



## Review – Prostate Cancer

# The Comparability of Models for Predicting the Risk of a Positive Prostate Biopsy with Prostate-Specific Antigen Alone: A Systematic Review

Fritz Schröder<sup>a,\*</sup>, Michael W. Kattan<sup>b</sup>

<sup>a</sup> Department of Urology, Erasmus Medical Centre, Rotterdam, The Netherlands

<sup>b</sup> Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, Ohio, USA

### Article info

#### Article history:

Accepted May 12, 2008

Published online ahead of print on May 22, 2008

#### Keywords:

Prostate cancer  
Screening  
Prostate specific antigen  
Nomogram  
Artificial neural network

### Abstract

**Context:** The sensitivity and specificity profile of measuring levels of prostate-specific antigen (PSA) to select men for prostate biopsy is not optimal. This has prompted the construction of nomograms and artificial neural networks (ANNs) to increase the performance of PSA measurements.

**Objective:** A systematic review of nomograms and ANNs designed to predict the risk of a positive prostate biopsy for cancer was conducted in order to determine their value versus measuring PSA levels alone.

**Evidence acquisition:** Medical Literature Analysis and Retrieval System Online (U.S. National Library of Medicine's life science database; MEDLINE) was searched using the terms "nomogram" "artificial neural network" and "prostate cancer" for dates up to and including July 2007 and was supplemented by manual searches of reference lists. Included studies used an assessment tool to examine the risk of a positive prostate biopsy in a man without a known cancer diagnosis. Intramodel comparisons with evaluation of PSA levels alone, and intermodel comparisons of area under the curve (AUC) from receiver operating characteristic (ROC) curves were conducted. Individual case examples were also used for comparisons.

**Evidence synthesis:** Twenty-three studies examining 36 models were included. With the exception of two studies, all the models had AUC values of 0.70 or greater, with eight reporting an AUC of  $\geq 0.80$  and four (all ANNs) reporting an AUC  $\geq 0.85$ , with variable validation status. Fourteen studies compared the AUC with PSA levels alone: all showed a benefit from using AUCs which varied from 0.02 to 0.26. Of the 16 external validation comparisons, in 13 the AUC was lower in the external population than in the model population.

**Conclusions:** Nomograms and ANNs produce improvements in AUC over measurement of PSA levels alone, but many lack external validation. Where this is available, the benefits are often diminished, although most remain significantly better than with evaluation of PSA levels alone. In men without additional risk factors, PSA cutoff values alone provide a relatively precise risk estimate, but if additional risk factors are known, PSA values alone are less accurate.

© 2008 European Association of Urology. Published by Elsevier B.V. All rights reserved.

\* Corresponding author. Erasmus Medical Centre, Department of Urology (SWOP), P.O. Box 2040, 3000 CA Rotterdam, The Netherlands. Tel. +31 10 4634328; Fax: +31 10 4635315. E-mail address: [m.h.vanderlinde@erasmusmc.nl](mailto:m.h.vanderlinde@erasmusmc.nl) (F. Schröder).

## 1. Introduction

In the last 20 years, the use of prostate-specific antigen (PSA) in screening for prostate cancer has become increasingly widespread in the Western world. However, limitations in the sensitivity and specificity of PSA values for the prediction of prostate cancer are well understood [1]. Although the balance of detection at different PSA levels remains intensely debated, the curvilinear relationship with risk and the lack of a clear value at which sensitivity and specificity are optimal, makes selecting a threshold level problematic [2]. Higher PSA levels used as a biopsy indicators increase specificity, but many cases of cancer are missed. At lower thresholds, sensitivity is improved, but the number of unnecessary biopsies is increased [1].

One approach to predicting the likelihood of a positive prostate biopsy is to combine PSA levels with other risk factors within a nomogram or artificial neural network (ANN). The former have been widely utilised for a number of roles in prostate cancer, including predicting cancer recurrence, seminal vesicle extension, extracapsular extension, lymph node invasion, and Gleason grade upgrading from biopsy to prostatectomy [3–5]. ANNs are computer-based systems that aim to mimic the learning process of a low-level operation of neural circuits in the brain. Data on risk factors and outcomes are fed into the system by a training programme that creates a system of weighting for these factors. Test cases are run to check the network, and, when it is complete, a validation set is also run. Comparative studies between regression-derived and ANN-derived predictive models show that they may differ in their accuracy in particular circumstances, with the latter being comparable to or inferior to Cox regression [6]. The aim of this review is to critically examine nomograms and ANNs designed to predict the risk of a positive prostate biopsy for cancer and to determine their value over and above evaluation of risk using PSA levels alone in clinical practice.

## 2. Methods

### 2.1. Type of studies considered in this review

Studies considered appropriate for inclusion were those that used a risk assessment tool to examine the prospective risk of a positive prostate biopsy in a man without a known prostate cancer diagnosis. Nomograms or ANNs were included if they utilised three or more variables for risk assessment: those that examined only two variables, such as age-modified PSA thresholds, were not included.

Although analyses have also been conducted through the construction of risk groups, where each variable is divided into discrete groups (eg, a PSA level of 0–2.0 ng/ml, 2.1–4.0 ng/ml etc), with groups being combined (eg, age <50 and a PSA level of 2.1–4.0 ng/ml), this type of assessment has limitations. It is difficult to demonstrate that individuals within these groups have a homogeneous level of risk, although this is dependent on the range of variables used [7,8]. In addition, these approaches do not easily allow flexibility in choosing cutoff levels of risk. For these reasons, this particular approach is not examined in this review.

### 2.2. Search strategy

MEDLINE was searched using the terms “nomogram,” “artificial neural network,” and “prostate cancer” for dates up to and including July 2007. Further studies were selected on the basis of manual searches of reference lists and review articles [7,9–11]; this approach was utilised because an earlier review of prostate cancer nomograms and ANNs identified the majority of appropriate studies by manual searching rather than by a database search alone [9].

### 2.3. Intramodel comparison

In order to compare the additional value of individual models over PSA levels alone, the accuracy of the models was compared with stated values for PSA within each study.

### 2.4. Intermodel comparison

**2.4.1. Comparison of parameters used to generate risk levels**  
The majority of analyses conducted to generate risk models utilise Cox or logistic regression techniques. The models are then typically refined by removal of variables that are not shown to have an independent, statistically significant effect on risk. For the purposes of this review, a comparison of the variables considered in the regression analyses, and those that were included in the final model, was conducted.

#### 2.4.2. Comparison of model accuracy

Estimates of area under the curve (AUC) from receiver operating characteristic (ROC) curves for each model and for PSA values alone were extracted from each publication for comparison, where available. This was done in accord with the view that a variable directly assessing classification accuracy should be utilised for marker comparisons [12]. Where AUC values were unavailable, other parameters such as sensitivity and specificity were examined.

#### 2.4.3. Construction of risk curves for finding prostate cancer at biopsy

Estimation of risk of a positive biopsy was calculated using different values for the variables defined in the models. Where these were available, values for PSA across the range 0.1–10 ng/ml were entered. For each PSA value, percentage risk values were modified by the inclusion of other risk factors where these could be entered: (1) age (40–70 yr); (2) digital rectal examination (DRE; suspicious or not suspicious); (3) previous

negative biopsy (yes or no); (4) family history of prostate cancer (present or absent). For each model, two curves were plotted for PSA level against percentage risk: the first represented the lowest levels of risk for any given PSA level, and the second represented the highest level of risk. This allowed an assessment of the range of risk for any given PSA level.

