreases global costs. The additional blood test costs were compensated by the savings on the costs of physician office visits and the avoidance of unnecessary biopsies [32, 33].

Several authors showed that %[–2] proPSA and phi may be related to prostate cancer aggressiveness, with higher levels of these tests in patients with Gleason score higher than 7 and in patients with locally advanced tumors [15, 17, 22, 23]. This is relevant information because about one-third of new diagnosed tumors have features of insignificant prostate cancer [34] and these patients can be candidates to active surveillance. However, the identification of these patients using the standard markers, including PSA, biopsy, Gleason score and number of positive biopsy cores, fails to predict accurately the prostate cancer aggressiveness and to choose the more adequate treatment. This point has been confirmed recently by the PIVOT study [35] comparing the effectiveness of radical prostatectomy versus observation in 731 men with localized prostate cancer. The authors showed absolute reductions in all-cause mortality with radical prostatectomy in patients with PSA higher than 10 µg/L and possibly for patients with intermediate- or high-risk tumors, but not in patients with low-risk prostate cancer.

These results underline the usefulness of risk factors in the management of patients with prostate cancer in order to select between a radical treatment and active surveillance. Results reported about %[–2] proPSA and phi suggest that these tests may distinguish low- and high-risk prostate cancer. Using a multivariate analysis, Guazzoni et al. [25] showed that the inclusion of %[–2] proPSA and phi significantly increased the predictive accuracy of a model based on patient age, PSA, %fPSA, clinical stage and biopsy Gleason score in the prediction of high pathologic stage or high pathologic Gleason

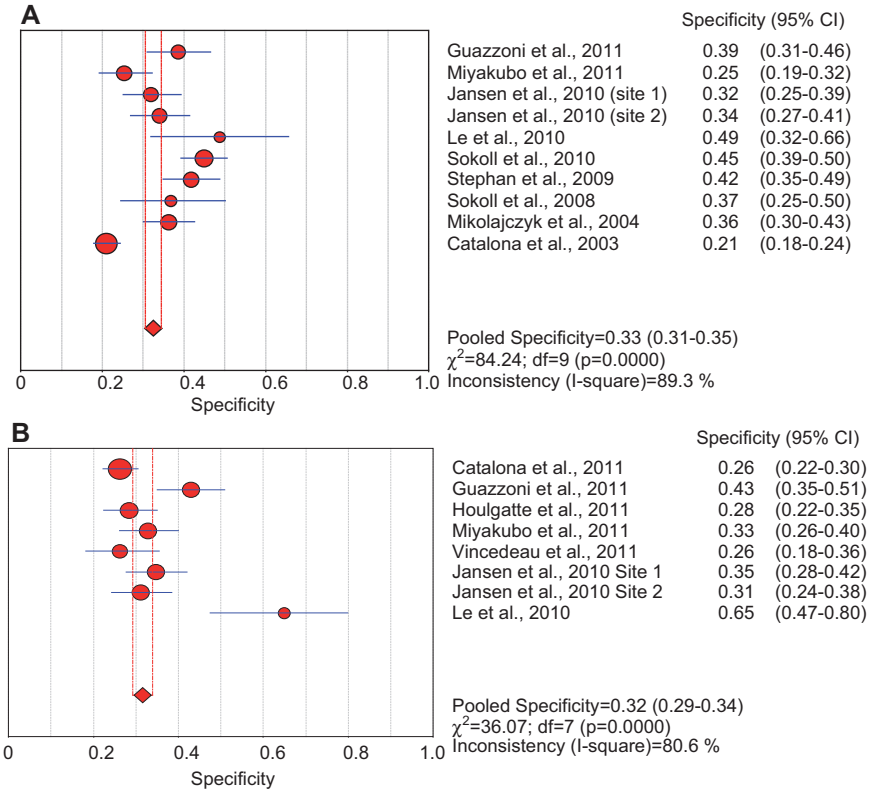

Figure 2 Specificities of %[–2] proPSA and phi. Forest plots showing pooled specificity results of %[–2] proPSA (A) and phi (B). Studies are ordered by author and year of publication. The circles and horizontal lines correspond to the recorded percentage of TN results among patients without prostate cancer and their respective 95% CIs. The area of circles reflects the weight each study contributes to the analysis. The diamond represents the pooled value with its 95% CI.

score. Similarly, de Vries et al. [26] indicated promising results for %[–2] proPSA in selecting treatment strategies for men with prostate cancer using Epstein's criteria to differentiate between non-aggressive and aggressive tumors. Finally, in a recently published study Isharwal et al. [27] described that %[–2] proPSA and phi predicts unfavorable biopsy conversion at an annual surveillance biopsy examination among men enrolled in an active surveillance program. According to this study, the probability of an unfavorable biopsy conversion is higher in patients with %[–2] proPSA higher than 0.7 or with phi higher than 34.2.

