

Aus der Klinik für Urologie
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Evaluierung des Einflusses der PSA-Test-Variabilität auf die Vorhersage eines Prostatakarzinoms mittels Nomogrammen

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von
Kerstin Siemßen
aus Berlin

Gutachter: 1. Herr PD Dr. med. Carsten Stephan
2. Herr Prof. Dr. med. Dietmar Schnorr
3. Herr PD Dr. med. Edwin Herrmann

Datum der Promotion: 24.02.2012

Teile dieser Dissertation sind in folgender Publikation erschienen:

Stephan C[§], Siemßen K[§], Cammann H, Deger S, Schrader M, Miller K, Lein M, Jung K, Meyer HA. Faulty risk prediction of prostate cancer by nomograms without considering PSA inter-assay variability. JCO 2009; re-submitted, November 2009.

[§]both authors contributed equally to the manuscript

Die Arbeit wurde im Rahmen der 20. Europäischen Studentenkonzferenz in Berlin, 4. bis 7. Oktober 2009, unter dem Titel „Use and misuse of nomograms for prostate cancer detection without consideration of analytical differences between PSA assays“ präsentiert.

Inhaltsverzeichnis

Inhaltsverzeichnis	I
Tabellenverzeichnis	III
Abbildungsverzeichnis	IV
Abkürzungsverzeichnis	V
1. Einleitung	1
1.1 Das Prostatakarzinom	1
1.2 Die PCa-Diagnostik	3
1.2.1 Prostataspezifisches Antigen.....	3
1.2.2 Molekulare Formen des PSA.....	5
1.2.3 Digitale rektale Untersuchung.....	7
1.2.4 Transrektaler Ultraschall und Prostatavolumen	8
1.2.5 Lebensalter	8
1.2.6 Verbesserungen der PSA-Diagnostik	9
1.3 Multivariate Auswertemodelle	10
1.3.1 Artificielle neuronale Netzwerke	10
1.3.2 Nomogramme	11
2. Zielsetzung der Arbeit	13
3. Material und Methoden.....	14
3.1 Patientenkollektiv.....	14
3.2 Probenmaterial	15
3.3 PSA-Bestimmungen	15
3.4 Klinische Untersuchungen	16
3.5 Nomogramme	17
3.6 Statistische Auswertungen	18
4. Ergebnisse	23
4.1 Charakterisierung des Patientenkollektivs	23
4.2 Analytik und Kalibrierung der Nomogramm-Ausgänge.....	25
4.3 Diskrimination	32
4.3.1 Analyse bei verschiedenen Cut-Off-Werten.....	32

4.3.2 ROC-Analyse.....	37
4.3.3 Analyse bei 90% und 95% Sensitivität.....	42
5. Diskussion	44
6. Zusammenfassung	54
7. Literaturverzeichnis.....	56
Eidesstattliche Erklärung	i
Lebenslauf	ii
Danksagung	iii

Tabellenverzeichnis

Tabelle 1. Übersicht über die PSA-Testsysteme mit den verwendeten Abkürzungen und den vollständigen Bezeichnungen sowie Herstellern.....	16
Tabelle 2. Übersicht über die angewandten Nomogramme mit ihren Eingangsvariablen und den zugrunde liegenden PSA-Tests.....	18
Tabelle 3. Charakteristika des Patientenkollektivs: Mediane und Bereiche für Alter, Prostatavolumen, positive DRU, tPSA und %fPSA für alle, PCa- und BPH-Patienten vergleichend für die Kohorte mit 798 und 780 Patienten.....	24
Tabelle 4. Mediane und Bereiche für die vorhergesagten PCa-Wahrscheinlichkeiten.	26
Tabelle 5. Flächen unter den ROC-Kurven (AUCs) mit Standardabweichung (in Klammern) für tPSA, %fPSA und die fünf Nomogramme mit den fünf PSA-Tests.	39
Tabelle 6. Spezifitäten und Sensitivitäten mit Konfidenzintervallen für die fünf Nomogramme mit den fünf PSA-Tests.	41
Tabelle 7. Nomogramm-Wahrscheinlichkeiten sowie positive und negative prädiktive Werte bei 90% und 95% Sensitivität.....	43

Abbildungsverzeichnis

Abbildung 1. Molekulare Formen des PSA	7
Abbildungen 2a-d. Kalibrierung des Nomogramms IV am Beispiel der Abbott-Daten für 5, 10, 20 und 50 Gruppen, die jeweils aus 156, 78, 39 und 15 Patienten bestehen.....	20-22
Abbildungen 3a-e. Darstellung der vorhergesagten PCa-Wahrscheinlichkeiten in Bezug auf die beobachteten PCa-Raten bei der Biopsie mittels Kalibrationskurven für die Nomogramme I bis V.	29-31
Abbildungen 4a-e. Darstellung der Sensitivitäten und Spezifitäten in Abhängigkeit von verschiedenen Cut-Off-Werten für die Nomogramm-Wahrscheinlichkeiten	34-36
Abbildung 5. Der Verlauf der ROC-Kurven ist für alle fünf PSA-Tests innerhalb des Nomogramms II dargestellt..	40

Abkürzungsverzeichnis

ACT-PSA	PSA gebunden an α_1 -Antichymotrypsin
ANN	artifizielles neuronales Netzwerk
API-PSA	PSA gebunden an α_1 -Proteaseinhibitor
AUC	Area under the curve (engl.) = Fläche unter der Kurve
A2M-PSA	PSA gebunden an α_2 -Makroglobulin
BPH	benigne Prostatahyperplasie
bPSA	benignes PSA (Subform des freien PSA)
bzw.	beziehungsweise
cPSA	komplexiertes PSA
DRU	digitale rektale Untersuchung
ERSPC	European Randomized Study of Screening for Prostate Cancer (engl.) = europäische Screeningstudie
fPSA	freies PSA
fPSA-i	enzymatisch inaktives freies PSA
%fPSA	prozentuales freies PSA oder fPSA/tPSA (Ratio)
ICC	Intraklassenkorrelationskoeffizient
LR	logistische Regression
p-Wert	Signifikanzwert
PCa	Prostatakarzinom
proPSA	aus 244 Aminosäuren bestehende Vorstufe des enzymatisch aktiven PSA
PSA	prostataspezifisches Antigen
ROC-Kurve	Receiver-operating-characteristic-Kurve
SD	Standardabweichung
tPSA	totales (gesamtes) PSA
TRUS	transrektaler Ultraschall

1. Einleitung

1.1 Das Prostatakarzinom

Mit einem Anteil von 25,4% an allen Krebserkrankungen und einer Inzidenz von 58.000 pro Jahr stellt das Prostatakarzinom (PCa) die häufigste bösartige Neubildung unter Männern in Deutschland dar [1]. In allen anderen westlichen Industriestaaten ist das PCa ebenfalls der häufigste maligne Tumor beim Mann. Die Neoplasie steht mit 10,1% an dritter Stelle bei den zum Tode führenden Krebserkrankungen [2].

Das mittlere Erkrankungsalter liegt bei etwas über 70 Jahren. Während das PCa vor dem 50. Lebensjahr äußerst selten ist, steigt die Prävalenz mit zunehmendem Lebensalter deutlich an [2]. Insbesondere aufgrund der verbesserten Methoden der Früherkennung ist die Inzidenz des PCa in den letzten Jahrzehnten stark angestiegen [1].

Die Ursachen für die Entstehung eines PCa sind bisher noch weitgehend unbekannt; als Risikofaktoren werden Übergewicht, Bewegungsmangel und Rauchen diskutiert [1, 3]. Bei jüngeren Patienten wird eine genetische Prädisposition vermutet; diese Hypothese wird durch die unterschiedliche Prävalenz und Mortalität in verschiedenen ethnischen Gruppen unterstützt [4]. Das Vorhandensein eines 2005 erstmals beschriebenen Genfusionstranskripts [5] ist mit erhöhter Tumoraggressivität und Mortalität assoziiert [6]. Etwa die Hälfte aller im Screening entdeckten PCa weisen diese Genfusion auf [7].

Aufgrund der bevorzugten Lokalisation des PCa in der peripheren Zone der Prostata tritt der Krebs klinisch in der Regel erst in späten Stadien in Erscheinung. Selten kann ein PCa zu Nykturie, Pollakisurie oder Harnverhalt führen; diese Symptome sind jedoch wesentlich häufiger auf eine benigne Prostatahyperplasie (BPH) zurückzuführen. Eine neu aufgetretene erektile Dysfunktion, Hämatospermie und Hämaturie sind ebenfalls mögliche Symptome, bei denen ein PCa in Betracht gezogen werden sollte [8]. Ein kleiner Teil der Patienten präsentiert bereits zum

Diagnosezeitpunkt durch Metastasen bedingte Symptome, wie zum Beispiel Knochenschmerzen. Das PCa metastasiert auf lymphogenem Wege vorrangig in die obturatorischen und iliakalen Lymphknoten, auf hämatogenem Wege vor allem in das Skelett.

Die Prognose eines Patienten mit PCa hängt stark von der Differenzierung und Ausbreitung des Tumors sowie möglichen Metastasen bei Diagnosestellung ab. Bei einem lokal begrenzten, nicht metastasierten Tumor kann mittels radikaler Prostatovesikulektomie oder Strahlentherapie eine kurative Therapie eingeleitet werden. Die Überlebenswahrscheinlichkeit nach 5 Jahren beträgt bei dieser Patientengruppe 100% [2]. Wird das Karzinom erst nach Kapselüberschreitung oder sogar Metastasierung diagnostiziert, sinken die 5-Jahres-Überlebensraten zwar deutlich ab [2], dennoch stehen verschiedene kurative und palliative Therapieoptionen zur Auswahl.

Die geringen Frühsymptome und die sinkenden Überlebensraten bei zu später Diagnosestellung verdeutlichen die große Relevanz einer zuverlässigen Früherkennung. 1971 wurde ein gesetzliches Früherkennungsprogramm für Männer ab dem 45. Lebensjahr eingeführt, das eine Anamnese und eine digitale rektale Untersuchung (DRU) der Prostata beinhaltet. Doch erst die Einführung des prostata-spezifischen Antigens (PSA) in die Früherkennung des PCa vor 15 Jahren führte zu einem deutlichen Anstieg an diagnostizierten Neuerkrankungen und zu einer vermehrten Diagnose lokal begrenzter Tumoren [9].

Heute können etwa 91% der Prostatatumoren im lokalisierten Stadium diagnostiziert werden [2]. Man nimmt jedoch an, dass die Anwendung des PSA-Tests in bis zu 50% eine Überbehandlung von Tumoren nach sich zieht, die möglicherweise weniger aggressiv oder gar nicht behandelt werden müssten [10]. Dahingegen zeigte die aktuelle Auswertung der europäischen Screeningstudie (ERSPC), dass die PCa-bedingte Sterblichkeit bei Patienten mit regelmäßigen PSA-Messungen 31% geringer ist [11]. Schröder et al. wiesen in der initialen Analyse der ERSPC bereits eine Reduktion der PCa-Mortalität um 20% nach [12].

Obwohl eine weitere Screeningstudie keine geringere Mortalität durch PSA-Screening offen legte [13], ist der Nutzen des PSA unbestritten.

Das PSA wird daher Thema des folgenden Abschnitts sein.

1.2 Die PCa-Diagnostik

1.2.1 Prostataspezifisches Antigen

Das PSA ist der entscheidende Serummarker für die frühzeitige Detektion eines PCa und auch der wesentliche Verlaufsparemeter nach radikaler Prostatektomie [14, 15]. Nachdem es in den 70er Jahren entdeckt worden war und zuerst unter anderem Namen bekannt war [16-19], wurde es 1979 von Wang et al. als „prostataspezifisch“ bezeichnet [20]. Anfang der 80er Jahre gelang der Nachweis des Antigens im Serum von PCa-Patienten [21]; des Weiteren konnte der Zusammenhang mit dem Fortschreiten der Erkrankung gesichert werden [22]. 1994 wurde das PSA in den USA durch die „Food and Drug Administration“ in Kombination mit der DRU zur Früherkennung eines PCa zugelassen.

Beim PSA handelt es sich um ein Glykoprotein mit einem Molekulargewicht von etwa 30 Kilodalton, welches aus 237 Aminosäuren und einer Kohlenhydratkette besteht [23]. Bedingt durch seine Lokalisation auf dem Chromosom 19, die Funktion als Serinprotease, die Regulierung durch Androgene und die ähnliche DNA- und Aminosäuresequenz wird das PSA den humanen Kallikreinen zugeordnet [24].

Innerhalb der Prostata konnte das PSA im endoplasmatischen Retikulum, in den Sekretgranula und den luminalen Zelloberflächen nachgewiesen werden [19] - es ist damit weitgehend organ-, nicht jedoch tumorspezifisch. In sehr geringem Maße findet sich das PSA auch in anderen Geweben, so zum Beispiel in den Paraurethraldrüsen, der Brustdrüse, der Schilddrüse, den Speicheldrüsen, der Lunge, den Hoden und dem Uterus [25].

Bei der Synthese des PSA wird zunächst ein Präprotein mit 261 Aminosäuren gebildet, das nach erfolgter Prozessierung als proPSA, bestehend aus 244 Aminosäuren, in das Drüsenlumen sezerniert wird [23]. Das enzymatisch aktive PSA entsteht nach proteolytischer Abspaltung von weiteren sieben Aminosäuren [26]. Das als Serinprotease fungierende Antigen spaltet Peptidbindungen an der Carboxylgruppe von Tyrosin und Leucin [27].

Nach erfolgter Synthese durch die Prostataepithelzellen wird das PSA durch Exozytose in das Drüsenlumen und anschließend in das Seminalplasma

abgegeben [28]. Dort ist es durch Proteolyse der gelbildenden Proteine Fibronectin und Semenogelin I und II für die Verflüssigung des Samens nach der Ejakulation verantwortlich [29].

Das PSA gelangt nur zu einem sehr geringen Anteil in den systemischen Kreislauf. Während die Konzentration im Seminalplasma bei 0,5-3 g/l liegt, beträgt die Serumkonzentration des PSA bei gesunden Männern in der Regel weniger als 4 µg/l [22, 28]. Somit ist die Konzentration im Serum etwa eine Million Mal geringer als die Konzentration in der Samenflüssigkeit. In der urologischen Praxis markiert der Wert von 4 µg/l die Referenzgrenze, ab der weitere Diagnostik in Hinblick auf ein PCa notwendig wird [30].