#### 2.4.4. Case comparison

Data from a series of eight individual patient cases were entered into selected models for a 'real world' view of model comparability.

#### 2.4.5. External validation

Where models were validated in populations beyond those in which they were derived, the AUC outcomes were tabulated for comparison. Additional comparisons with outcomes using PSA levels alone in these populations were also tabulated where available.

### 3. Results

#### 3.1. Included studies

A total of 23 studies (26 publications) were included in the review [13–39]. These studies examined 36 different nomograms and ANNs. The characteristics of the models and the populations on which they are based are outlined in Table 1.

The populations used to construct the risk models differed significantly among studies. Of the 23 studies, 5 studies included men from screening studies [13,18–20,26,27], 2 studies were of mixed populations [35,37], 15 studies were from referral populations [14–17,21–25,28,29,32,34,36,38,39], and 1 study was of a longitudinal follow-up population [33]. Of the 15 referral population studies, 2 studies exclusively examined men who had had a previous negative biopsy [23,24]. In all but three of the studies [13,16,18], all men underwent prostate biopsy. The majority of studies therefore examined men with at least one reason to undergo a prostate biopsy; that is men above a certain PSA threshold, men with a positive DRE, or clinically apparent disease. Conversely, tumours in men below the PSA threshold, without a suspicious DRE, or without clinical suspicion of prostate cancer were not biopsied. A prominent exception to this was the Prostate Cancer Prevention Study (PCPT), where all men were offered an end-of-study biopsy [32].

The sample sizes varied from 151 to 8851 and were derived mostly from US cohorts, but included populations from Europe and Japan. Ethnicity and family history data, where available, showed substantial variance in the nature of the individual cohorts. The risk of a positive biopsy also varied from 9% to 44%. A history of screening prior to the

current evaluation was absent in 17 of the studies: of the remaining 6 studies, 4 studies contained men undergoing first-round screening [13,18–20] and two studies contained previously biopsied men [23,24].

The number and nature of included variables differed substantially among the models examined (Table 2). The simplest models utilised 3 variables, and the most complex used 10 variables. All of the models included PSA as a variable: this was expressed as a measure of total PSA in all but two models. The range of PSA values available in the studies varied: some examined risk in men with a PSA level below typical biopsy thresholds, others examined risk above them, whilst some traversed the two. Age and DRE status were the next most commonly used variables. In 16 of the models, all the variables tested in any initial model were utilised in the final model: in the 7 remaining models some variables were removed. Of these, age, family history, and the velocity of increase in PSA level (PSA velocity) were the most frequent variables to lack independent significance in the final models.

Three of the studies also included additional models which examined the risk of a high-grade tumour (Gleason score  $\geq 7$ ) [27,30–32,38]. Two studies examined the issue of whether biopsy core number was a predictive factor for cancer detection [29,39]; two additional studies were conducted in biopsy-negative populations and included the previous number of negative cores as a variable [23,24].

#### 3.2. Model accuracy

A comparison of model accuracy is presented in Table 3. With the exception of two studies [33,36], the models examined that reported AUC outcomes had values of 0.70 or greater, with eight reporting an AUC of  $\geq 0.80$  [13,14,19,21,28,34,35,37] and four (all ANNs) reporting an AUC  $\geq 0.85$  [13,19,35,37]. Of the eight models with an AUC  $\geq 0.85$ , four were internally validated on the same population used for the model construction, two were validated in a group from the model database set aside for this purpose, and two were externally validated. Of the 23 studies examined, a direct AUC comparison with PSA alone (total PSA or another PSA assessment) was available for 14 of the studies [14,16,19,21,22,24,28,32–35,37–39]: all showed a benefit from the nomogram or the ANN tested. The magnitude of the increase in AUC over PSA level alone varied substantially between models, from 0.02 in the PCPT [32] to 0.26 in the study of Stephan et al [37]. However, when the type of ANN used in this study was re-evaluated in comparison with a nomogram, the AUC dropped from 0.84 to 0.67 [36].

Table 1 – Characteristics of models and the populations on which they were based

Study	Type	n	Location	Patients	Validation	Mean age (yr)	Mean PSA (ng/ml)	White (%)	Family history (%)	All men biopsied	Pos. biopsy (%)	Data for analysis	Parameter range over which risk calculable
Snow et al 1994 [13]	Neural network	1787	Washington, DC, USA	PSA >4.0 ng/ml and positive USS and/or DRE; underwent biopsy in a first-round screen	Constructed and validated against the “PSA-1” database	62.6	N/A	N/A	N/A	No	33.5 in database; 27 in test group	N/A	Age 50–90 yr; PSA >4.0 ng/ml
Optenberg et al 1997 [14]	Logistic regression	703 633	Model pop: Texas, USA Test pop: Seattle, WA, USA	Biopsy population; no other details; round of screening unclear	Retrospective analysis against further clinic database; “CAPRI” test	Model: 64.5; test: 67.1	Model: 7.39; test: 14.72	Model: 88.8; test: 93.9	N/A	Yes	Model: 23.4; test: 29.1	Derived from plots	Age 35–85 yr; PSA ≤20 g/ml
Carlson et al 1998 [15]	Logistic regression	3773 525	Model pop: Baltimore, MD, USA Test pop: USA	Age ≥45 yr; no history of PCa; PSA 4–20 ng/ml; sextant biopsy; round of screening unclear	Retrospective analysis validated against further database	Model: ~67	Model: ~8	N/A	N/A	Yes	Model: 32.7	Specific values only from tables	Age ≥45 yr; PSA 4–20 ng/ml
Babaian et al 2000 [16]	3 neural networks compared with PSAD, PSAD-TZ and %fPSA	151	Houston, TX, USA	Age 40–75 yr; PSA 2.5–4.0 ng/ml; no previous biopsy; 11-core biopsy; round of screening unclear	Comparison of single variables vs neural network	N/A	N/A	N/A	N/A	No (1998); yes (2000)	24.5	N/A	Age 40–75 yr; PSA 2.5–4.0 ng/ml
Eastham et al 1999 [17]	Logistic regression nomogram presented	700	Shreveport, LA, USA	Abnormal DRE and PSA <4 ng/ml; sextant biopsy plus extra cores if needed; round of screening unclear	Internal bootstrapping	~63	~1.7	63.6	N/A	Yes	9	Yes	PSA <4 ng/ml; DRE abnormal