Conclusions

The available data shows that %[–2] proPSA and the derivative test phi may be useful in the detection of prostate cancer reducing the number of negative biopsies and improving results obtained with %fPSA and total PSA. Recent published data, concerning cost-effectiveness of these tests also suggests a positive budget impact of their generalized implementation in the management of prostate cancer. Results about the relationship of %[–2] proPSA and phi with the aggressiveness of the tumor corroborate the clinical usefulness of these tests. However, more studies are necessary in order to confirm these data and, specially, in order to define the most appropriate cut-off for %[–2] proPSA and phi.

Acknowledgments: The authors wish to thank Ms. Patricia Vigues for correcting the English version of this article.

Conflict of interest statement

Authors' conflict of interest disclosure: The authors stated that there are no conflicts of interest regarding the publication of this article.
Research funding: None declared.
Employment or leadership: None declared.
Honorarium: None declared.

Received June 25, 2012; accepted October 12, 2012; previously published online November 15, 2012

References

1. Catalona WJ, Partin AW, Slawin KM, Brawer MK, Flanigan RC, Patel A, et al. Use of the percentage of free prostate-specific antigen to enhance differentiation of prostate cancer from benign prostatic disease: a prospective multicenter clinical trial. J Am Med Assoc 1998;279:1542–7.
2. Heidenreich A, Bellmunt J, Bolla M, Joniau S, Mason M, Matveev V, et al. EAU guidelines on prostate cancer. Part 1: screening, diagnosis, and treatment of clinically localised disease. Eur Urol 2011;59:61–71.
3. Roddam AW, Duffy MJ, Hamdy FC, Ward AM, Patnick J, Price CP, 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.
4. Mikolajczyk SD, Marks LS, Partin AW, Rittenhouse HG. Free prostate-specific antigen in serum is becoming more complex. Urology 2002;59:797–802.
5. Mikolajczyk SD, Marker KM, Millar LS, Kumar A, Saedi MS, Payne JK, et al. A truncated precursor form of prostate-specific antigen is a more specific serum marker of prostate cancer. Cancer Res 2001;61:6958–63.
6. Sokoll LJ, Chan DW, Mikolajczyk SD, Rittenhouse HG, Evans CL, Linton HJ, et al. Proenzyme PSA for the early detection of prostate cancer in the 2.5–4.0 ng/ml total PSA range: preliminary analysis. Urology 2003;61:274–6.
7. Khan MA, Sokoll LJ, Chan DW, Mangold LA, Mohr P, Mikolajczyk SD, et al. Clinical utility of proPSA and 'benign' PSA when percent free PSA is less than 15%. Urology 2004;64:1160–4.
8. Filella X, Alcover J, Molina R, Luque P, Corral JM, Augé JM, et al. Usefulness of proprostate-specific antigen in the diagnosis of prostate cancer. Anticancer Res 2007;27:607–10.
9. Stephan C, Meyer HA, Paul EM, Kristiansen G, Loening SA, Lein M, et al. Serum (–5, –7) proPSA for distinguishing stage and grade of prostate cancer. Anticancer Res 2007;27:1833–6.
10. 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.
11. 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.
12. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, 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.
13. 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.
14. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol 2006;6:31.
15. 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.
16. 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.
17. 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.
18. Houlgatte A, Vincendeau S, Desfemmes F, Ramirez J, Benoist N, Bensalah K, et al. Place du –2proPSA et de l'index phi dans la détection précoce du cancer de prostate: évaluation sur unes série de 452 patients. Prog Urol 2011;22:279–83.
19. Miyakubo M, Ito K, Yamamoto T, Suzuki K. Diagnostic significance of [–2]proPSA, total and transition zone prostate volume adjusted PSA-related indices in Japanese men with total PSA in the 2.0 to 10.0 ng/ml range. Eur Urol Suppl 2011;10:65.
20. Vincendeau S, Stephan C, Houlgatte A, Semjonow A. The Beckman Coulter Prostate Health Index (phi) increases the specificity of detection of prostate cancer and reduces the number of negative biopsies. IFCC, WorldLab, EuroMedLab Berlin 2011. Berlin, 15–19 May 2011. Clin Chem Lab Med 2011;49:S874.
21. Le BV, Griffin CR, Loeb S, Carvalhal GF, Kan D, Baumann NA, 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.
22. Sokoll LJ, Sanda MG, Feng Z, Kagan J, Mizrahi IA, Broyles DL, et al. A prospective, multicenter, National Cancer Institute Early Detection Research Network study of [–2]proPSA: improving prostate cancer detection and correlating with cancer aggressiveness. Cancer Epidemiol Biomarkers Prev 2010;19:1193–200.
23. Stephan C, Kahrs AM, Cammann H, Lein M, Schrader M, Deger S, et al. A [–2]proPSA-based artificial neural network significantly improves differentiation between prostate cancer and benign prostatic diseases. Prostate 2009;69:198–207.
24. Sokoll LJ, Wang Y, Feng Z, Kagan J, Partin AW, Sanda MG, 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.
25. 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.
26. de Vries SH, Raaijmakers R, Blijenberg BG, Mikolajczyk SD, Rittenhouse HG, Schröder FH. Additional use of [–2] precursor prostate-specific antigen and "benign" PSA at diagnosis in screen-detected prostate cancer. Urology 2005;65:926–30.
27. Isharwal S, Makarov DV, Sokoll LJ, Landis P, Marlow C, Epstein JI, et al. ProPSA and diagnostic biopsy tissue DNA content