Ein maligner Prostatatumor führt zu einer Störung der „Prostata-Blut-Schranke“ und einem aktiven Sezernieren des PSA in das Interstitium und bedingt so die vielfach höheren PSA-Konzentrationen im Serum der PCa-Patienten verglichen mit den Serumkonzentrationen gesunder Männer [31]. Allerdings werden auch bei gutartigen Prostataerkrankungen wie bei der BPH und der Prostatitis sowie nach diagnostischen und therapeutischen Maßnahmen an der Prostata (zum Beispiel nach der DRU) erhöhte PSA-Konzentrationen detektiert [32].

Obwohl die PSA-Werte im Serum pro Gewichtseinheit an Prostatagewebe beim PCa zehnfach höher liegen, ist die Interpretation erhöhter Werte erschwert [15, 33, 34].

Letztlich kann die endgültige Diagnosestellung nur durch die Entnahme einer Prostatagewebeprobe erfolgen [35, 36].

Generell wird eine Prostatabiopsie, nach nochmaliger PSA-Kontrolle, bei einem PSA-Wert > 4 µg/l empfohlen - unabhängig vom DRU-Befund. Der positive prädiktive Wert für ein PCa beträgt bei einem Serum-PSA-Wert zwischen 4 und 10 µg/l circa 20% [37]. Bei ausschließlicher Beurteilung des PSA-Wertes wird demnach in vier von fünf Fällen die Prostatabiopsie unnötig durchgeführt.

Als problematisch ist ebenso zu bewerten, dass 20% der PCa-Patienten zum Diagnosezeitpunkt einen PSA-Wert < 4 µg/l haben und somit durch das PSA-Screening als falsch-negativ bewertet werden [28, 38]. Die Prävalenz des PCa ist im PSA-Bereich 2-3 und 3-4 µg/l mit 24 bzw. 27% ähnlich hoch wie im Bereich 4-10 µg/l [39]. Im PSA-Bereich bis 4 µg/l [40] und von 4-10 µg/l wurde zumeist ein PCa im nicht-organüberschreitenden Stadium diagnostiziert, während bei mehr als 50% der

Patienten mit PSA-Werten $> 10 \mu\text{g/l}$ ein bereits fortgeschrittenes Karzinom festgestellt wurde [37].

Aufgrund der aufgeführten Limitationen des PSA-Wertes in der Differenzierung zwischen benignen und malignen Erkrankungen der Prostata finden zusätzlich zur PSA-Bestimmung verschiedene weitere Methoden und Marker in der Diagnostik des PCa Anwendung.

Bekannt sind der Einfluss des Prostatavolumens und des Lebensalters auf den PSA-Wert. Die Bestimmungen der molekularen Formen (Isoformen) des PSA im Serum stellen einen weiteren viel versprechenden Ansatz in der PCa-Diagnostik dar. Die PSA-Anstiegsgeschwindigkeit ist aktuell ebenfalls Gegenstand der Forschung.

Alle diese diagnostischen Verfahren inklusive DRU und transrektalem Ultraschall (TRUS) sollen in den folgenden Abschnitten kurz erörtert werden. Ebenso soll der Einsatz von multivariaten Auswertemodellen in der heutigen Diagnostik des PCa dargestellt werden.

1.2.2 Molekulare Formen des PSA

Das PSA zirkuliert im Serum sowohl als ein an Proteaseinhibitoren gebundenes Molekül als auch in freier, ungebundener Form (fPSA), welche mit einem Anteil von 10-30% am Gesamt-PSA (tPSA) die deutlich kleinere Fraktion darstellt [41-43].

Während das Serumprotein α_1 -Antichymotrypsin etwa 70-90% des PSA bindet (ACT-PSA), geht der α_1 -Proteaseinhibitor (entspricht α_1 -Antitrypsin) nur mit sehr geringen Mengen des PSA eine Bindung ein (API-PSA) [41]. Die Summe aus ACT- und API-PSA bezeichnet man als komplexiertes PSA (cPSA) [44].

Ein größerer Anteil des PSA wird auch vom α_2 -Makroglobulin gebunden: A2M-PSA [45-48]. Weit geringere PSA-Mengen binden an Antithrombin III, Inter- α -Trypsin-Inhibitor und an Protein C, wobei der PSA-Protein C-Komplex bisher nur im Seminalplasma nachgewiesen werden konnte [49].

Vermutlich erfolgt im Serum die Bindung an Proteaseinhibitoren, da das PSA in seiner aktiven Form als Serinprotease arbeitet und diese Funktion im Blut unterbunden werden soll [50].

Es konnte nachgewiesen werden, dass der relative Anteil des ACT-PSA bei Karzinompatienten höher liegt als bei BPH-Patienten [42, 43] - trotzdem konnte sich die alleinige Bestimmung des ACT-PSA in der klinischen Routine nie durchsetzen.

1998 wurde eine Methode zur Bestimmung des komplexierten PSA entwickelt: dabei wird sowohl ACT- als auch API-PSA gemessen [44]. Entgegen der anfänglichen Vermutung, der Parameter cPSA könne den Quotienten %fPSA ersetzen, zeigte nur der Quotient cPSA/tPSA vergleichbare Resultate mit dem %fPSA [51, 52].

Das fPSA existiert im Prostatagewebe sowie im Serum in verschiedenen molekularen Formen. Nachgewiesen wurden das proPSA, das benigne PSA (bPSA), das „nicked“ PSA und das inaktive freie PSA (fPSA-i) [53].

Das aus 237 Aminosäuren bestehende fPSA bildet sich durch Abspaltung von sieben Aminosäuren aus dem Vorläufermolekül (-7) proPSA. Es existieren verschiedene Formen des proPSA, die je nach zusätzlicher Anzahl an Aminosäuren, bezogen auf das fPSA, als (-1) bis (-7) proPSA bezeichnet werden [54, 55]. Hierbei sind insbesondere das (-2) proPSA und das (-4) proPSA von Bedeutung für die frühe Detektion des PCa, da sie sich bei den Karzinompatienten vermehrt anreichern [56]. Seit kurzem ist ein automatisierter (-2) proPSA Test kommerziell erhältlich [57, 58]. Im Jahr 2000 gelang erstmals der Nachweis des bPSA, welches wie das fPSA aus 237 Aminosäuren besteht, jedoch einen Bruch der Kette bei Lysin 182 und Lysin 145 aufweist [59]. Das bPSA (auch als BPH assoziiertes PSA bezeichnet: BPHA) ist eher mit einer gutartigen Vergrößerung der Prostata assoziiert [60].

Beim fPSA-i handelt es sich um eine intakte, jedoch enzymatisch inaktive Subform des fPSA, wobei sich die Konzentrationen im Serum von PCa- und BPH-Patienten nicht unterscheiden [61]. Signifikant erhöht war dagegen der fPSA-i/fPSA-Quotient bei den PCa-Patienten [61, 62]. Das sogenannte „nicked“ PSA besteht aus einer in sich gespaltenen Aminosäurekette.

Eine signifikante Verbesserung in der Diskrimination zwischen PCa und BPH konnte durch die Messung des fPSA und die anschließende Berechnung des prozentualen freien PSA ($\%fPSA = fPSA/tPSA$) erreicht werden [63-68].

Der Anteil des fPSA ist bei den BPH-Patienten relativ höher, bei den PCa-Patienten relativ geringer [41-43, 69, 70]. Mit sinkenden %fPSA-Werten steigt die Wahrscheinlichkeit, an einem PCa erkrankt zu sein, signifikant an [67, 71]. Diese

quantitativen fPSA-Unterschiede bei benignen und malignen Erkrankungen der Prostata liegen wahrscheinlich darin begründet, dass das fPSA bei der BPH aufgrund einer partiellen proteolytischen Spaltung nach seiner zellulären Freisetzung (zumindest teilweise) nicht mehr von ACT im Serum gebunden werden kann [72]. Es werden allerdings auch andere Hypothesen diskutiert [69, 73]. Im tPSA-Bereich von 4-10 µg/l können durch den Einsatz des %fPSA etwa 20-25% aller Prostatabiopsien vermieden werden [74, 75]. Dies gilt mit Einschränkungen auch für den niedrigeren tPSA-Bereich von < 4 µg/l [76-78].

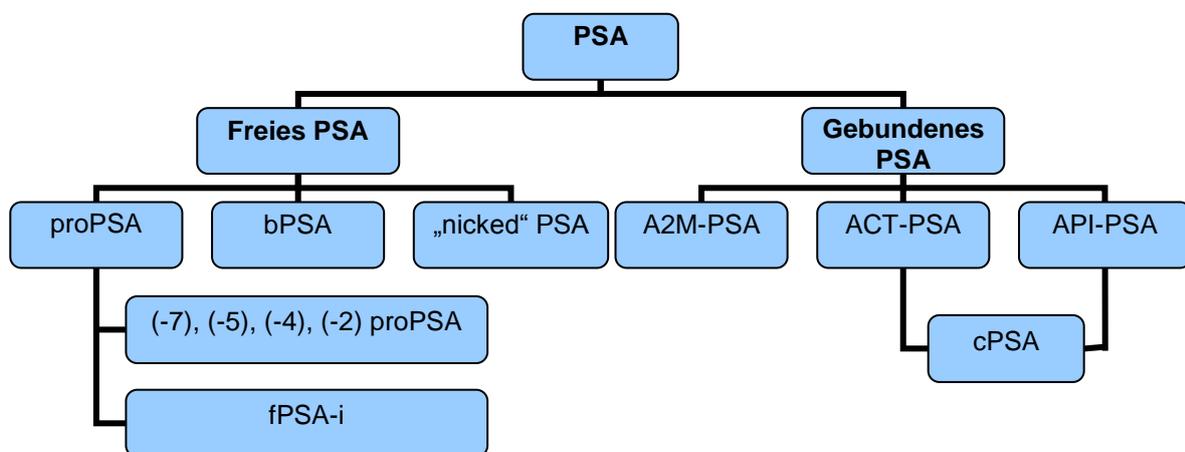


Abbildung 1. Molekulare Formen des PSA (modifiziert nach einer Abbildung von Stephan et al. [70]).

1.2.3 Digitale rektale Untersuchung

Gemeinsam mit der Bestimmung des PSA bildet die DRU die Grundlage der PCa-Diagnostik. Die digital-rektale Palpation der Prostata wird ab einem Alter von 45 Jahren einmal pro Jahr zur Früherkennung des PCa empfohlen [79]. Obwohl inzwischen gezeigt werden konnte, dass sich ein PCa mit Hilfe des Serummarkers PSA wesentlich früher diagnostizieren lässt als durch die alleinige DRU [38, 80, 81], bleibt die DRU ein wichtiger Bestandteil der urologischen Untersuchung, da einige Karzinome nur durch einen suspekten Tastbefund entdeckt werden. Zudem ist ein suspektes DRU-Ergebnis bei PCa-Patienten deutlich häufiger als bei Patienten mit

benignen Erkrankungen der Prostata und muss zwangsläufig mittels Prostatastanzbiopsie abgeklärt werden.

Als karzinomverdächtig werden Indurationen, Knotenbildungen und eine schlechte Abgrenzbarkeit der Prostata bewertet, während eine prall-elastische Struktur der Drüse den Normalbefund darstellt.

Nachteile der DRU sind, dass diese ein subjektives Untersuchungsverfahren darstellt, das schwer zu standardisieren ist und dass etwa 70% der Tumoren, die durch einen suspekten Tastbefund auffällig geworden sind, schon lokal fortgeschritten oder bereits metastasiert sind und damit einer kurativen Therapie kaum noch zugänglich sind [82].

1.2.4 Transrektaler Ultraschall und Prostatavolumen

Die transrektale Ultraschalluntersuchung (TRUS) der Prostata bildet neben der PSA-Bestimmung und der DRU eine weitere wichtige Säule in der PCa-Diagnostik [83].

Der TRUS kann durch die Darstellung suspekter Areale einen Hinweis auf ein PCa geben. Außerdem dient die Ultraschalluntersuchung der Bestimmung des Prostatavolumens und der Orientierung bei der Stanzbiopsie [84].

Während insbesondere bei den BPH-Patienten das tPSA in positiver Korrelation zum Volumen steht, findet sich bei den PCa-Patienten eher eine positive Korrelation zwischen dem %fPSA und dem Prostatavolumen [85].

Es konnte gezeigt werden, dass die diagnostische Effizienz des %fPSA in der Diskrimination zwischen benignen und malignen Erkrankungen nur bis zu einem Prostatavolumen von 40 cm³ bzw. 40 ml besser ist als die des tPSA [71, 86].

1.2.5 Lebensalter

Bereits 1993 konnte gezeigt werden, dass es einen signifikanten Zusammenhang zwischen höherem Lebensalter und steigendem PSA-Wert gibt [87].

Die Hauptursache der höheren PSA-Werte bei älteren Männern ist in der erhöhten Prävalenz der BPH im fortgeschrittenen Lebensalter zu sehen [88]. Die

hyperplastischen Prostataepithelzellen bilden aufgrund der größeren Gewebemasse mehr PSA.

Anhand dieser Erkenntnisse wurden altersspezifische Referenzgrenzen postuliert [87]. Problematisch ist aber, dass diese zu einer geringeren Spezifität bei jüngeren Männern führen, während bei älteren Männern eine sinkende Sensitivität die Folge ist. Ihre Verwendung wird von den Herstellern kommerzieller PSA-Testsysteme daher nicht empfohlen [50].

1.2.6 Verbesserungen der PSA-Diagnostik

Die PSA-Anstiegsgeschwindigkeit oder auch PSA-Velocity stellt eine Methode zur Optimierung der PSA-Diagnostik dar. Carter et al. stellten fest, dass eine Anstiegsgeschwindigkeit von über 0,75 µg/l pro Jahr mit einem signifikant erhöhten Karzinomrisiko assoziiert ist [89]. Verschiedene Folgestudien kamen zu widersprüchlichen Ergebnissen [90-93].