**Table 1 (Continued)**

Study	Type	n	Location	Patients	Validation	Mean age (yr)	Mean PSA (ng/ml)	White (%)	Family history (%)	All men biopsied	Pos. biopsy (%)	Data for analysis	Parameter range over which risk calculable
Finne et al 2000 [18]	Logistic regression vs neural network	656	Finland	PSA 4–10 ng/ml; age 55–67 yr; sextant biopsy; 1st-round screening	Retrospective analysis of men with PSA $\geq 4$ who had a biopsy in the Finnish screening study	~62	~5.5	N/A	6	No	27.1	N/A	Age 55–67 yr; PSA 4–10 ng/ml
Djavan et al 2002 [19]	Neural network	1246	Vienna, Austria	Age 31–89 yr; no FH of PCa; PSA 2.5–4 ng/ml and 4–10 ng/ml (two groups combined); sextant biopsy; 1st-round screening	Prospective analysis of men with PSA $\geq 2.5$ who had a biopsy in the EPCDS; sample split into training, test, and validation groups	~67	~3.1 in the 2.5–4 group and 6.7 in the 4–10 group	100	N/A	Yes	2.5–4: 31; 4–10: 35	N/A	Age 31–89 yr; No FH of PCa; PSA 2.5–10 ng/ml
Finne et al 2002 [20]	Logistic regression Nomogram	758	Finland	PSA 4–20 ng/ml; age 55–67 yr; sextant biopsy; 1st-round screening	Retrospective analysis of men with PSA $\geq 4$ who had a biopsy in the Finnish screening study	~62	~5.2	N/A	~6.7	Yes	26.4	Derived from plots	Age 55–67 yr; PSA 4–20 ng/ml
Stephan et al 2002 [21]	Neural network and %fPSA alone	1188	Germany, Netherlands, Canada	PSA 2–20 ng/ml; sextant or octant biopsy; screening round unclear	Retrospective analysis of men with PSA 2–20 ng/ml who had a biopsy following referral for possible PCa or LUTS	~65	N/A	N/A	N/A	Yes	N/A	N/A	Age 40–86 yr; PSA 2–20 ng/ml
Garzotto et al 2003 [22]	Logistic regression Nomogram	1239	Portland, OR, USA	PSA $\leq 10$ ng/ml; sextant biopsy; screening round unclear	Men referred for biopsy 20% of sample used for validation, 80% for nomogram construction	66 (med)	4.93	93.5	16.9	Yes	24.3	Yes	PSA $\leq 10$ ng/ml

Lopez-Corona et al 2003 [23]	Logistic regression Nomogram	343	New York, USA	≥1 previous negative biopsy conducted due to HGPIN, ASAP, positive DRE, persistently elevated PSA, PSA slope >0.75 ng/ml yearly; 6-22 biopsy cores	Men undergoing repeat biopsy	62.1	8.4	N/A	17	Yes	N/A	Yes	N/A
Yanke et al 2005 [24]	Logistic regression Nomogram	230	Test pop: New York, USA	External validation of Lopez-Corona et al 2003 [23]; sextant or 12-core biopsy dependent on when biopsied between 1993 and 2003; screening round unclear	Model constructed on data from one centre and validated against data from another	66.3	10.3	38.2	10	Yes	33.9	Yes	N/A
Karakiewicz et al 2005 [25]	Logistic regression nomograms ×2	41 93 1 762 514	Three cohorts: Montreal, Canada; Hamburg, Germany; Montreal, Canada	PSA < 50 ng/ml; men referred for sextant biopsy; screening round unclear	3 cohorts of men referred for evaluation with suspicious PSA/DRE who underwent sextant biopsy: 1st cohort for development of 1st nomogram, 3rd for its validation; 3rd cohort for development of 2nd nomogram, validated on 2nd cohort	Cohort 1: 64.2; cohort 2: 64.8; cohort 3: 63.2	Cohort 1: 9.8; cohort 2: 8.9; cohort 3: 8.1	N/A	N/A	Yes	Cohort 1: 35.2; cohort 2: 41.9; cohort 3: 36.8	Yes	PSA <50 ng/ml

Table 1 (Continued)

Study	Type	n	Location	Patients	Validation	Mean age (yr)	Mean PSA (ng/ml)	White (%)	Family history (%)	All men biopsied	Pos. biopsy (%)	Data for analysis	Parameter range over which risk calculable
Nam et al 2006 [27,43]	Logistic regression recursive partitioning modelling with 6 risk groups created nomogram	2637	Toronto, Canada	PSA >2.5 ng/ml or abnormal DRE; 6–15 core biopsy; examination of any and aggressive PCa; referred from screening programme, but no details of round	Retrospective analysis of men with PSA >2.5 ng/ml who had one or more biopsies in the Toronto screening study; 2000 bootstraps run	~65	~11	~50	~50	Yes	44.2 at initial biopsy, 28.4 of those with repeat biopsy	Yes	Age 40–94 yr; PSA >2.5 ng/ml
Suzuki et al 2006 [28]	Logistic regression nomogram	834	Chiba, Japan	Japanese men referred for prostate biopsy; screening round unclear	20% of sample used for validation, 80% for nomogram construction	70	13.1	Japanese	N/A	Yes	28.9	Yes	N/A
Yanke et al 2006 [29]	Logistic regression nomogram	8851	Shreveport, LA; Detroit, MI; and New York, NY, USA	Men referred for prostate biopsy; nomogram designed for African American men	Bootstrapping with 200 samples	~65	Cohort 1: 10.3; cohort 2: 11.8; cohort 3: 11.1	40.2	N/A	Yes	33, African American; 26, Caucasian	Yes	N/A
PCPT 2006 [30–32]	Logistic regression nomogram	5519	Multicentre USA	PSA ≤3.0 ng/ml; age ≥55 yr; ≥sextant biopsy; many men likely to have been previously screened	Placebo group from PCPT; model constructed and validated against same set	N/A	1.5 (median)	95.6	16.7	Yes	All: 21.9; HGT: 4.7	Yes	Age ≥55
Parekh et al 2006 [33]	Logistic regression nomogram	446	San Antonio, TX, USA	Referred biopsy population; screening round unclear	External validation of PCPT nomogram[31]	N/A	2.8 (median)	~80	30.7	Yes	33.2	PCPT model used	N/A
Stephan et al 2006 [34]	Neural network ×3, and individual consideration of %fPSA, proPSA, and proPSA/%fPSA	898	Aarau, Switzerland and Berlin, Germany	Men with known PCa and those with BPH, chronic prostatitis or normal tissue on biopsy; round of screening unclear	Model constructed and validated against same set: men with and without PCa from screened populations	N/A	N/A (range: 1–10)	N/A	N/A	Yes	N/A	Yes	PSA 1–10 ng/ml

Stephan et al 2007 [35]	Neural network x7 (including one previously tested) [21] derived from and compared with PSA and %fPSA derived from different assays	4480	Hamburg, Germany Vienna, Austria Rotterdam, Netherlands Münster, Germany Helsinki, Finland	Mixed population of referred men and those from screening studies	Model constructed and validated against same set: men with and without PCa from screened populations	N/A	N/A (range 2–10)	N/A	N/A	Yes	35.4	N/A	Age 31–91 yr; PSA 2–10 ng/ml
Chun et al 2007 [36]	Comparison of neural network and nomogram	3980	Hamburg, Germany	Referred population with an elevated PSA and/or abnormal DRE; round of screening unclear	Comparison of existing ANN [21] and nomogram [25] in the same population	64.8	7.2	N/A	N/A	Yes	42.4	N/A	Age 34–85 yr; PSA 2–20 ng/ml
Stephan et al 2007 [37]	Construction and testing of multiple ANNs tested against each other and PSA	656 and 606	Helsinki, Finland and a multicenter group	Screening population (Finland) and referred population (multicentre)	External validation	63–65	5.4–6.7	N/A	N/A	Yes	22.5–61.7	N/A	Age 55–67 yr; PSA 4–10 ng/ml
Nam et al 2007 [38]	Logistic regression nomogram	2700	Toronto, Canada	Referred population	Internal validation	63–65	8.1–9.7	43–57	48–52	Yes	42	Yes	Age 40–94 yr; PSA 0.05–48.7 ng/ml
Chun et al 2007 [39]	Logistic regression nomogram	1262, 2900, and 1738	Milan, Italy, Hamburg, Germany, Seattle, WA, USA	Referred population	Internal validation of 10-core nomogram; external validation of Karakiewicz 2005 in a 10-core biopsy population [25]; external validation of 10-core nomogram	65	6.8–8.2	N/A	N/A	Yes	37.9–46.2	Yes	Age 37–92 yr; PSA 0.1–48.9 ng/ml