combination improves accuracy to predict need for prostate cancer treatment among men enrolled in an active surveillance program. Urology 2011;77:763.e1–6.
28. Mikolajczyk SD, Catalona WJ, Evans CL, Linton HJ, Millar LS, Marker KM, et al. Proenzyme forms of prostate-specific antigen in serum improve the detection of prostate cancer. Clin Chem 2004;50:1017–25.
29. Catalona WJ, Bartsch G, Rittenhouse HG, Evans CL, Linton HJ, Amirkhan A, et al. Serum pro prostate specific antigen improves cancer detection compared to free and complexed prostate specific antigen in men with prostate specific antigen 2 to 4 ng/ml. J Urol 2003;170:2181–5.
30. Semjonow A, Köpke T, Eltze E, Pepping-Schefers B, Bürgel H, Darte C. Pre-analytical in-vitro stability of [–2]proPSA in blood and serum. Clin Biochem 2010;43:926–8.
31. Vickers AJ, Ulmert D, Serio AM, Björk T, Scardino PT, Eastham JA, et al. The predictive value of prostate cancer biomarkers depends on age and time to diagnosis: towards a biologically-based screening strategy. Int J Cancer 2007;121:2212–7.
32. Nichol MB, Wu J, An JJ, Huang J, Denham D, Frencher S, et al. Budget impact analysis of a new prostate cancer risk index for prostate cancer detection. Prostate Cancer Prostatic Dis 2011;14:253–61.
33. 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.
34. Roemeling S, Roobol MJ, Postma R, Gosselaar C, van der Kwast TH, Bangma CH, et al. Management and survival of screen-detected prostate cancer patients who might have been suitable for active surveillance. Eur Urol 2006;50:475–82.
35. Wilt TJ, Brawer MK, Jones KM, Barry MJ, Aronson WJ, Fox S, et al. Prostate Cancer Intervention versus Observation Trial (PIVOT) Study Group. Radical prostatectomy versus observation for localized prostate cancer. N Engl J Med 2012;367:203–13.

Xavier Filella is doctor of Medicine from the University of Barcelona. He is a specialist in Clinical Biochemistry. At present, he is Senior Consultant in the Department of Biochemistry and Molecular Genetics (CDB) in the Hospital Clinic of Barcelona. He is currently involved in the research of the clinical usefulness of tumor markers, and he is leader of the research line on the study of tumor markers in prostate cancer. He is member of the International Society for Oncology and Biological Medicine and the European Group on Tumor Markers. He has published more than 130 original articles, and he has an accumulated impact factor of more than 390.

Nuria Giménez received her MD and PhD in Medicine from the University of Barcelona and completed a Master of Research methodology in the Autonomous University of Barcelona (UAB). She is a specialist in Clinical Biochemistry and a Research Consultant in the Foundation Research Mútua Terrassa and Associate Professor at the UAB. She is a member of the Evidence Based Laboratory Medicine Commission of the Spanish Society of Clinical Biochemistry and Molecular Pathology (SEQC). Her main fields of research include cancer, toxicology and public health. She has presented over 150 papers in congresses and published over 40 scientific articles in peer reviewed journals.