Ein Nachteil dieser Methode ist, dass eine diagnostische Aussagekraft aufgrund von intraindividuell stark unterschiedlichen Serumwerten erst nach zweijähriger, engmaschiger PSA-Kontrolle erreicht werden kann [94, 95]. Die Variationen, die durch die Nutzung verschiedener PSA-Testsysteme entstehen, schränken die Anwendung der PSA-Anstiegsgeschwindigkeit ebenfalls ein [96].

Ein weiterer Ansatz zur Verbesserung der Diagnostik eines PCa ist die PSA-Dichte oder auch PSA-Density. Die PSA-Dichte, die definiert ist als der Quotient aus dem Gesamt-PSA (tPSA) und dem Prostatavolumen, soll durch die Einbeziehung der Prostatagröße die Diskrimination zwischen BPH- und PCa-Patienten erleichtern [33]. Grundlage dieses Ansatzes ist die Tatsache, dass die PSA-Werte sowohl beim PCa als auch bei der BPH erhöht sein können, dies bei der BPH jedoch auf die Zellvermehrung und Volumenzunahme zurückzuführen ist, beim PCa dagegen auf die zerstörte Matrix. Die PSA-Dichte ist demzufolge beim Karzinom größer. Einige Studien belegten die Wirksamkeit dieser Methode [33, 97].

Die Prostatavolumenmessung mittels TRUS stellt bei diesem Verfahren einen limitierenden Faktor dar, da bereits geringe Messungenauigkeiten zu erheblichen Fehleinschätzungen der Prostatagröße führen [98]. Des Weiteren muss das

Prostatavolumen auch nicht zwangsläufig mit dem tPSA korreliert sein, da das PSA ausschließlich vom Prostataepithel produziert wird, das Epithel-Stroma-Verhältnis jedoch interindividuell schwankt [99].

1.3 Multivariate Auswertemodelle

Multivariate Modelle, wie zum Beispiel artifizielle neuronale Netzwerke (ANN) oder auf logistischer Regression (LR) basierte Nomogramme, die eine kombinierte Bewertung verschiedener Parameter ermöglichen, stellen eine weitere Option der Spezifitätsverbesserung des tPSA und des %fPSA dar.

1.3.1 Artifizielle neuronale Netzwerke

Bei artifiziellen neuronalen Netzwerken handelt es sich um mathematische Modelle, die im Falle einer Klassifikationsaufgabe bestimmte, gegebenenfalls unterschiedlich gewichtete Eingangsvariablen bewerten und die Wahrscheinlichkeit für die Zugehörigkeit zu einer bestimmten Klasse ermitteln. Nachdem 1958 das erste Modell, das sogenannte Perzeptron, erfolgreich entwickelt worden war [100], wurde das Prinzip des ANN zunächst in den Bereichen der Neurophysiologie, klinischen Chemie und Biochemie angewendet [101].

1994 gelang Snow et al. die Entwicklung des ersten ANN für die PCa-Diagnostik [102]. Seitdem wurden verschiedene neuronale Netzwerke für die Prädiktion des PCa erstellt [103-107]. Für die ANN-Modelle zeigte sich eine signifikante Verbesserung der Spezifität verglichen mit den Einzelvariablen [108, 109].

Mit dem Computerprogramm „ProstataClass“ wurde 2002 an der Charité ein ANN entwickelt, das auf der Basis der fünf Eingangsvariablen tPSA, %fPSA, DRU-Befund, Prostatavolumen und Lebensalter die Wahrscheinlichkeit für das Vorliegen eines PCa berechnet und somit eine Hilfe bei der Entscheidung für oder gegen eine Prostatastanzbiopsie darstellt [110].

In einer anschließenden multizentrischen Studie mit insgesamt 1188 Patienten konnte gezeigt werden, dass sich durch die Nutzung dieses Programms etwa 20-30% unnötig durchgeführter Biopsien vermeiden lassen [107].

1.3.2 Nomogramme

Ein Nomogramm, dem als statistische Methode meist die LR zugrunde liegt, ist definiert als die graphische Darstellung einer mathematischen Formel. Es besteht aus mehreren Skalen, auf denen bekannte Variablen aufgetragen sind, und einer Skala, an der das Ergebnis abgelesen werden kann [111].

Die der Prädiktion des PCa dienenden Nomogramme messen jeder Variable, beispielsweise dem tPSA, einen ihrer diagnostischen Wichtigkeit entsprechenden Punktwert zu. Aus diesen Punktwerten der Einzelvariablen entsteht das Ergebnis, das als Prozentzahl angegeben wird [112].

Innerhalb der PCa-Forschung wurden Nomogramme zur Vorhersage einer Reihe verschiedener Größen entwickelt. So wurden LR-Modelle zur Prädiktion eines PCa bei erstmaliger wie auch wiederholter Prostatastanzbiopsie entwickelt; sogenannte Staging-Nomogramme prognostizieren das pathologische Stadium bei radikaler Prostatektomie. Weitere Nomogramme dienen der Vorhersage eines biochemischen Rezidivs nach radikaler Prostataentfernung bzw. Radiotherapie oder prognostizieren die metastatische Progression und die Lebenserwartung der PCa-Patienten [113].

In der vorliegenden Arbeit werden nur Nomogramme, die die Wahrscheinlichkeit für das Vorliegen eines PCa bei erstmaliger Prostatastanzbiopsie angeben, näher untersucht. Einleitend erfolgt dazu ein kurzer Überblick.

Eastham et al. entwickelten 1999 das erste Nomogramm zur Vorhersage eines PCa bei erstmaliger Biopsie an Männern mit einem suspekten DRU-Befund und PSA-Werten im Bereich von 0-4 µg/l [114]. Kritikpunkte dieses Nomogrammes, das in der eigenen Studienpopulation eine gute Treffsicherheit aufwies, sind die fehlende externe Validierung und die eingeschränkte Möglichkeit einer generalisierten Anwendung, da die Mehrheit der urologischen Patienten entweder unauffällige DRU-Befunde oder aber PSA-Werte von > 4 µg/l aufweist.

Ein weiteres Nomogramm mit klinischen Daten, Laborparametern und Ultraschall-ergebnissen als Eingangsvariablen wurde von Garzotto et al. entwickelt [115]. Die Nutzung der Sextanten-Biopsie schränkt seine Anwendbarkeit ein, da heutzutage eher erweiterte Biopsieschemata zum Einsatz kommen. Eine externe Validierung erfolgte auch bei diesem Nomogramm nicht.

Zwei Nomogramme für die Vorhersage eines PCa wurden von Karakiewicz et al. an einer Studienpopulation von > 5000 Patienten entwickelt und extern validiert. Auch diese beiden Nomogramme basieren nur auf durchgeführten Sextanten-Biopsien [116].

Chun et al. wiesen nach, dass die Vorhersagekraft von Nomogrammen, die für Patienten mit Sextanten-Biopsie bestimmt waren, bei Anwendung auf eine Population mit erweitertem Biopsieschema signifikant abnimmt. Daher entwickelten sie für diese Patientengruppe ein Nomogramm auf der Basis der Eingangsvariablen Alter, tPSA, %fPSA, DRU-Befund und „Sampling density“ (Quotient aus Prostatavolumen und geplanter Anzahl an Biopsien) [117]. 2008 veröffentlichten Kawakami et al. zwei weitere Nomogramme für die Vorhersage eines PCa bei erweitertem Biopsieschema [118].

Im direkten Vergleich von ANN und Nomogrammen zur Vorhersage eines PCa berichteten einige Autoren über Vorteile von Nomogrammen [118-122], andere wiederum konnten eine Überlegenheit der ANN zeigen [104, 105, 123, 124].

Vergleicht man die beiden Methoden in sehr großen Studienpopulationen (>5000), wird in sieben von acht Studien kein Unterschied zwischen ANN und Nomogrammen gefunden [125]. Dies unterstützt die Hypothese, dass die Unterschiede nicht durch einen mathematischen Vorteil einer der beiden Methoden bedingt sind, sondern vielmehr die Eingangsvariablen, Populationen und die Nichtberücksichtigung der Methodenabhängigkeit von PSA-Werten ausschlaggebend sind [120, 124].

2. Zielsetzung der Arbeit

Damit das PCa einer kurativen Therapie zugänglich ist, ist die Früherkennung des Tumors wichtig. Die geringe Spezifität des PSA-Wertes führt allerdings dazu, dass bei etwa 60-80% aller durchgeführten Prostatastanzbiopsien kein Karzinom nachgewiesen wird [50]. Diese hohe Anzahl an unnötigen Biopsien führt zu enormen Kosten für das Gesundheitssystem und stellt für die Patienten eine große Belastung dar.

Nomogramme, mit denen die Wahrscheinlichkeit für das Vorliegen eines PCa ermittelt werden kann, sind nichtinvasive Verfahren zur Reduzierung dieser unnötigen Biopsien. Eine aktuelle Übersichtsarbeit zeigt, dass sie die Effizienz des tPSA deutlich verbessern [109]. Trotzdem ist die Diskrimination zwischen BPH und PCa bei Weitem nicht optimal.

Die Variabilität zwischen den verschiedenen PSA-Testsystemen ist ein bekanntes Problem im klinischen Alltag. Trotz der erfolgten Kalibrierung der unterschiedlichen tPSA- und fPSA-Tests mit dem WHO-Referenzmaterial konnte keine endgültige Harmonisierung der Testsysteme erreicht werden [96, 126, 127].

Eine Studie hat gezeigt, dass vor allem die Berücksichtigung des PSA-Testsystems für die optimale Nutzung eines ANN wichtig war [128]. Als Folge dessen wurde bereits ein ANN für fünf verschiedene PSA-Testsysteme entwickelt [129]. Dagegen ist bisher nicht bzw. äußerst unzureichend [130] untersucht, welchen Einfluss die PSA-Test-Variabilität auf Nomogramme ausübt. In externen Validierungen von Nomogrammen wurden die PSA-Testsysteme entweder nicht [116, 117] oder nur teilweise [118, 131] berücksichtigt.

Die Zielsetzung der vorliegenden Arbeit ist daher die Evaluierung der analytischen Differenzen zwischen den verschiedenen PSA-Testsystemen und deren Auswirkungen auf die Vorhersage eines PCa mittels Nomogrammen.

Diese Fragestellung wird im Folgenden anhand einer 798 Patienten umfassenden Population bearbeitet. Die tPSA- und fPSA-Werte wurden mit fünf häufig verwendeten Testsystemen gemessen. Anschließend erfolgte die retrospektive Analyse dieser Daten mit fünf anerkannten Nomogrammen.

Zusätzlich dient diese Arbeit der externen Validierung der fünf Nomogramme.

3. Material und Methoden

3.1 Patientenkollektiv

Es wurden retrospektiv die Daten von insgesamt 798 Patienten ausgewertet, bei denen die Blutentnahmen im Zeitraum von 2001 bis 2004 erfolgt waren.

Die Probanden waren aufgrund von erhöhten PSA-Werten, einer auffälligen DRU, Symptomen einer BPH, einem bereits bioptisch gesicherten PCa oder zur urologischen Vorsorgeuntersuchung in die Klinik und Poliklinik für Urologie der Charité - Universitätsmedizin Berlin, Campus Charité Mitte, überwiesen worden.

Bei allen Patienten wurde eine DRU durchgeführt und das Prostatavolumen mittels TRUS bestimmt.

Zum Zeitpunkt der Blutentnahmen waren alle Probanden bezüglich einer Prostataerkrankung weder medikamentös noch operativ vorbehandelt. Die Diagnosen wurden mittels einer 8- oder 10-fachen Stanzbiopsie gesichert.

Da mit dem PSA-Test von Bayer Diagnostics für 18 der 798 Patienten ein höherer cPSA- als tPSA-Wert ermittelt wurde und demzufolge negative %fPSA-Ratios resultieren würden, erfolgte eine Unterteilung in zwei Kohorten: eine mit 798 und eine mit 780 Patienten.

Von der Gesamtheit der 798 Patienten hatten 468 (58,6%) ein histologisch gesichertes PCa. Die übrigen 330 Probanden wiesen größtenteils die Diagnose einer BPH oder Prostatitis auf. Einige Biopate enthielten auch normales Prostatagewebe. Alle Patienten, die nicht an einem PCa leiden, werden im Folgenden als BPH-Gruppe zusammengefasst. Bei der Population aus 780 Patienten wurde bei 455 (58,3%) Männern ein PCa diagnostiziert.

Die mit dem Abbott-Test ermittelten tPSA Werte der 798 Probanden befanden sich im Bereich von 0,54 bis 23,4 µg/l, während die %fPSA-Ratio den Bereich von 3,69 bis 73,9 % umfasste.

3.2 Probenmaterial

Die Blutproben wurden mindestens drei bis vier Wochen nach Manipulationen an der Prostata, wie der rektalen Untersuchung oder dem transrektalen Ultraschall, sowie vor der Prostatastanzbiopsie entnommen, um mögliche Fehlmessungen zu vermeiden. Anschließend wurden die Proben 15 Minuten bei 4° Celsius und 1600 G zentrifugiert und das aus dem Überstand gewonnene Serum bei -80° Celsius eingefroren. Am Tag der Messung wurden die Serumproben aufgetaut und innerhalb von drei Stunden analysiert.

3.3 PSA-Bestimmungen

Die Messungen der PSA-Werte sind von Juni bis August 2004 an den folgenden fünf vollautomatischen Immunoassay-Analysegeräten nach den Anweisungen der Hersteller erfolgt:

Für die Messung von tPSA und fPSA sind der AxSYM[®]- (Abbott Diagnostics GmbH, Wiesbaden, Deutschland), der Access[®]- (Hybritech[®]-PSA, TANDEM R, Beckman Coulter Inc., Fullerton, USA), der Immulite 2000[®]- (Diagnostic Products Corp, DPC-Biermann GmbH, Bad Nauheim, Deutschland) sowie der Elecsys 2010[®]-Test (Roche Diagnostics GmbH, Mannheim, Deutschland) verwendet worden. Mit dem ADVIA Centaur[®]-Test (Bayer Diagnostics, Tarrytown, USA) wurden tPSA und cPSA bestimmt.