PSA: prostate-specific antigen; N/A: not available; USS: ultrasound scan; PCa: prostate cancer; PSAD; PSA density; PSAD-TZ: PSAD transition zone; %fPSA: percent free PSA; DRE: digital rectal examination; FH: family history; EPCDS: epicardium-derived cells; LUTS: lower urinary tract symptoms; HGPN: high-grade prostatic intraepithelial neoplasia; ASAP: atypical small acinar proliferation; PCPT: Prostate Cancer Prevention Study; BPH: benign prostatic hyperplasia; ANN: artificial neural network; CAPRI: Casodex vs placebo in non-metastatic early prostate cancer.

Table 2 – Variables used in model construction

Study	Age	PSA	DRE	Ethnicity	Prev. neg. biopsy	%fPSA	TRUS +	TRUS PV	FH	Other
Snow et al 1994 [13]	-	-	-				-			PSA velocity
Optenberg et al 1997 [14]	-	-	-	-						
Carlson et al 1998 [15]	-	-	-			-				
Babaian et al 2000 (3 ANNs) [16]	-	-	-			-				PAP, CK
Eastham et al 1999 [17]	-	-	-	-						
Finne et al 2000 [18]	-	-	-			-		-	-	
Djavan et al 2002 [19]	-	-	-			-		-		PSA-TZ, PSAD, TZV, PSAV
Finne et al 2002 [20]	-	-	-			-		-		
Stephan et al 2002 [21]	-	-	-			-		-		
Garzotto et al 2003 [22]	-	-	-	-			-	-	-	PSAD, prostate asymmetry, prior vasectomy
Lopez-Corona (previously biopsied population) [23]	-	-	-		All				-	PSA vel., neg cores, HG-PIN, ASAP, months from first and prev. neg. biopsy
Yanke et al 2005 (previously biopsied population) [24]	-	-	-		All				-	PSA vel., neg cores, HG-PIN, ASAP, months from first and prev. neg. biopsy
Karakiewicz et al 2005 N1 [25]	-	-	-							
Karakiewicz et al 2005 N2 [25]	-	-	-			-				
Nam et al 2006 (any PCa) [27]	-	-	-	-	-			-	-	LUTS
Nam et al 2006 (aggressive PCa) [27]	-	-	-	-	-			-	-	LUTS
Suzuki et al 2006 [28]	-	-	-			-		-		
Yanke et al 2006 [29]	-	-	-	-						Year of biopsy, total number of cores
PCPT et al 2006 (any PCa) [30–32]	-	-	-	-	-				-	PSA velocity
PCPT 2006 (aggressive PCa) [30–32]	-	-	-	-	-				-	
Parekh et al 2006 (external validation of PCPT) [33]	-	-	-	-	-				-	
Stephan et al 2006 [34]		-	-			-		-		proPSA, proPSA/fPSA, proPSA/%fPSA
Stephan et al 2007 [35]	-	-	-			-		-	-	
Chun et al 2007 [36]	-	-	-			-		- (ANN only)		
Stephan et al 2007 [37]	-	-	-			-		-		
Nam et al 2007 [38]	-	-	-	-		-			-	LUTS score
Chun et al 2007 (10-core biopsy) [39]	-	-	-			-				Core sampling density

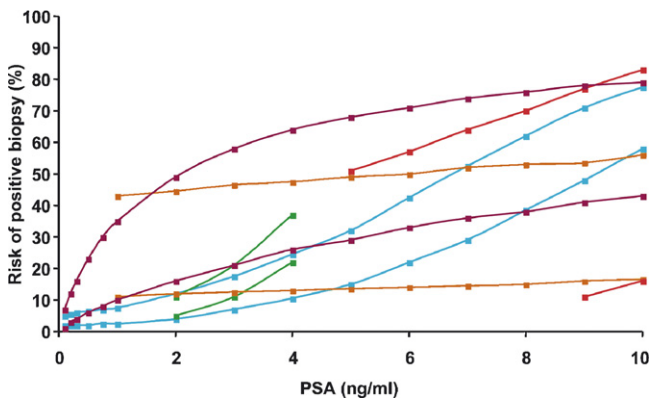
Crosses, and letters in italics, represent analysed variables that were not included in the final model due to lack of statistical significance as independent variables.

PSA: prostate-specific antigen; DRE: digital rectal examination; %fPSA: percent free PSA; TRUS: transrectal ultrasound; TRUS-PV: TRUS prostate volume; FH: family history; PAP: prostatic acid phosphatase; CK: creatine kinase; PSAD: PSA density; PSA-TZ: PSAD transition zone; TZV: transition zone volume; PSAV: PSA velocity; HG-PIN: high-grade prostatic intraepithelial neoplasia; ASAP: atypical small acinar proliferation; LUTS: lower urinary tract symptoms; PCPT: Prostate Cancer Prevention Study; proPSA: subform of fPSA.

**Table 3 – Model area-under-the-curve (AUC) values versus prostate-specific antigen (PSA) levels alone, where reported**

Study	AUC for model	AUC for PSA	Increase in AUC vs PSA levels alone
Snow 1994 [13]	0.87	N/A	N/A
Optenberg 1997 [14]	0.81	0.78	0.03
Carlson 1998 [15]		11% increase in sensitivity vs %fPSA alone	
Babaian et al 2000 [16,57]	0.74	%fPSA: 0.64; PSAD: 0.74	0.10
Eastham et al 1999 [17]	0.75	N/A	N/A
Finne et al 2000 [18]		6–8% increase in specificity at a sensitivity of 80–90% over f/tPSA alone	
Djavan et al 2002 [19]	PSA 2.5–4.0 ng/ml: 0.88; PSA 4.0–10.0 ng/ml: 0.91	N/A	ANN AUC greater than AUC for PSA; numbers not given
Finne et al 2002 [20]		Significant reduction in false positive rate	
Stephan et al 2002 [21]	PSA 2.5–10.0 ng/ml: 0.84	PSA 2.5–10 ng/ml: 0.66	0.18
		%fPSA 2.5–10.0 ng/ml: 0.71	0.13
Garzotto et al 2003 [22]	0.73	0.62	0.11
Lopez-Corona et al 2003 [23]	0.70	N/A	N/A
Yanke et al 2005 [24]	0.71	0.64	0.07
Karakiewicz et al 2005 [25]	N1: 0.69; N2: 0.77	N/A	N/A
Nam et al 2006 [27,43]	Any PCa: 0.77; aggressive PCa: 0.74	N/A	N/A
Suzuki et al 2006 [28]	0.82	0.70	0.12
Yanke et al 2006 [29]	0.75	N/A	N/A
PCPT 2006 [30–32]	Any PCa: 0.70 Aggressive PCa: 0.70	0.68	0.02
Parekh et al 2006 [33]	0.66	0.64	0.02
Stephan et al 2006 [34]	PSA 1.0–10.0 ng/ml: 0.77–0.82	PSA (tPSA) 1.0–10.0: 0.66	0.11–0.16
		PSA (proPSA/%fPSA) 1.0–10.0: 0.74	0.03–0.08
Stephan et al 2007 [35]	0.77	0.56	0.21
	0.86	0.64	0.22
	0.85	0.80	0.05
	0.88	0.74	0.14
	0.90	0.77	0.13
	0.77	0.56	0.21
Chun et al 2007 [36]	ANN: 0.67; nomogram: 0.71	N/A	N/A
Stephan et al 2007 [37]	0.83	0.57	0.26
Nam et al 2007 [38]	0.74 (all tumours); 0.77 (high-grade)	0.62 <sup>*</sup> 0.69 <sup>*</sup>	0.12 0.08
Chun et al 2007 [39]	0.77 (internal validation); 0.75, 0.73, 0.76 (external validation at 3 sites)	61	0.12–0.16