Um die fPSA-Testsysteme besser vergleichen zu können, wurde aus den cPSA-Werten des ADVIA Centaur[®]-Tests das fPSA wie folgt berechnet:

$fPSA = tPSA - cPSA$. Die Formel $\%fPSA = (fPSA/tPSA) \times 100$ diente anschließend der Ermittlung des prozentualen Anteils des fPSA am tPSA.

Tabelle 1 gibt einen Überblick über die PSA-Testsysteme und erklärt die im Folgenden verwendeten Abkürzungen.

Tabelle 1. Übersicht über die PSA-Testsysteme mit den verwendeten Abkürzungen und den vollständigen Bezeichnungen sowie Herstellern.

Abkürzung	Vollständige Bezeichnung und Hersteller
Abbott	AxSYM [®] -Test Abbott Diagnostics GmbH, Wiesbaden, Deutschland
Bayer (jetzt Siemens)	ADVIA Centaur [®] -Test Bayer Diagnostics, Tarrytown, USA
Beckmann	Access [®] -Test (Hybritech [®] -PSA, TANDEM R) Beckman Coulter Inc., Fullerton, USA
DPC (jetzt Siemens)	Immulite 2000 [®] -Test Diagnostic Products Corp, DPC-Biermann GmbH, Bad Nauheim, Deutschland
Roche	Elecsys 2010 [®] -Test Roche Diagnostics GmbH, Mannheim, Deutschland

3.4 Klinische Untersuchungen

Bei allen Probanden wurde der klinische Status erhoben. Sämtliche Patienten unterzogen sich einer DRU. Eine gute Abgrenzbarkeit der Prostata, eine prall-elastische Struktur, eine glatte Oberfläche und keine Induration der Drüse galten hierbei als Kriterien für einen unauffälligen, negativen Befund: DRU = 0.

Dagegen wurden Drüsen mit palpablen Gewebeverhärtungen, Knotenbildungen und schlechter Abgrenzbarkeit als auffällig und karzinomsuspekt und somit positiv definiert: DRU = 1.

Das Prostatavolumen wurde mittels TRUS bestimmt. Bei der Berechnung fand die prolata Ellipsoid-Formel (Höhe x Breite x Länge x 0,52) Anwendung. Bei allen Patienten erfolgte eine 8- bzw. 10-fache Stanzbiopsie zur histologischen Sicherung der Diagnose.

3.5 Nomogramme

Der vollständige Datensatz der Studienpopulation umfasst das Alter, das Prostatavolumen, den TRUS-Befund, den DRU-Status (0 = unauffällig, 1 = suspekt)

Nomogramm	I	II	III	IV	V
-----------	---	----	-----	----	---

sowie die gemessenen tPSA- und berechneten %fPSA-Werte analysiert mit den fünf verschiedenen PSA-Testsystemen von Abbott, Bayer, Beckmann, DPC und Roche.

Alle Patientendaten wurden für jeden der fünf PSA-Tests einzeln mit fünf verschiedenen Nomogrammen ausgewertet und die Ergebnisse hinsichtlich möglicher Differenzen zwischen den PSA-Testsystemen analysiert.

Mit Hilfe der Nomogramme wurde für jeden einzelnen Patienten die Wahrscheinlichkeit ermittelt auf Grund seines individuellen Datensatzes an einem PCa erkrankt zu sein.

Tabelle 2 gibt einen Überblick über die fünf angewandten Nomogramme bezüglich ihrer Eingangsvariablen, den Patientenzahlen, dem PSA-Bereich und den PSA-Tests, auf deren Basis die Nomogramme entwickelt wurden. Die Nomogramme I und II sind von Karakiewicz et al. entwickelt worden [116]. Während Nomogramm I die Wahrscheinlichkeit aus den Größen Alter, DRU und tPSA errechnet, enthält das zweite Nomogramm zusätzlich %fPSA als Eingangsvariable. Das Nomogramm III ist erhältlich unter <http://www.nomogram.org/> und wurde geschaffen auf der Grundlage von Alter, DRU, tPSA, %fPSA und der sogenannten „Sampling density“ (Quotient aus Prostatavolumen und geplanter Anzahl an Biopsien) [117]. Kawakami et al. entwickelten die Nomogramme IV und V mit den Eingangsvariablen Alter, DRU, tPSA und %fPSA (Nomogramm IV und V) und den zusätzlichen Faktoren TRUS-Befund und Prostatavolumen (nur Nomogramm V) [118].

Autoren	Karakiewicz et al. [116]		Chun et al. [117]	Kawakami et al. [118]	
Variablen	Alter DRU tPSA	Alter DRU tPSA %fPSA	Alter DRU tPSA %fPSA Sampling density	Alter DRU tPSA %fPSA	Alter DRU tPSA %fPSA TRUS-Befund Prostatavolumen
PSA-Test	nicht angegeben		nicht angegeben	Beckmann, später Abbott	
Patientenanzahl	4193	1762	1162	1509	1509
PSA-Bereich	0-50		nicht bekannt	0-20	

Tabelle 2. Übersicht über die angewandten Nomogramme mit ihren Eingangsvariablen und den zugrunde liegenden PSA-Tests (Erläuterungen siehe Text).

3.6 Statistische Auswertungen

Die statistischen Berechnungen erfolgten mit den Softwareprogrammen MedCalc[®] Version 10.4.8.0 (MedCalc Software, Mariakerke, Belgien) und SPSS 17.0 für Windows (SPSS, Chicago, USA). Das Programm MATLAB (The MathWorks, Inc., Natick, Mass., USA) diente der Entwicklung der Abbildungen.

Die Diskriminationsfähigkeit der Parameter tPSA und %fPSA und der Nomogramme wurde mit der Receiver-operating-characteristic (ROC)-Analyse bewertet [132], wobei die Flächen unter der ROC-Kurve (Area under the curve (engl.) = AUC) und die Spezifitäten bei 95% Sensitivität bzw. die Sensitivitäten bei 95% Spezifität verglichen wurden. Die dafür nötigen Berechnungen wurden ebenfalls mit MedCalc[®] Version 10.4.8.0 (MedCalc Software, Mariakerke, Belgien) durchgeführt.

Als statistische Tests wurden der Mann-Whitney-U-, der Friedman- und der Wilcoxon-Test angewendet und $P < 0,05$ wurde als statistisch signifikant definiert.

Die Kalibrierung der Nomogramme erfolgte durch den Vergleich der vorhergesagten PCa-Wahrscheinlichkeiten und der beobachteten PCa-Raten bei der Biopsie [133].

Hierbei wurde die Kohorte mit den 780 Patienten nach den jeweiligen PCa-Wahrscheinlichkeiten aufsteigend sortiert und anschließend zunächst in 5, dann in 10, 20 und 50 Gruppen geteilt, die demzufolge aus 156, 78, 39 und 15 Patienten bestanden. Für jede Gruppe wurde die mittlere vorhergesagte und beobachtete PCa-Wahrscheinlichkeit berechnet. Der Intraklassenkorrelationskoeffizient (ICC) [134] wurde errechnet um die Übereinstimmung dieser Wahrscheinlichkeiten zu beurteilen. Ein ICC von 1 wäre optimal.

Die Abbildungen 2a-d veranschaulichen den Prozess der Gruppenbildung am Beispiel der Abbott-Test-Daten innerhalb des Nomogramms IV. Teilt man die Studienpopulation in 5, 10 oder 20 Gruppen, so ist der ICC auf hohem Niveau stabil (0,96-0,97) (Abbildungen 2a-c). Er sinkt dagegen deutlich ab, wenn 50 Gruppen gebildet werden und damit weniger als 2% aller Patienten in einer Gruppe enthalten sind (Abbildung 2d). Zum Vergleich der verschiedenen PSA-Testsysteme innerhalb der fünf Nomogramme wurde daher eine mittlere Gruppengröße (39 Patienten, also 5% der Studienpopulation, pro Gruppe) gewählt.

Zur besseren graphischen Darstellung der Beziehung zwischen vorhergesagten PCa-Wahrscheinlichkeiten und beobachteten PCa-Raten wurde eine Ausgleichskurve (Spline-Approximation) durch die Punkte gelegt. Diese Art der Darstellung wird in der Literatur als Kalibration bezeichnet. An dieser Kurve kann abgelesen werden, wie das Modell bei einem bestimmten Wert der Wahrscheinlichkeit die wirkliche Situation unterschätzt, überschätzt oder richtig einschätzt (45°-Linie).

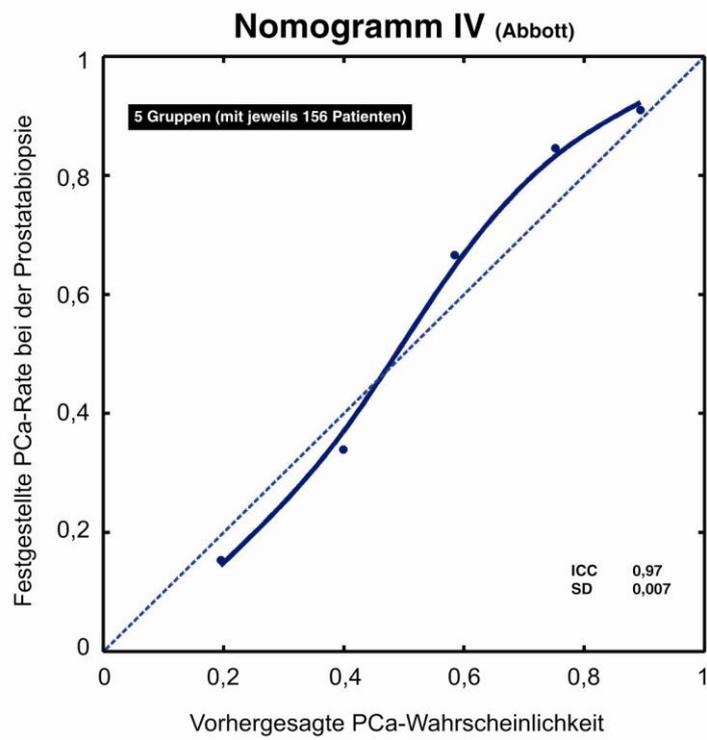


Abbildung 2a.

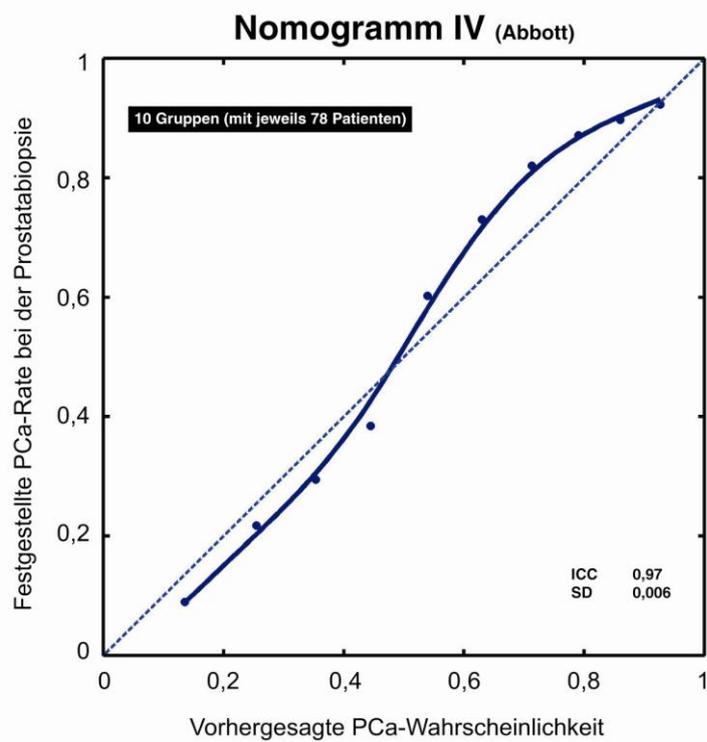


Abbildung 2b.

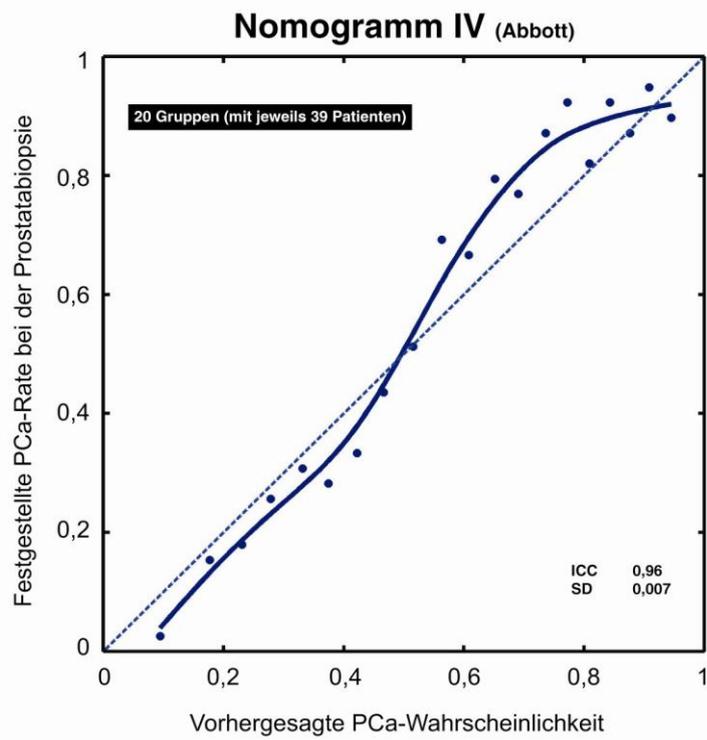


Abbildung 2c.

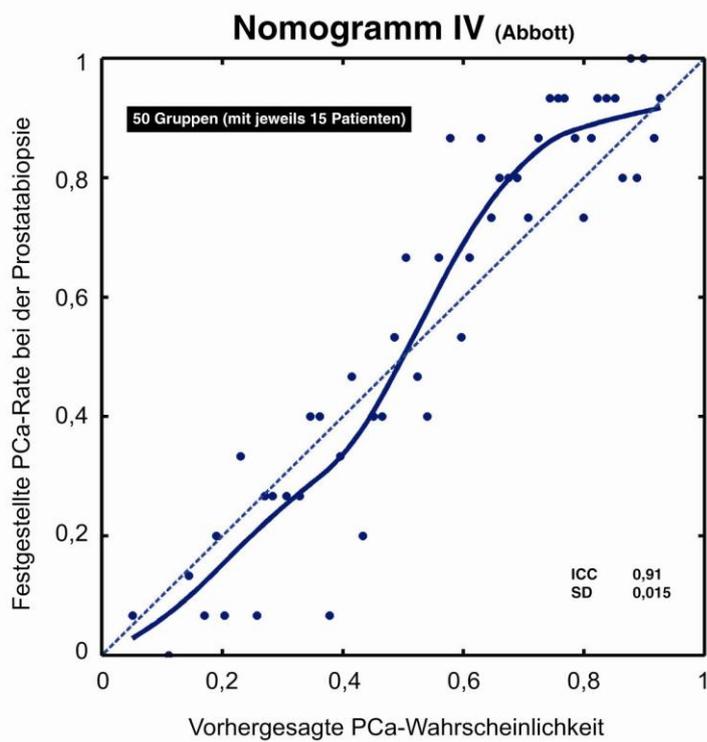


Abbildung 2d.

Abbildungen 2a-d.

Kalibrierung des Nomogramms IV am Beispiel der Abbott-Daten für 5, 10, 20 und 50 Gruppen, die jeweils aus 156, 78, 39 und 15 Patienten bestehen. Der jeweilige Intraklassenkorrelationskoeffizient (ICC) und die Standardabweichung (SD) sind angegeben. Im Fall der vollständigen Übereinstimmung der vorhergesagten Wahrscheinlichkeiten mit den beobachteten PCa-Raten würden sich alle Punkte auf der 45°-Linie befinden. Zur Veranschaulichung des Zusammenhangs sind Ausgleichskurven (Spline-Approximation) eingezeichnet. Weitere Erläuterungen siehe Text.

4. Ergebnisse

	Kohorte n=798	Kohorte n=780
--	---------------	---------------

4.1 Charakterisierung des Patientenkollektivs

Tabelle 3 stellt die Kohorten mit 798 und 780 Patienten gegenüber und bietet eine Übersicht über die tPSA- und %fPSA-Werte, die Altersverteilung, die Prostata-volumina und die DRU-Befunde der gesamten Population, der PCa- und der BPH-Patienten. Da im Vergleich der beiden Kohorten für keine der Variablen ein statistisch signifikanter Unterschied gefunden wurde ($P= 0,44$ bis $0,88$), soll im Folgenden nur das aus 780 Patienten bestehende Kollektiv analysiert werden.

Die Population bestand aus 455 (58,3%) Patienten mit einem PCa und 325 mit einer BPH. Diese beiden Gruppen unterschieden sich in jedem Parameter signifikant (P immer $< 0,0001$). Die BPH-Patienten waren mit einem Altersmedian von 66 Jahren drei Jahre älter als die Karzinompatienten und hatten ein deutlich größeres medianes Prostata-volumen (46 ml). Während 60,7% der PCa-Patienten einen positiven Tastbefund aufwiesen, war dieser nur bei 8,62% der BPH-Patienten suspekt. Die tPSA-Werte der PCa-Patienten waren signifikant höher als die der BPH-Patienten; die %fPSA-Werte lagen bei den Karzinompatienten dagegen signifikant niedriger.

Der Vergleich der tPSA- und %fPSA-Werte erbrachte fast ausnahmslos statistisch signifikante Unterschiede zwischen den fünf PSA-Testsystemen ($P < 0,0001$ bis $< 0,003$). Nur die mit dem Abbott- und Bayer-Test ermittelten medianen tPSA-Werte unterschieden sich nicht ($P= 0,86$). Die größten Unterschiede wurden, sowohl bei den tPSA- als auch bei den %fPSA-Werten, zwischen den PSA-Tests von Bayer und DPC beobachtet. Betrachtet man die gesamte Patientenpopulation, wurde mit dem Bayer-Test der niedrigste mediane tPSA-Wert ($5,5 \mu\text{g/l}$) gemessen (kein signifikanter Unterschied zum Abbott-Test), während der DPC-Test mit $7,04 \mu\text{g/l}$ den höchsten medianen Wert aufwies. Der höchste %fPSA-Median der gesamten Patientengruppe betrug 16,4% und wurde mit dem Bayer-Test ermittelt, die mit 10,6% niedrigste %fPSA-Ratio war beim DPC-Test sichtbar.

	Gesamt	PCa	BPH	Gesamt	PCa	BPH
n	798	468	330	780	455	325
Alter [Jahre]						
Median	64	63	66	64	63	66
Bereich	38-85	43-79	38-85	40-85	43-79	40-85
Prostatavolumen [ml]						
Median	38	33,8	45,5	38	34	46
Bereich	10-180	10-110	13-180	10-180	10-110	13-180
Anzahl positiver DRU						
n	314	284	30	304	276	28
%	39,4	60,7	9,09	39,0	60,7	8,62
tPSA [$\mu\text{g/l}$]						
Abbott						
Median	5,68	6,68	4,01	5,6	6,63	3,98
Bereich	0,54-23,4	0,67-23,4	0,54-22,9	0,54-23,4	0,67-23,4	0,54-22,9
Bayer						
Median	*	*	*	5,5	6,54	4,01
Bereich	*	*	*	0,69-24,5	0,70-20,7	0,69-24,5
Beckmann						
Median	6,63	7,76	4,69	6,54	7,68	4,61
Bereich	0,49-27,0	0,86-24,0	0,49-27,0	0,49-27,0	0,86-24,0	0,49-27,0
DPC						
Median	7,2	8,55	5,23	7,04	8,48	5,15
Bereich	0,76-24,9	0,96-24,3	0,76-24,9	0,76-24,9	0,96-24,3	0,76-24,9
Roche						
Median	6,63	7,75	4,85	6,53	7,69	4,71
Bereich	0,56-29,5	1,18-25,7	0,56-29,5	0,56-29,5	1,18-25,7	0,56-29,5
%fPSA [%]						
Abbott						
Median	15,9	12,4	22,3	16,2	12,7	22,5
Bereich	3,69-73,9	3,69-71,7	3,74-73,9	3,74-73,9	3,85-71,66	3,74-73,9
Bayer[#]						
Median	*	*	*	16,4	12,9	22,2
Bereich	*	*	*	0,51-72,4	0,51-63,2	2,17-72,4
Beckmann						
Median	12,8	10,3	18,3	13,1	10,4	18,4
Bereich	2,52-69,4	3,10-49,9	2,52-69,4	2,52-69,4	3,10-49,9	2,52-69,4
DPC						
Median	10,4	8,02	14,5	10,6	8,22	14,6
Bereich	0,70-54,3	0,70-45,3	2,29-54,3	0,70-54,3	0,70-45,3	2,29-54,3
Roche						
Median	12,3	10,1	16,8	12,5	10,3	16,9
Bereich	2,07-51,9	2,07-36,9	2,13-51,9	2,13-51,9	3,05-36,9	2,13-51,9

Tabelle 3. Charakteristika des Patientenkollektivs: Mediane und Bereiche für Alter, Prostatavolumen, positive DRU, tPSA und %fPSA für alle, PCa- und BPH-Patienten vergleichend für die Kohorte mit 798 und 780 Patienten.

* keine Angabe, da bei 18 Patienten cPSA > tPSA und somit negative %fPSA-Ratios resultieren würden

[#] %fPSA bei Bayer-Test wie folgt berechnet: $100 - (c/tPSA \times 100)$

4.2 Analytik und Kalibrierung der Nomogramm-Ausgänge

Tabelle 4 präsentiert die medianen PCa-Wahrscheinlichkeiten, die die fünf Nomogramme für die 780 Patienten anhand deren PSA-Werten (gemessen mit den PSA-Testsystemen von Abbott, Bayer, Beckmann, DPC und Roche) sowie den klinischen Parametern vorhergesagt haben.

Nomogramm I, das die Wahrscheinlichkeit für das Vorliegen eines PCa aus den Eingangsvariablen Alter, DRU-Befund und tPSA errechnet, gab mit medianen Werten im Bereich von 0,24-0,26 die niedrigsten Wahrscheinlichkeiten an. Die höchsten medianen Wahrscheinlichkeiten präsentierte das Nomogramm IV (0,59-0,76).

Je nach Nutzung des Nomogramms und PSA-Testsystems wurde also für die gleichen Patienten eine PCa-Wahrscheinlichkeit im Bereich von 0,24-0,76 angegeben.

Betrachtet man die Nomogramme einzeln und analysiert die vorhergesagten PCa-Wahrscheinlichkeiten hinsichtlich der Differenzen zwischen den verschiedenen Testsystemen, werden große Unterschiede deutlich. Die Nutzung von Nomogramm I mit den Abbott- und Bayer-Daten hatte zwar identische mediane Werte zur Folge ($P = 0,76$), alle anderen Vergleiche der Nomogramm-Ausgänge führten jedoch zu statistisch signifikanten Unterschieden (P immer $< 0,0001$, nur beim Abbott- und Bayer-Test $P < 0,05$).

Nomogramm IV zeigte hierbei die stärksten Abweichungen zwischen den PSA-Testsystemen. Während der Einsatz der Abbott-Daten zu einem medianen Ergebnis von 0,59 führte, wurde den gleichen Patienten unter Berücksichtigung der DPC-Daten eine PCa-Wahrscheinlichkeit von 0,76 prognostiziert.

Bei den Nomogrammen II, III und V waren es ebenfalls die PSA-Tests von Abbott und DPC, die zu den größten Unterschieden in den medianen Nomogramm-Ausgängen führten.

Das Nomogramm I gab die niedrigsten PCa-Wahrscheinlichkeiten bei Nutzung der PSA-Tests von Bayer (0,24) und Abbott (0,25) an. Der DPC-Test, mit dem die höchsten tPSA- und niedrigsten %fPSA-Werte gemessen wurden, führte in allen fünf Nomogrammen zu den höchsten PCa-Wahrscheinlichkeiten.

Tabelle 4. Mediane und Bereiche für die vorhergesagten PCa-Wahrscheinlichkeiten.

	Vorhergesagte PCa-Wahrscheinlichkeiten		
Nomogramm und PSA-Test [§]	Median	Bereich	Standardabweichung
Nomogramm I			
Abbott	0,25 [#]	0,11-0,72	0,15
Bayer	0,24 [#]	0,11-0,70	0,15
Beckmann	0,26	0,11-0,73	0,15
DPC	0,26	0,11-0,74	0,16
Roche	0,26	0,11-0,77	0,15
Nomogramm II			
Abbott	0,42*	0-0,92	0,27
Bayer	0,45*	0-0,93	0,27
Beckmann	0,49	0-0,93	0,26
DPC	0,54	0,01-0,94	0,24
Roche	0,50	0,02-0,94	0,25
Nomogramm III			
Abbott	0,59*	0,02-0,95	0,24
Bayer	0,59*	0,02-0,96	0,23
Beckmann	0,65	0,02-0,97	0,22
DPC	0,70	0,06-0,97	0,21
Roche	0,66	0,07-0,98	0,21
Nomogramm IV			
Abbott	0,59*	0,02-0,97	0,25
Bayer	0,61*	0,02-1,00	0,26
Beckmann	0,68	0,01-0,98	0,24
DPC	0,76	0,03-1,00	0,23
Roche	0,71	0,03-0,98	0,23
Nomogramm V			
Abbott	0,48*	0,01-1,00	0,29
Bayer	0,50*	0,01-0,99	0,29
Beckmann	0,56	0,01-1,00	0,29
DPC	0,62	0,03-1,00	0,28
Roche	0,57	0,02-1,00	0,28

[§] Signifikanzniveau von $P < 0,0001$ zwischen den Wahrscheinlichkeiten der verschiedenen PSA-Tests innerhalb eines Nomogramms soweit nicht anders gekennzeichnet

[#] kein signifikanter Unterschied der Nomogramm-Wahrscheinlichkeiten ($P=0,76$)

* Signifikanzniveau von $P < 0,05$

Das Maß der Übereinstimmung der vorhergesagten PCa-Wahrscheinlichkeiten durch die Nomogramme und der tatsächlichen PCa-Raten bei der Prostatabiopsie stellt eine Möglichkeit dar, die Qualität von Nomogrammen als PCa-Vorhersagemodelle zu beurteilen. Die Kalibrationskurven (Abbildungen 3a-e) zeigen die diagnostische Genauigkeit der fünf Nomogramme. Zusätzlich erlauben die Abbildungen den direkten Vergleich der verschiedenen PSA-Testsysteme.

Abbildung 3a zeigt für alle PSA-Testsysteme eine deutliche Abweichung von der 45°-Linie. Die ICC-Werte der PSA-Tests befanden sich hier im Bereich von 0,46-0,52. Nomogramm I prognostizierte den Patienten durchgehend eine zu niedrige PCa-Wahrscheinlichkeit; teilweise war die bei der Biopsie festgestellte PCa-Rate doppelt so groß wie die vorausgesagte Wahrscheinlichkeit.

Diese Darstellung macht deutlich, dass es sich beim Nomogramm I ohne Berücksichtigung des %fPSA um ein ungenaues Modell zur Vorhersage eines PCa handelt, das sich jedoch nicht von den Unterschieden zwischen den PSA-Testsystemen beeinflussen lässt.

Die Abbildungen 3b-e lassen erkennen, dass die Nomogramme II-V relativ exakte Modelle sind; teilweise waren jedoch große Differenzen hinsichtlich der verschiedenen Testsysteme ersichtlich.

Verwendete man im Nomogramm II mit dem DPC-Test gemessene PSA-Werte, so sagte das Modell teilweise eine zu hohe Wahrscheinlichkeit vorher, während bei Nutzung der Abbott-Test-Daten zu niedrige Werte prognostiziert wurden (Abbildung 3b).

Die Verlaufskurven dieser beiden Testsysteme differierten im Nomogramm II am stärksten; die den PSA-Tests von Beckmann und Roche zugeordneten Kurven waren dagegen nahezu identisch.

Die Abbildung 3c zeigt wieder die größten Unterschiede zwischen dem DPC- und Abbott-Test, in Abbildung 3d weicht der Bayer-Test noch etwas stärker als der von Abbott vom DPC-Test ab. In diesen beiden Abbildungen, wie auch in Abbildung 3b, wird deutlich, dass sowohl die Verlaufskurven von Abbott und Bayer als auch die von Beckmann und Roche nahezu identisch sind, während die dem DPC-Test zugehörige Kurve einzeln verläuft.