N/A: not available; %fPSA: percent free PSA; PSAD: PSA density; tPSA: total serum PSA; ANN: artificial neural network; PCa: prostate cancer; proPSA: subform of PSA; DRE: digital rectal examination.  
<sup>\*</sup> PSA and DRE.



**Fig. 1 – Comparison of risk curves predicting positive biopsies for prostate cancer from five nomograms. Purple: PCPT [31]; red: Nam et al [27,43]; blue: Optenberg et al[14]; green: Eastham et al [17]; orange: Karakiewicz et al [25]. See Methods section for description of upper and lower curves for each model.**

Of the three studies not reporting AUC values, one study reported an increase in sensitivity [15], a second study reported an increase in specificity [18], and the last study reported a reduction in the rate of false negatives versus that of PSA values alone [20]. The AUC values for the three studies that examined high-grade tumours separately were similar to those for models examining all tumours [27,30–32,38].

**3.3. Risk curves for finding prostate cancer at biopsy**

A comparison of risk levels conducted among five models is presented in Fig. 1. For any given PSA level, the upper and lower limits of risk vary substantially among the models. Furthermore, the overall shape of the curves differs substantially among models: some are linear, others are logarithmic, and some are sigmoidal.

**3.4. Case comparison**

A comparison of risk values from patient cases is shown in Table 4. The specific risk values generated from different models vary substantially. For example, for a 55-year-old Caucasian man with a PSA level of 2.4 ng/ml, there is a three-fold variation in risk level across the three models tested. In addition, ethnic status substantially effects risk in one model, has a lesser effect in another, but has no effect in the remaining two.

**3.5. External validation**

Of the models examined, ten were externally validated (Table 5) [14,15,22,24,25,31–33,35–37,39].

**Table 4 – Cases and positive biopsy risk levels provided by different models for individual patient cases**

Case	Age (yr)	Race	PSA	FH	DRE suspicious	Prev. neg. biopsy	Eastham et al 1999 [17]	Karakiewicz et al 2005 [25]	Nam et al 2006 [27,43]*	Yanke et al 2006 [29]	PCPT 2006 [31]
1	40	African American	2.5	Yes	No	No	8.0%	11.0%	N/A	28%	N/A
2	55	Caucasian	0.5	No	No	No	N/A	15.0%	N/A	27%	8%
3	55	Caucasian	2.4	No	No	No	9.0%	16.0%	N/A	29%	26%
4	55	Caucasian	2.4	Yes	No	No	9.0%	16.0%	N/A	29%	31%
5	60	Caucasian	3.5	No	No	No	22.5%	17.0%	N/A	32%	32%
6	60	Caucasian	3.5	Yes	Yes	No	22.5%	40.0%	N/A	44%	61%
7	60	African American	3.5	No	No	No	24.0%	17.0%	N/A	39%	32%
8	60	African American	3.5	Yes	Yes	No	24.0%	40.0%	N/A	52%	61%

PSA: prostate-specific antigen; FH: family history; DRE: digital rectal examination; PCPT: Prostate Cancer Prevention Study; N/A: nomogram was unable to provide result because the parameters were outside of range.  
 \* Assumes 40-ml prostate and no symptoms; risk values below a PSA of 10.0 ng/ml unclear.

**Table 5 – Changes in area-under-the-curve (AUC) values for models applied to external populations and comparison with values for prostate-specific antigen (PSA)**

Study	Difference in AUC between external and model populations	Difference in AUC between nomogram/ANN and PSA in external population
Optenberg et al 1997 [14]	–0.05	N/A
Carlson et al 1998 [15]	N/A	N/A
Garzotto et al 2003 [22]	N/A	N/A
Yanke et al 2003* [24]	+0.01	N/A
Karakiewicz et al 2005 [25]	+0.01	N/A
(2 nomograms examined)	–0.01	
Parekh et al 2006 [33]	–0.03	+0.02
Stephan et al 2007† [35]	–0.11	+0.17
(6 population comparisons)	–0.01	+0.13
	–0.05	–0.01
	–0.01	+0.19
	0	+0.07
	–0.09	+0.19
Chun et al 2007‡ [36]	–0.17	N/A
(ANN and nomogram)	–0.06	
Stephan et al 2007† [37]	–0.10	+0.18
(comparison with a single centre and a multicentre cohort)	–0.01	+0.26
Chun et al 2007 [39]	–0.02	N/A
	–0.04	
	–0.01	

N/A: not available; ANN: artificial neural network.  
\* External validation of Lopez-Corona 2003 [23].  
† External validation of Stephan et al 2002 [21].  
‡ External validation of Karakiewicz 2005 [25].

Of the 16 comparisons available, in 13 comparisons the AUC was lower in the external population compared with the model population; in 2 comparisons it was increased by 0.01, and in 1 comparison it was the same. The greatest variations were observed with an ANN [21]. For nine models, data comparing the AUC for PSA levels alone with the models examined were available from external populations (Table 5). The change to the AUC ranged between –0.01 and +0.26.

#### 4. Discussion

A key question concerning the use of any model to prospectively determine the risk of a positive biopsy for prostate cancer is whether it provides a more accurate method of discerning risk versus single risk factors alone. In reality, this means it needs to perform better than measurement of PSA levels as a marker of risk. PSA level has a positive, curvilinear relationship with prostate cancer risk [40] if a previous value has not been applied as a biopsy indication [41]; however, the AUC value for PSA level alone is lower than that of many commonly employed screening tests for malignancy [42]. In this review, we conducted intrastudy comparisons

of accuracy in predicting positive biopsies between AUC values from models and PSA levels alone to examine levels of benefit of nomograms and ANNs.

Data from six studies demonstrate favourable AUC values for the models constructed [13,17,25,27,29,43], but because no value for PSA level alone was available for comparison, the value of these models is unclear. Twelve studies did, however, report benefits of prediction models over PSA measurement alone, with benefits of AUC varying from 0.02 [32] to 0.26 [37]. External validation had been conducted in only five of these models, with the remainder internally validated. On average, the models examined added ~0.10 to the AUC over using PSA levels alone. It is interesting to note that the smallest incremental benefit (0.02) for a nomogram on AUC amongst these studies was observed in one of the largest and the most systematic studies, which was based on >5500 men who were biopsied regardless of PSA value and DRE status [32]. One of the reasons for this small size of this benefit may be the impact on the predictive value of PSA level on systematically biopsied patients, as opposed to the use of a cutoff level. It is also evident that the detection of prostate cancer in men with low PSA values (<3.0 ng/ml) may have limited relevance due to the overdiagnosis of indolent tumours [44]. With

regard to the AUC associated with ANNs, of those where a comparison with PSA levels was available, the average increase in AUC was of a magnitude higher than that observed for nomograms ( $\sim 0.15$ ). However, a recent comparison of one of these ANNs demonstrated that it was marginally inferior to a nomogram [36]. A comparison of ANNs and Cox regression methods for prediction conducted in 2003 found the latter to be similar to or inferior to the former [6].