In Abbildung 3e sind die Unterschiede zwischen den PSA-Testsystemen deutlich geringer, trotzdem lassen sich noch analoge Kurvenverläufe der Tests von Abbott und Bayer sowie von Beckmann und Roche erkennen. Der DPC-Test zeigte mit einem ICC von 0,99 eine exzellente Darstellung im Nomogramm V.

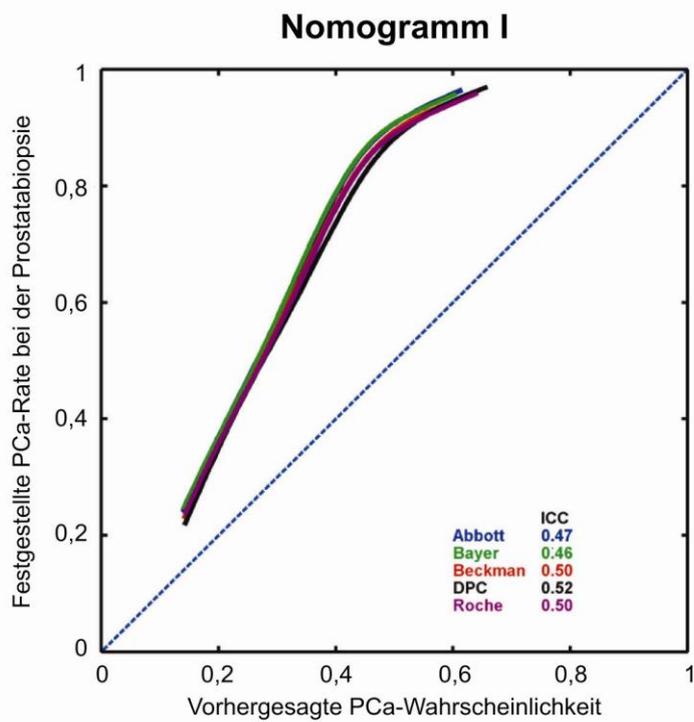


Abbildung 3a.

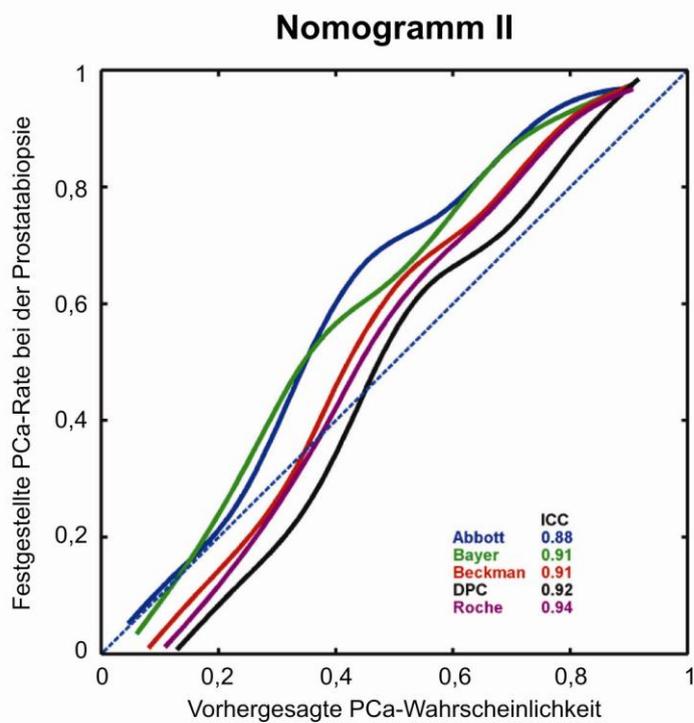


Abbildung 3b.

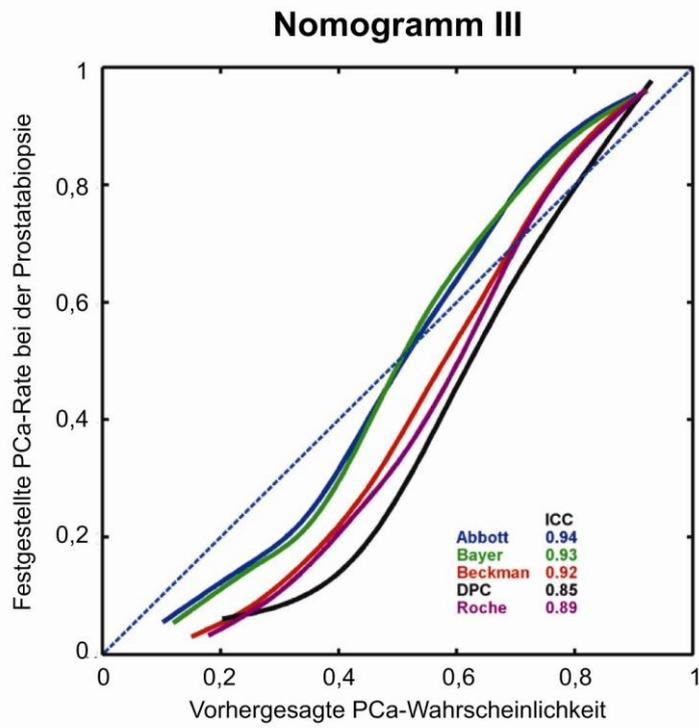


Abbildung 3c.

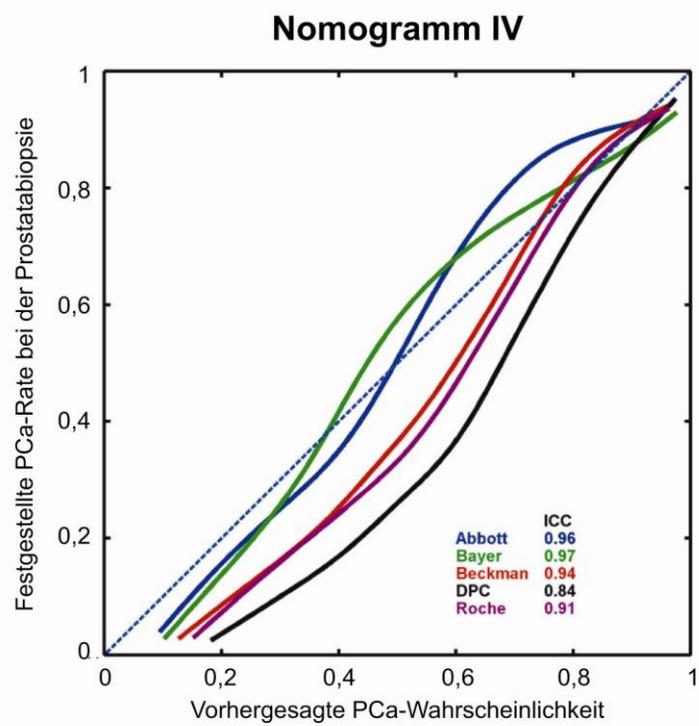


Abbildung 3d.

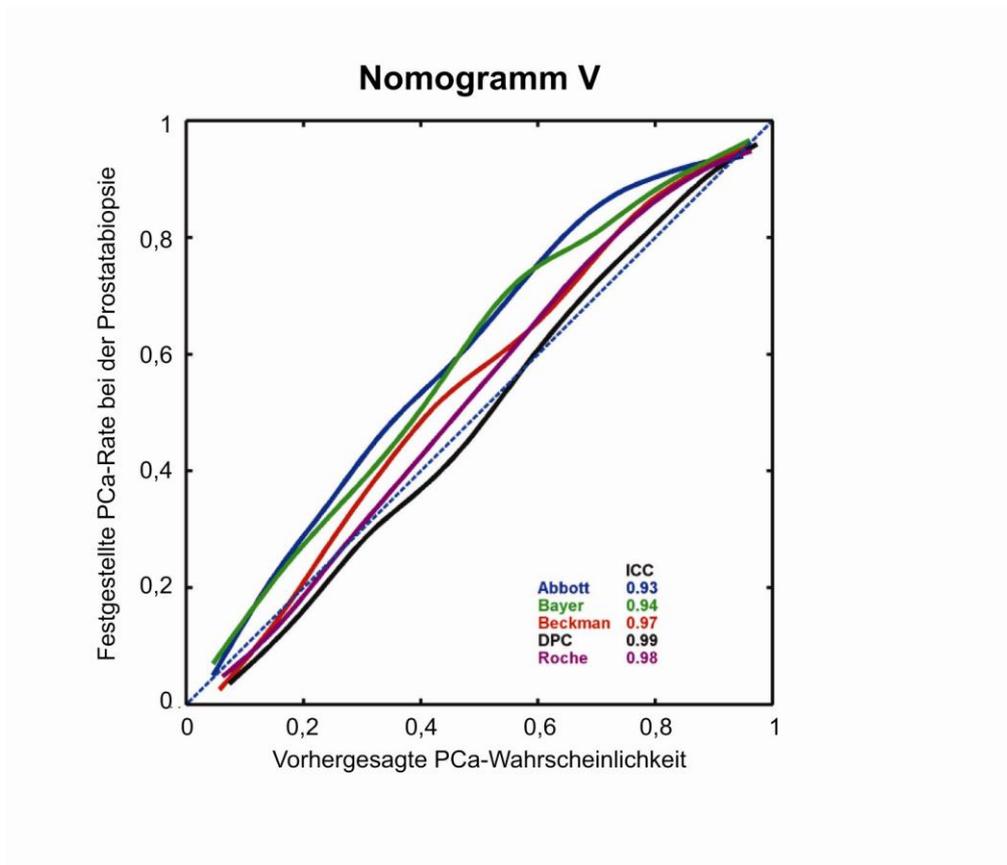


Abbildung 3e.

Abbildungen 3a-e. Darstellung der vorhergesagten PCa-Wahrscheinlichkeiten in Bezug auf die beobachteten PCa-Raten bei der Biopsie mittels Kalibrationskurven für die Nomogramme I bis V. Der Intraklassenkorrelationskoeffizient (ICC), ein Maß für die Beständigkeit der beobachteten und vorhergesagten Werte, ist für jeden der fünf PSA-Tests angegeben.

4.3 Diskrimination

4.3.1 Analyse bei verschiedenen Cut-Off-Werten

Die Abbildungen 4a-e zeigen die Sensitivitäten und Spezifitäten bei verschiedenen Cut-Off-Werten für die Nomogramm-Wahrscheinlichkeit und verdeutlichen Unterschiede in der Diskrimination zwischen PCa und BPH für die fünf Nomogramme mit den fünf verschiedenen PSA-Testsystemen.

Im Folgenden dienen die Cut-Off-Werte 0,25, 0,5 und 0,75 der Demonstration der teilweise erheblichen Unterschiede in den Sensitivitäten und Spezifitäten je nach verwendetem PSA-Test. Bei einem gewählten Cut-Off-Wert von z.B. 0,25 wurden die maximalen Spezifitätsunterschiede innerhalb des Nomogramms II gefunden: Während die Spezifität mit den PSA-Werten von Bayer 54% betrug, sank sie bei Nutzung der DPC-Werte auf 20% ab (Abbildung 4b). Die Abweichungen in den Sensitivitäten waren bei einem Cut-Off von 0,25 relativ geringer. Beim Nomogramm V war mit 7% Differenz zwischen dem DPC- und Abbott-Test die größte Abweichung zu verzeichnen (Abbildung 4e).

Die größte Variabilität sowohl in den Sensitivitäten als auch in den Spezifitäten wurde bei einem Cut-Off-Wert von 0,5 innerhalb des Nomogramms IV beobachtet. Wurden die PSA-Werte der Patienten mit dem Abbott-Test gemessen, gelang es mit dem Nomogramm IV 73% der BPH-Patienten (die in diesem Fall die Gesunden darstellen) richtig zu klassifizieren. War dagegen der DPC-Test zur PSA-Messung verwendet worden, sank die Spezifität an diesem fixen Grenzwert von 0,5 um 30% auf 43% ab. Die Sensitivität schwankte innerhalb des Nomogramms IV im Bereich von 81% (Bayer) bis 95% (DPC) (Abbildung 4d).

Die Analyse der Sensitivitäten und Spezifitäten bei einem Cut-Off-Wert von 0,75 zeigte die größten Unterschiede wiederum innerhalb des Nomogramms IV (Abbildung 4d). Dienten die mit dem DPC-Test gemessenen PSA-Werte als Eingangsvariablen im Nomogramm IV, so konnten 74,5% der PCa-Patienten richtig diagnostiziert werden. Die Anwendung der Bayer-PSA-Werte im Nomogramm IV

fürte dagegen zu einem Absinken der Sensitivität auf 46%. Die Spezifität befand sich im Bereich von 80% (DPC) bis 92% (Abbott).

Das Nomogramm I wies nur sehr geringe Unterschiede zwischen den fünf Testsystemen auf (Abbildung 4a). Bei einem Cut-Off-Wert von 0,5 reichten die Sensitivitäten von 23% (Abbott und Bayer) bis 33% (DPC), die Spezifitäten lagen noch näher beieinander und befanden sich im Bereich von 95% (DPC) bis 97% (Abbott).