It is likely that methodological issues, as well as the populations examined, resulted in the observation of wide variations in AUC improvement over PSA values alone [45]. One key issue is the choice of variables for inclusion in the models, which appears largely to be driven by data availability. If, for example, PSA density (PSAD), complexed PSA (cPSA) or percent free PSA (%fPSA) confer advantages in accuracy over total PSA level alone, as has been reported [46,47], failing to test these parameters within a model will limit likely improvements in accuracy. The omission of prostate volume from the PCPT predictive model may explain the modest increase in AUC observed over PSA level alone.

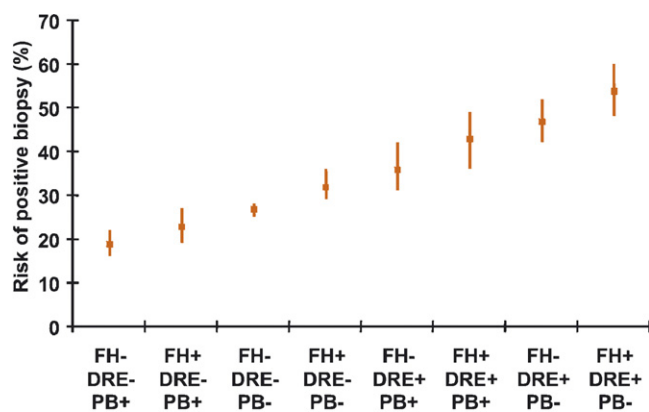
In addition to this, the method of construction of the nomograms and the fact that these are not compared in a “head-to-head fashion” may have resulted in quantitative differences on one side and in failure to appropriately judge their quality on the other. We cannot judge which of the nomograms is best.

Whilst the studies differed in the availability of parameters on which to base their models, they also differed with regard to the parameters providing an independent, statistically significant effect on risk. Although it is likely that individual population differences played a part with regard to variables such as ethnicity, and possibly family history, it is difficult to conceive that parameters such as age and PSA velocity should vary dramatically between the patient populations examined. It is more likely that the nature of the patient population (variable distribution) and its size impacted the likelihood of the parameter reaching statistical significance. An example of this phenomenon can be seen in the PCPT with regard to the effect of African American ethnicity on positive biopsy risk. The odds ratio (OR) for all tumour risk for African American ethnicity was 1.42, but this narrowly missed statistical significance (95% CI, 1.0–2.01) [32]. Thus in the PCPT model, the risk of any grade of tumour is unaltered by African American ethnicity, although more patients within the study are likely to find this parameter to be statistically significant. This highlights the danger of using significance values to

determine inclusion of parameters within a model. A further issue is the value of different variables according to the previous history of screening and/or prostate biopsy. For example, data from the Rotterdam Section of the European Randomized study of Screening for Prostate Cancer (ERSPC) demonstrate that the value of PSA as a predictor of a positive biopsy is reduced in men with a previous negative biopsy [26].

Even when broadly similar variables are included in the models, variations in the model construction can be driven by the statistical analysis. Although adding risk factors to the model may enhance the AUC, such an increase in complexity may afford only marginal benefits, if any at all, over PSA value alone. This is because a very strong association between a variable and a positive biopsy is required to have a significant impact on the AUC [12,48]. ORs that may appear to have import for prediction, and have statistical significance, can have little or no impact on the predictive accuracy of a model. Furthermore, the existing elements of a model can impact the significance of new variables if, for example, they used cutoff levels [49]. Overall, it is difficult to judge the extent to which an individual parameter may impact the AUC. Even if the OR is large, many different combinations of true positive and false positive fractions are possible for a given OR, and therefore it does not describe the discriminant value of the variable [12].

A more complex issue is whether variables are truly independent. The well-understood relationships between age, prostate volume, and PSA levels mean that PSA may effectively subsume other risk factors: it is likely that the elimination of age from a number of models reflects this association. The relationship between PSA value, prostate volume, and cancer detection is particularly complex. Although PSA level has a correlation with prostate volume, positive prostate biopsies are more common in men with smaller prostate volumes, most likely through a combination of an increased chance of a biopsy core hitting tumour tissue and a real association of smaller volume with elevated tumour risk [50,51]. Recent evidence also demonstrates that PSA velocity and tPSA are closely linked [52] and that PSA doubling time adds little to prostate cancer detection [53]; observations that may account for the lack of independent significance of PSA velocity observed in the PCPT model. Ethnicity may also be associated with a higher serum PSA [29]; evidence for a relationship with family history is, however, lacking. Although PSA level is a single variable, its value may therefore reflect a number of risk factors.



**Fig. 2 – Risk of a positive biopsy and 95% confidence intervals at a PSA level of 2.5 ng/ml in the Prostate Cancer Prevention Study (PCPT) population (for-cause and end-of-study biopsies) modified by other risk factors [31].**

**FH = family history (+, yes; –, no); DRE = digital rectal examination (+, suspicious; –, unsuspecting); PB = previous biopsy (+, yes; –, no).**

One critical observation about the precision of prostate cancer risk estimates based on PSA levels has been provided by the PCPT. In Fig. 2, the risk predictions and 95% confidence intervals are presented for men with a PSA level of 2.5 ng/ml modified by other risk factors, namely family history of prostate cancer (yes/no), DRE status (suspicious/not suspicious), and previous negative biopsy (yes/no) [31]. Not only does the risk level alter with different risk factors, but the degree of precision of the estimate also varies because the proportion of the PCPT population with those risk factors also differs. The key observation is that, for the largest group of men, who have no family history, an unsuspecting DRE, and no previous negative biopsy, the confidence intervals are narrow: the risk estimate at a PSA level of 2.5 ng/ml is 27% (95% CI, 25%–28%). These data suggest that, with all the limitations mentioned, for the majority of men, PSA value alone provides a level of precision that may be acceptable for prostate cancer risk assessment. In men with additional risk factors, the use of a more complex assessment will yield a more accurate estimation.

Another key question regarding nomograms or ANNs is their applicability to populations outside of those on which they were modelled [45]. The majority of models did not undergo external validation, and their value outside of the model populations is therefore unknown. Of those that were validated externally, most had a reduction in performance, although this was typically minor. The ANN of Stephan et al has been extensively validated in external populations, with interesting

results. Minor reductions in performance were noted in some comparisons, but some reductions are in excess of 10% [35–37]. Interestingly, a reduction of 10% was observed when the ANN was tested in a single centre, but the performance was almost identical when validated in a multi-centre population [37]. This suggests that significant, population-driven differences may occur between individual centres, which are less likely to be observed in a wide population. In their studies, they also established that ANNs based on the local population and the specific PSA assay used generally performed better than an ANN based on a multi-centre population [35–37]. The recently published, 10-core nomogram of Chun et al was externally validated in diverse populations and found to have a very similar AUC [39]: this may reflect the higher positive biopsy rate observed with an increased core number. This study also afforded a valuable insight into the impact of altering biopsy core number on the AUC of a nomogram. The study demonstrated that a nomogram derived from a population who underwent six-core biopsies was less predictive in men who underwent 10-core biopsies, supporting earlier evidence that increasing the biopsy core number can negatively impact the AUC of PSA level [54]. Only one model included core number as a predictive factor for cancer detection, based on a modest but significant hazard ratio of 1.19 [29].