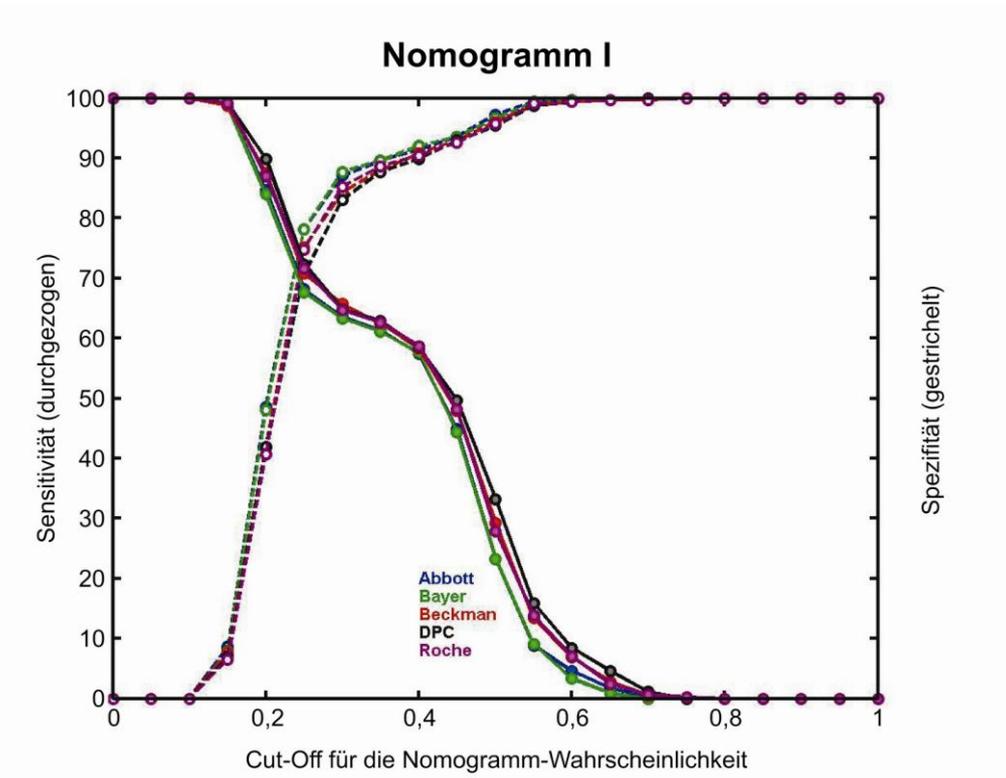


Abbildung 4a.

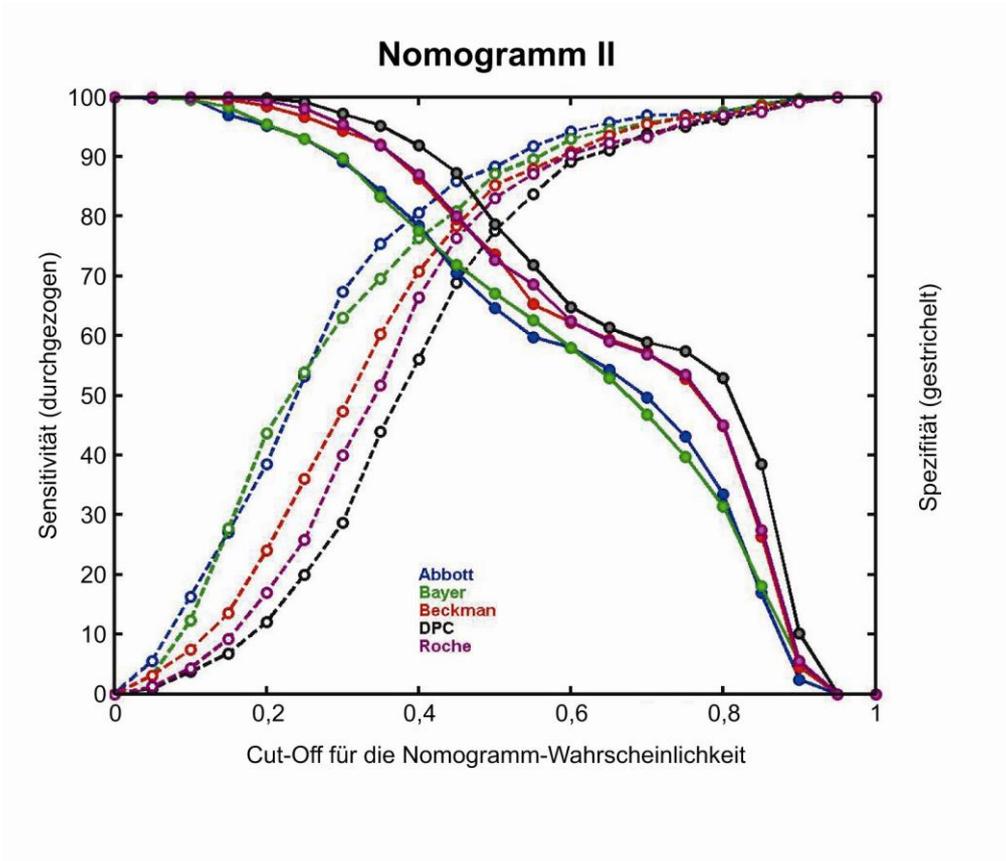


Abbildung 4b.

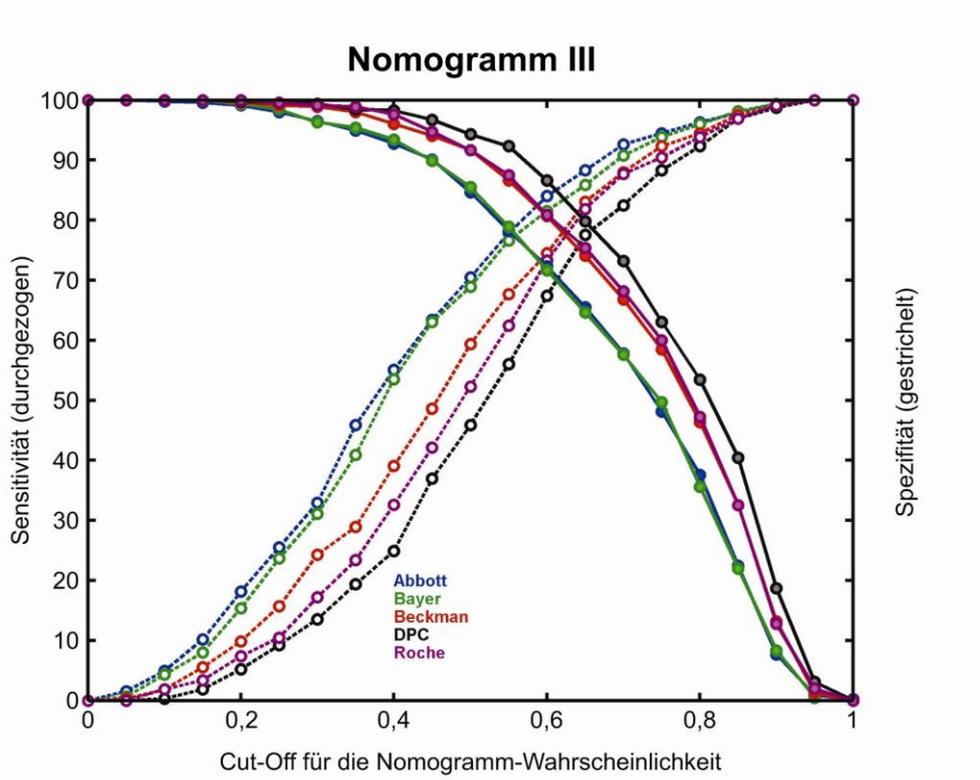


Abbildung 4c.

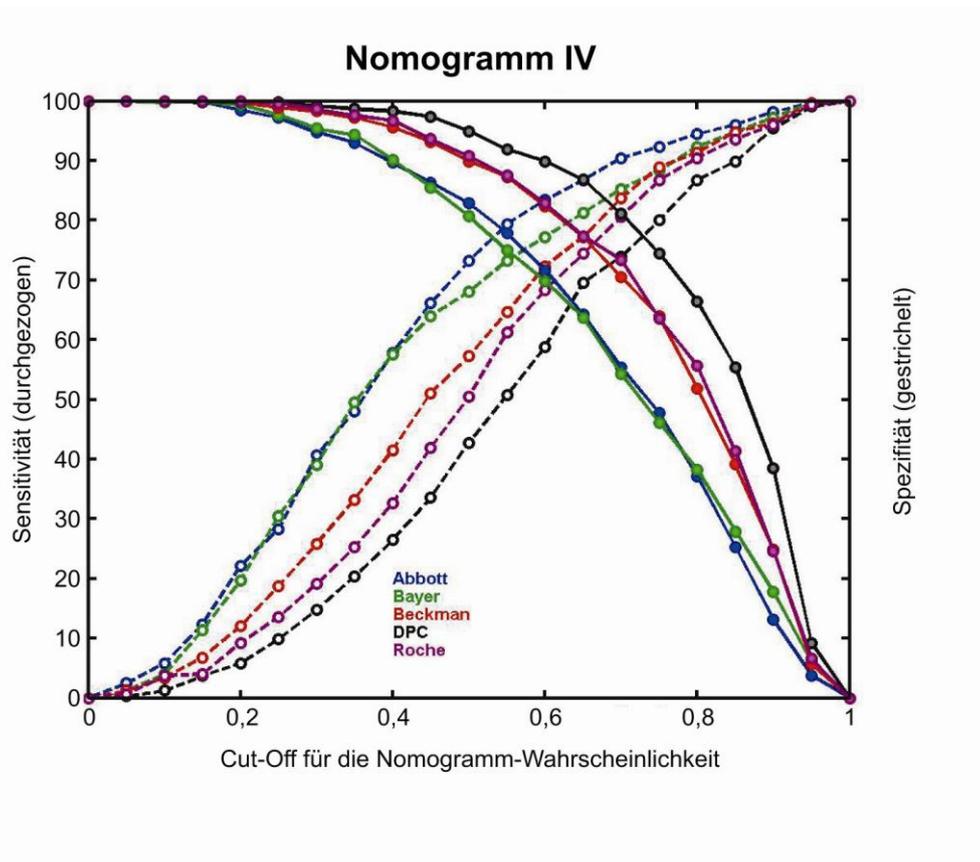


Abbildung 4d.

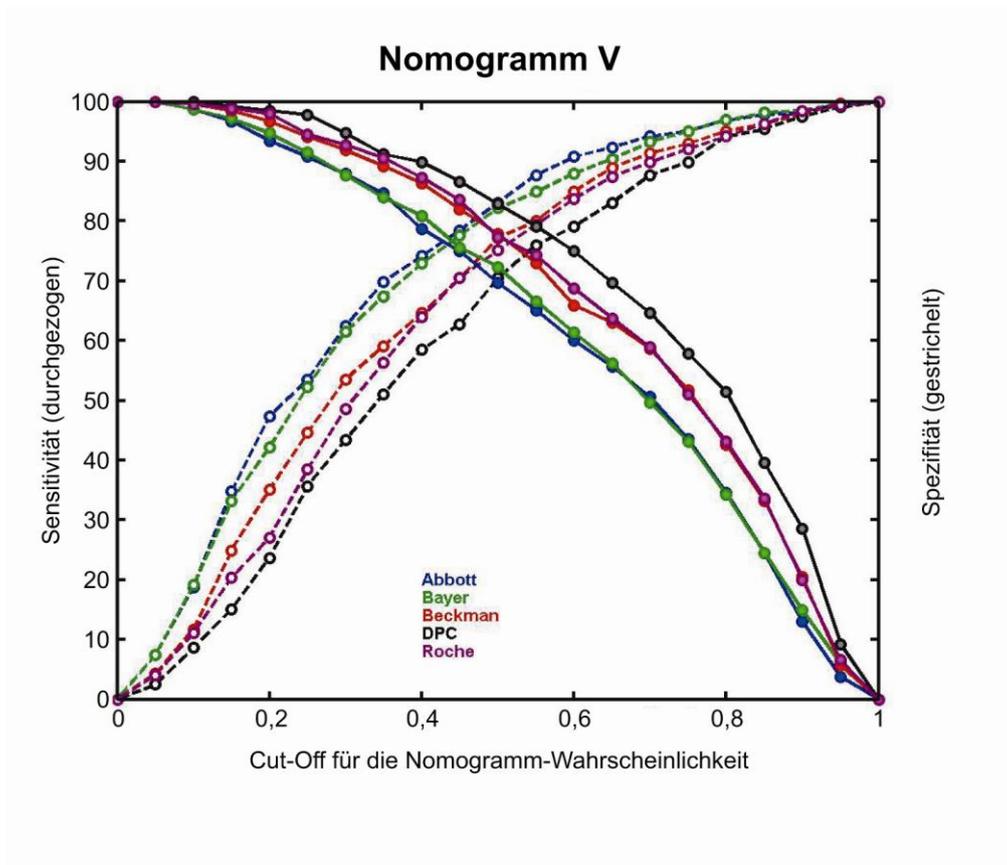


Abbildung 4e.

Abbildungen 4a-e. Darstellung der Sensitivitäten und Spezifitäten in Abhängigkeit von verschiedenen Cut-Off-Werten für die Nomogramm-Wahrscheinlichkeiten. Die Abbildungen zeigen Unterschiede in der Diskrimination zwischen PCa und BPH für die fünf Nomogramme mit den fünf verschiedenen PSA-Testsystemen.

4.3.2 ROC-Analyse

Die ROC-Kurven-Analyse, die zur Beurteilung der Diskriminationsfähigkeit der Nomogramme herangezogen wird, zeigte vergleichbare Ergebnisse für die einzelnen PSA-Testsysteme innerhalb der Nomogramme (Tabelle 5).

Betrachtet man die Flächen unter den Kurven (AUCs) für tPSA, %fPSA und die Nomogramme I-V getrennt, so lassen sich kaum Unterschiede erkennen.

Die Analyse der AUCs für tPSA und %fPSA ergab für jeweils zwei Testsysteme (Abbott und Bayer sowie Beckmann und DPC) bessere Ergebnisse als für den Roche-Test. Beim %fPSA schnitt der Bayer-PSA-Test signifikant schlechter ab als die Testsysteme von Abbott, Beckmann und DPC.

Der Vergleich der verschiedenen PSA-Testsysteme innerhalb der einzelnen Nomogramme ergab für die Nomogramme III und V keine Unterschiede zwischen den Tests ($P = 0,09$ bis $0,87$); im Nomogramm II lieferten nur die PSA-Tests von Roche und Beckmann signifikant unterschiedliche AUCs ($P = 0,03$).

Das Nomogramm I präsentierte die niedrigste AUC-Variabilität mit einer Differenz von nur $0,01$ zwischen niedrigster und höchster AUC - trotzdem wurden drei signifikante Unterschiede gefunden ($P = 0,02$ bis $0,04$). Die größten Abweichungen zwischen den PSA-Testsystemen zeigte das Nomogramm IV mit fünf Signifikanzen ($P = 0,0001$ bis $0,03$).