## 5. Conclusion

In conclusion, paper-based nomograms, or their electronic equivalents, have produced improvements in AUC over PSA values alone in predicting positive biopsies in a number of studies, but many lack external validation. The most recent of these, the prediction model based on data from the PCPT, is associated with a minimal increment in performance over PSA value alone; an observation confirmed in external validation. ANNs have demonstrated significant improvements in AUC over PSA value alone, although the magnitude of these benefits appears diminished when they are validated externally. Nevertheless, performance improvements over PSA levels appear to be largely maintained in external populations, although further external validation would be useful to confirm this. Available evidence suggests that models may not apply well to populations outside of those from which they were derived: certainly locally derived ANNs have better performance than those derived at other centres. This is exemplified by widely varying risk levels generated for similar

patient types by different models, part of which is attributable to the populations on which they are constructed. However, it is clear that a number of nomograms and ANNs have significant advantages over PSA levels alone with regard to the AUC when they are externally validated.

Nomograms and ANNs appear to have a valuable role to play in assessing patient risk [55], although internal validation against PSA value alone and external validation in populations beyond the model cohort are critical [56]. Despite their potential, their clinical utility in everyday practice may have limitations. It is unclear whether many practitioners will have access to all the parameters (transrectal ultrasound [TRUS]-derived volume, %fPSA, etc) that many of the nomograms and ANNs require to achieve an enhanced precision over PSA level alone, although some nomograms can still be utilised without all the parameters being available, albeit at a loss of accuracy compared with the full model. Where these are not incorporated into a nomogram, such as the one derived from the PCPT data, the incremental precision over PSA alone is slight. Another key question is what role estimating risk to the single-percent level has in clinical practice. As the decision to biopsy a patient is typically based on a threshold level of risk, knowing the exact level of risk, beyond that it is above threshold, may add little to the decision-making process, although a more accurate assessment of risk may result in individual men changing position around threshold values. PSA value alone may therefore offer an approach to risk estimation for men without additional risk factors, but overall nomograms afford a more accurate assessment of risk.

**Author contributions:** Fritz Schröder had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

*Study concept and design:* Schröder, Kattan.

*Acquisition of data:* Schröder, Kattan.

*Analysis and interpretation of data:* Schröder, Kattan.

*Drafting of the manuscript:* Schröder, Kattan.

*Critical revision of the manuscript for important intellectual content:* Schröder, Kattan.

*Statistical analysis:* Schröder, Kattan.

*Obtaining funding:* Schröder, Kattan.

*Administrative, technical, or material support:* Schröder, Kattan.

*Supervision:* Schröder, Kattan.

*Other (specify):* Schröder, Kattan.

**Financial disclosures:** I certify that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/ affiliation,

grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: None.

**Funding/Support and role of the sponsor:** GlaxoSmithKline provided funding, financial support and/or material support which included design and conduct of this study, collection of the data, management of the data, analysis, interpretation of the data, preparation, review, and approval of the manuscript.

**Acknowledgment statement:** The authors acknowledge Alexander Gray for providing medical writing support.

## References

- [1] Welch HG, Schwartz LM, Woloshin S. Prostate-specific antigen levels in the United States: implications of various definitions for abnormal. *J Natl Cancer Inst* 2005;97:1132–7.
- [2] Thompson IM, Ankerst DP, Chi C, et al. Operating characteristics of prostate-specific antigen in men with an initial PSA level of 3.0 ng/ml or lower. *JAMA* 2005;294:66–70.
- [3] Chun FK, Karakiewicz PI, Briganti A, et al. A critical appraisal of logistic regression-based nomograms, artificial neural networks, classification and regression-tree models, look-up tables and risk-group stratification models for prostate cancer. *BJU Int* 2007;99:794–800.
- [4] Chun FK, Karakiewicz PI, Huland H, Graefen M. Role of nomograms for prostate cancer in 2007. *World J Urol* 2007;25:131–42.
- [5] Chun FK-H, Karakiewicz PI, Briganti A, et al. Prostate cancer nomograms: an update. *Eur Urol* 2006;50:914–26.
- [6] Kattan MW. Comparison of Cox regression with other methods for determining prediction models and nomograms. *J Urol* 2003;170(Suppl):6–10.
- [7] Eastham JA, Kattan MW, Scardino PT. Nomograms as predictive models. *Semin Urol Oncol* 2002;20:108–15.
- [8] Kattan MW. Nomograms are superior to staging and risk grouping systems for identifying high-risk patients: pre-operative application in prostate cancer. *Curr Opin Urol* 2003;13:111–6.
- [9] Ross PL, Scardino PT, Kattan MW. A catalog of prostate cancer nomograms. *J Urol* 2001;165:1562–8.
- [10] Ohori M, Swindle P. Nomograms and instruments for the initial prostate evaluation: the ability to estimate the likelihood of identifying prostate cancer. *Semin Urol Oncol* 2002;20:116–22.
- [11] Hammerer PG, Kattan MW, Mottet N, Prayer-Galetti T. Using prostate-specific antigen screening and nomograms to assess risk and predict outcomes in the management of prostate cancer. *BJU Int* 2006;98:11–9.
- [12] Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol* 2004;159:882–90.
- [13] Snow PB, Smith DS, Catalona WJ. Artificial neural networks in the diagnosis and prognosis of prostate cancer: a pilot study. *J Urol* 1994;152:1923–6.
- [14] Optenberg SA, Clark JY, Brawer MK, et al. Development of a decision-making tool to predict risk of prostate cancer:

- the Cancer of the Prostate Risk Index (CAPRI) test. *Urology* 1997;50:665–72.
- [15] Carlson GD, Calvanese CB, Partin AW. An algorithm combining age, total prostate-specific antigen (PSA), and percent free PSA to predict prostate cancer: results on 4298 cases. *Urology* 1998;52:455–61.
- [16] Babaian RJ, Fritsche H, Ayala A, et al. Performance of a neural network in detecting prostate cancer in the prostate-specific antigen reflex range of 2.5 to 4.0 ng/ml. *Urology* 2000;56:1000–6.
- [17] Eastham JA, May R, Robertson JL, Sartor O, Kattan MW. Development of a nomogram that predicts the probability of a positive prostate biopsy in men with an abnormal digital rectal examination and a prostate-specific antigen between 0 and 4 ng/ml. *Urology* 1999;54:709–13.
- [18] Finne P, Finne R, Auvinen A, et al. Predicting the outcome of prostate biopsy in screen-positive men by a multilayer perceptron network. *Urology* 2000;56:418–22.
- [19] Djavan B, Remzi M, Zlotta A, et al. Novel artificial neural network for early detection of prostate cancer. *J Clin Oncol* 2002;20:921–9.
- [20] Finne P, Auvinen A, Aro J, et al. Estimation of prostate cancer risk on the basis of total and free prostate-specific antigen, prostate volume and digital rectal examination. *Eur Urol* 2002;41:619–27.
- [21] Stephan C, Cammann H, Semjonow A, et al. Multicenter evaluation of an artificial neural network to increase the prostate cancer detection rate and reduce unnecessary biopsies. *Clin Chem* 2002;48:1279–87.
- [22] Garzotto M, Hudson RG, Peters L, et al. Predictive modeling for the presence of prostate carcinoma using clinical, laboratory, and ultrasound parameters in patients with prostate specific antigen levels  $\leq 10$  ng/ml. *Cancer* 2003;98:1417–22.
- [23] Lopez-Corona E, Ohori M, Scardino PT, et al. A nomogram for predicting a positive repeat prostate biopsy in patients with a previous negative biopsy session. *J Urol* 2003;170:1184–8.
- [24] Yanke BV, Gonen M, Scardino PT, Kattan MW. Validation of a nomogram for predicting positive repeat biopsy for prostate cancer. *J Urol* 2005;173:421–4.
- [25] Karakiewicz PI, Benayoun S, Kattan MW, et al. Development and validation of a nomogram predicting the outcome of prostate biopsy based on patient age, digital rectal examination and serum prostate specific antigen. *J Urol* 2005;173:1930–4.
- [26] Roobol MJ, Schroder FH, Kranse R. A comparison of first and repeat (four years later) prostate cancer screening in a randomized cohort of a symptomatic men aged 55–75 years using a biopsy indication of 3.0 ng/ml (results of ERSPC, Rotterdam). *Prostate* 2006;66:604–12.
- [27] Nam RK, Toi A, Klotz LH, et al. Nomogram prediction for prostate cancer and aggressive prostate cancer at time of biopsy: utilizing all risk factors and tumor markers for prostate cancer. *Can J Urol* 2006;13(Suppl 2):2–10.
- [28] Suzuki H, Komiya A, Kamiya N, et al. Development of a nomogram to predict probability of positive initial prostate biopsy among Japanese patients. *Urology* 2006;67:131–6.
- [29] Yanke BV, Carver BS, Bianco Jr FJ, et al. African-American race is a predictor of prostate cancer detection: incorporation into a pre-biopsy nomogram. *BJU Int* 2006;98:783–7.
- [30] Klein EA, Tangen CM, Goodman PJ, Lippman SM, Thompson IM. Assessing benefit and risk in the prevention of prostate cancer: the Prostate Cancer Prevention Trial revisited. *J Clin Oncol* 2005;23:7460–6.
- [31] National Cancer Institute Early Detection Research Network Web site. Cancer Risk Calculator for Prostate Cancer. <http://www.compass.fhcr.org/edrnci/bin/calculator/main.asp> April 2008.
- [32] Thompson IM, Ankerst DP, Chi C, et al. Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. *J Natl Cancer Inst* 2006;98:529–34.
- [33] Parekh DJ, Ankerst DP, Higgins BA, et al. External validation of the Prostate Cancer Prevention Trial risk calculator in a screened population. *Urology* 2006;68:1152–5.
- [34] Stephan C, Meyer H-A, Kwiatkowski M, et al. A (–5, –7) proPSA-based artificial neural network to detect prostate cancer. *Eur Urol* 2006;50:1014–20.
- [35] Stephan C, Xu C, Cammann H, et al. Assay-specific artificial neural networks for five different PSA assays and populations with PSA 2–10 ng/ml in 4480 men. *World J Urol* 2007;25:95–103.
- [36] Chun FK-H, Graefen M, Briganti A, et al. Initial biopsy outcome prediction—head-to-head comparison of a logistic regression-based nomogram versus artificial neural network. *Eur Urol* 2007;51:1236–43.
- [37] Stephan C, Xu C, Finne P, et al. Comparison of two different artificial neural networks for prostate biopsy indication in two different patient populations. *Urology* 2007;70:596–601.
- [38] Nam RK, Toi A, Klotz LH, et al. Assessing individual risk for prostate cancer. *J Clin Oncol* 2007;25:3582–8.
- [39] Chun FK-H, Briganti A, Graefen M, et al. Development and external validation of an extended 10-core biopsy nomogram. *Eur Urol* 2007;52:436–45.
- [40] Thompson IM, Pauler DK, Goodman PJ, et al. Prevalence of prostate cancer among men with a prostate-specific antigen level  $\leq 4.0$  ng/ml. *N Engl J Med* 2004;350:2239–46.
- [41] Raaijmakers R, Wildhagen MF, Ito K, et al. Prostate-specific antigen change in the European Randomized Study of Screening for Prostate Cancer, section Rotterdam. *J Urol* 2004;63:316–20.
- [42] Crawford ED, Abrahamsson P-A. PSA-based screening for prostate cancer: how does it compare with other cancer screening tests? *Eur Urol* 2008;54:262–73.
- [43] Nam RK, Toi A, Trachtenberg J, et al. Making sense of prostate specific antigen: improving its predictive value in patients undergoing prostate biopsy. *J Urol* 2006;175:489–94.
- [44] Van der Crujisen-Koeter IW, Vis AN, Roobol MJ, et al. Comparison of screen detected and clinically diagnosed prostate cancer in the European randomized study of screening for prostate cancer, section Rotterdam. *J Urol* 2005;174:121–5.

- [45] Guillonneau B. *Ceteris paribus* and nomograms in medicine. *Eur Urol* 2007;52:1287-9.
- [46] 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.
- [47] Vickers AJ, Ulmert D, Serio AM, 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.
- [48] Wald NJ, Hackshaw AK, Frost CD. When can a risk factor be used as a worthwhile screening test? *BMJ* 1999;319: 1562-5.
- [49] Kattan MW. Judging new markers by their ability to improve predictive accuracy. *J Natl Cancer Inst* 2003;95: 634-5.
- [50] Al-Azab R, Toi A, Lockwood G, Kulkarni GS, Fleshner N. Prostate volume is strongest predictor of cancer diagnosis at transrectal ultrasound-guided prostate biopsy with prostate-specific antigen values between 2.0 and 9.0 ng/ml. *Urology* 2007;69:103-7.
- [51] Remzi M, Fong YK, Dobrovits M, et al. The Vienna nomogram: validation of a novel biopsy strategy defining the optimal number of cores based on patient age and total prostate volume. *J Urol* 2005;174:1256-61.
- [52] Yu X, Loeb S, Roehl KA, Han M, Catalona WJ. The association between total prostate specific antigen concentration and prostate specific antigen velocity. *J Urol* 2007;177: 1298-302.
- [53] Spurgeon SE, Mongoue-Tchokote S, Collins L, et al. Assessment of prostate-specific antigen doubling time in prediction of prostate cancer on needle biopsy. *Urology* 2007;69:931-5.
- [54] Canto EI, Singh H, Shariat SF, et al. Effects of systematic 12-core biopsy on the performance of percent free prostate specific antigen for prostate cancer detection. *J Urol* 2004;172:900-4.
- [55] Bianco Jr FJ. Nomograms and medicine. *Eur Urol* 2006;50:884-6.
- [56] Ross RW, Kantoff PW. Predicting outcomes in prostate cancer: how many more nomograms do we need? *J Clin Oncol* 2007;25:3563-4.
- [57] Babaian RJ, Fritsche HA, Zhang Z, et al. Evaluation of prostAsure index in the detection of prostate cancer: a preliminary report. *Urology* 1998;51:132-6.