Die ROC-Kurven-Analyse ergab für den Quotienten %fPSA signifikant höhere AUCs ($0,765$ bis $0,81$) als für den Parameter tPSA ($0,705$ bis $0,72$). Des Weiteren zeigte die ROC-Analyse eine deutliche Überlegenheit der Nomogramme II-V: sie erzielten signifikant höhere AUCs als das %fPSA (P immer $< 0,001$).

Eine Ausnahme bildete hier das Nomogramm I, das signifikant geringere AUCs als die anderen Nomogramme erreichte (P immer $< 0,001$) und bei dem kein Unterschied zum %fPSA bestand ($P = 0,30$ bis $0,79$).

Das Nomogramm II erzielte unterdessen die höchsten AUCs und erreichte mit dem Beckmann-Test das beste Gesamtergebnis in der ROC-Analyse (AUC = $0,87$).

Abbildung 5 zeigt den Verlauf der ROC-Kurven für alle fünf PSA-Tests innerhalb dieses Nomogramms.

Vergleicht man die AUCs der Nomogramme II-V mit den Daten jeweils eines PSA-Tests miteinander, so wurde innerhalb der DPC-Daten nur ein signifikanter Unterschied gefunden, innerhalb der Roche-Daten waren es zwei Signifikanzen.

Die Analyse der anderen Test-Daten ergab trotz ähnlicher AUCs im Bereich von 0,82 bis 0,87 drei bzw. vier signifikante Unterschiede.

Tabelle 6 enthält die Spezifitäten bei einer Sensitivität von 95% und die Sensitivitäten bei einer Spezifität von 95% für die fünf Nomogramme mit den fünf verschiedenen PSA-Testsystemen. Während die Spezifitäten innerhalb eines Nomogramms im Bereich von 4,3 bis 7,1% zwischen den Testsystemen differierten, variierten die Sensitivitäten zwischen 5 und 15,4%. Diese Unterschiede waren jedoch nicht statistisch signifikant.

Tabelle 5. Flächen unter den ROC-Kurven (AUCs) mit Standardabweichung (in Klammern) für tPSA, %fPSA und die fünf Nomogramme mit den fünf PSA-Tests.

PSA-Test	Nomogramm						
	tPSA	%fPSA [§]	I	II [#]	III	IV	V
Abbott	0,72 [§] (0,018)	0,80* (0,017)	0,79 [^] (0,016)	0,87 (0,012)	0,86 (0,013)	0,84" (0,014)	0,85 (0,013)
Bayer	0,715 [§] (0,018)	0,765 (0,018)	0,79 [^] (0,016)	0,86 (0,013)	0,85 (0,013)	0,82 (0,015)	0,85 (0,014)
Beckmann	0,71 (0,018)	0,81* ⁺ (0,016)	0,79 (0,016)	0,87 [°] (0,012)	0,86 (0,013)	0,85" [¤] (0,014)	0,85 (0,014)
DPC	0,71 (0,018)	0,80* ⁺ (0,017)	0,80 (0,016)	0,87 (0,013)	0,86 (0,013)	0,85" [¤] (0,013)	0,85 (0,013)
Roche	0,705 (0,018)	0,785 (0,017)	0,79 [^] (0,016)	0,86 (0,013)	0,86 (0,013)	0,84 (0,014)	0,84 (0,014)

[§] signifikant besser als Roche tPSA

[§] %fPSA bei Bayer-Test wie folgt berechnet: $100 - (c/tPSA \times 100)$

* signifikant besser als Bayer %fPSA

⁺ signifikant besser als Roche %fPSA

[#] das beste Ergebnis in der ROC-Analyse erzielte der Beckmann-Test mit dem Nomogramm II

[^] signifikant schlechter als DPC ($P < 0,04$)

[°] signifikant besser als Roche ($P = 0,03$)

" signifikant besser als Bayer ($P < 0,003$)

[¤] signifikant besser als Roche ($P < 0,03$)

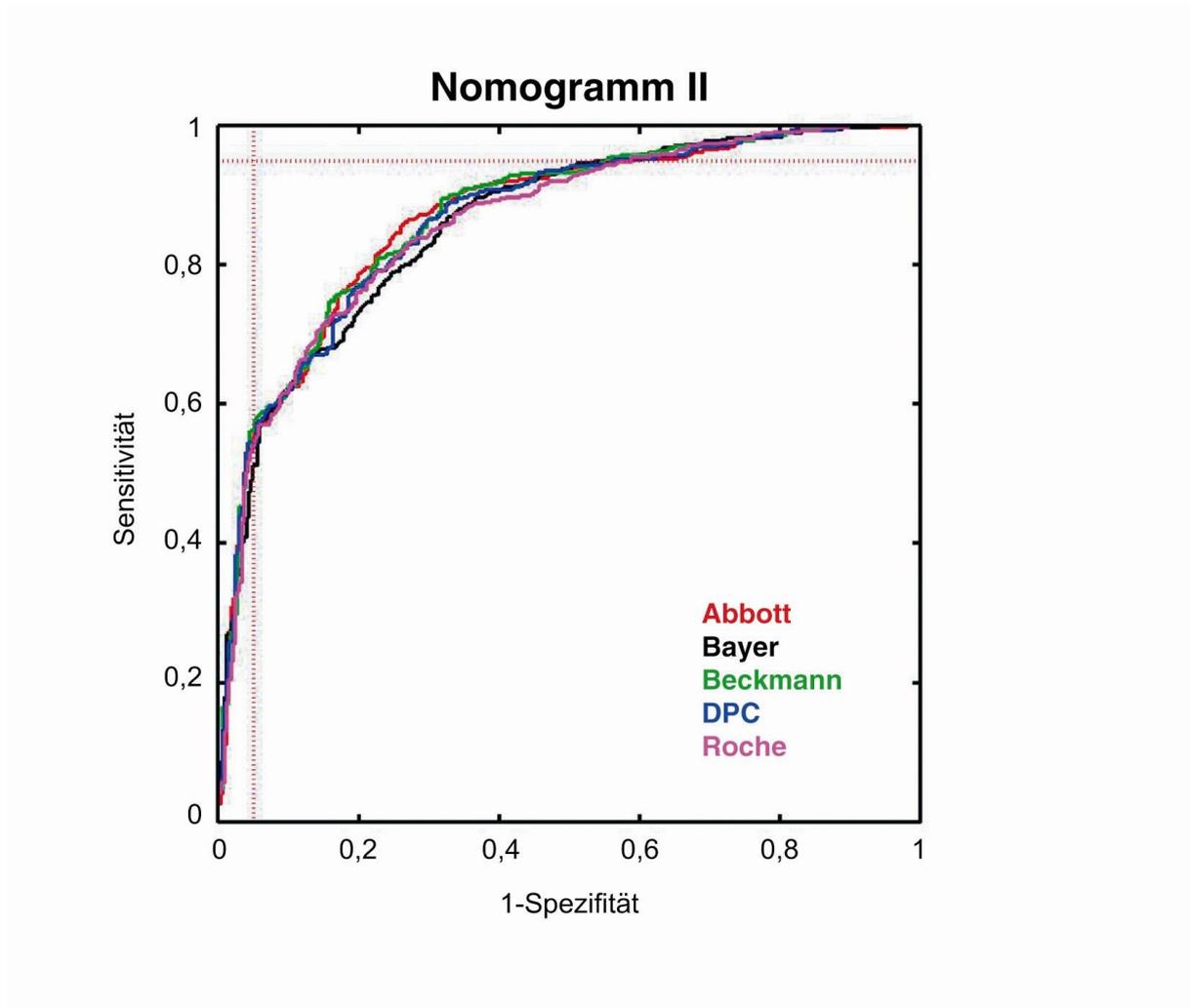


Abbildung 5. Der Verlauf der ROC-Kurven ist für alle fünf PSA-Tests innerhalb des Nomogramms II dargestellt. Das Nomogramm II erreichte die höchsten AUCs (Bereich: 0,86-0,87). Die eingezeichneten gestrichelten Linien markieren die 95%ige Sensitivität bzw. Spezifität.

Tabelle 6. Spezifitäten und Sensitivitäten mit Konfidenzintervallen für die fünf Nomogramme mit den fünf PSA-Tests.

Nomogramm und PSA-Test	Spezifität und Konfidenzintervall bei 95% Sensitivität	Sensitivität und Konfidenzintervall bei 95% Spezifität
Nomogramm I		
Abbott	18,8 (14,7-23,4)	33,0 (28,7-37,5)
Bayer	17,9 (13,8-22,5)	35,6 (31,2-40,2)
Beckmann	20,6 (16,3-25,4)	33,4 (29,1-37,9)
DPC	21,5 (17,2-26,4)	35,6 (31,2-40,2)
Roche	22,2 (17,8-27,1)	29,0 (24,9-33,4)
Nomogramm II		
Abbott	40,3 (34,9-45,9)	55,8 (51,1-60,4)
Bayer	44,6 (39,1-50,2)	51,2 (46,5-55,9)
Beckmann	44,9 (39,4-50,5)	58,0 (53,3-62,6)
DPC	44,3 (38,8-49,9)	57,4 (52,7-62,0)
Roche	41,9 (36,4-47,4)	55,4 (50,7-60,0)
Nomogramm III		
Abbott	42,8 (37,3-48,3)	46,6 (41,9-51,3)
Bayer	43,1 (37,6-48,7)	38,7 (34,2-43,3)
Beckmann	45,2 (39,7-50,8)	44,6 (40,0-49,3)
DPC	44,3 (38,8-49,9)	46,2 (41,5-50,9)
Roche	40,6 (35,2-46,2)	39,3 (34,8-44,0)
Nomogramm IV		
Abbott	39,4 (34,0-44,9)	34,1 (29,7-38,6)
Bayer	40,3 (34,9-45,9)	26,4 (22,4-30,7)
Beckmann	43,7 (38,2-49,3)	36,0 (31,6-40,6)
DPC	42,8 (37,3-48,3)	41,8 (37,2-46,4)
Roche	36,6 (31,4-42,1)	34,5 (30,1-39,1)
Nomogramm V		
Abbott	40,3 (34,9-45,9)	43,7 (39,1-48,4)
Bayer	41,5 (36,1-47,1)	43,1 (38,5-47,8)
Beckmann	42,8 (37,3-48,3)	43,1 (38,5-47,8)
DPC	43,1 (37,6-48,7)	40,2 (35,7-44,9)
Roche	37,5 (32,3-43,1)	38,7 (34,2-43,3)

4.3.3 Analyse bei 90% und 95% Sensitivität

In der Tabelle 7 sind die Nomogramm-Wahrscheinlichkeiten sowie die positiven und negativen prädiktiven Werte bei 90% und 95% Sensitivität dargestellt.

Vergleicht man die fünf Nomogramme bei Nutzung desselben PSA-Tests miteinander, fallen große Unterschiede in ihren Wahrscheinlichkeiten sowohl bei 90% als auch bei 95% Sensitivität auf.

Deutliche Unterschiede wurden auch beim Vergleich der fünf verschiedenen PSA-Testsysteme innerhalb der Nomogramme II bis V sichtbar. Bei einer Sensitivität von 90% wurden die maximalen Differenzen zwischen den Nomogramm-Wahrscheinlichkeiten für Abbott mit 0,39 und DPC mit 0,59 innerhalb des Nomogramms IV gefunden. Das Nomogramm I gab mit allen PSA-Tests sehr ähnliche Wahrscheinlichkeiten an.

Dagegen unterschieden sich die positiven und negativen prädiktiven Werte bei Nutzung verschiedener Testsysteme innerhalb eines Nomogramms nicht. Die PPV und NPV der Nomogramme II bis V lagen mit den Daten des gleichen PSA-Tests ebenfalls nah beieinander.

Tabelle 7. Nomogramm-Wahrscheinlichkeiten sowie positive und negative prädiktive Werte bei 90% und 95% Sensitivität.

PSA-Test und Charakteristika	90% Sensitivität Nomogramm					95% Sensitivität Nomogramm				
	I	II	III	IV	V	I	II	III	IV	V
Abbott										
Cut-Off	0,18	0,29	0,45	0,39	0,28	0,17	0,21	0,35	0,29	0,18
PPV	65,1	78,8	77,5	74,4	75	62,2	69,7	71,1	68,9	70,1
NPV	70	82,7	82,1	80,3	80,7	72,9	85,6	86,6	85	86
Bayer										
Cut-Off	0,18	0,3	0,45	0,4	0,27	0,17	0,21	0,36	0,31	0,2
PPV	64,6	76,8	76,9	74,8	74,3	61,8	71,1	70	69	69,6
NPV	69	81,7	81,8	80,6	80,3	71,6	86,6	85,9	85,1	85,5
Beckmann										
Cut-Off	0,19	0,38	0,52	0,5	0,33	0,17	0,29	0,43	0,41	0,24
PPV	64,9	79,2	77,2	74,7	74,1	62,7	70,9	70,8	70,2	69,9
NPV	69,6	82,8	81,9	80,5	80,2	74,7	86,5	86,5	86,1	85,8
DPC										
Cut-Off	0,2	0,43	0,57	0,59	0,4	0,18	0,35	0,49	0,5	0,3
PPV	67,9	77,8	76,5	74,4	75	62,9	70,5	70,5	69,9	70
NPV	74,4	82,2	81,6	80,3	80,7	75,3	86,2	86,2	85,8	85,9
Roche										
Cut-Off	0,19	0,37	0,53	0,52	0,36	0,18	0,31	0,44	0,43	0,24
PPV	65,4	74,1	75	73,9	75	63,1	69,6	69,1	68	68
NPV	70,6	80,2	80,7	80	80,7	75,8	85,5	85,2	84,1	84,1

PPV: positive predictive value (engl.) = positiver prädiktiver Wert

NPV: negative predictive value (engl.) = negativer prädiktiver Wert

alle Angaben in % (bis auf den Cut-Off-Wert, welcher ein Absolutwert ist)

5. Diskussion

Anhand einer Population von 798 urologischen Patienten, deren Datensätze verschiedene klinische Parameter und die mit fünf häufig verwendeten PSA-Testsystemen gemessenen tPSA- und fPSA-Werte umfassten, wurde in der vorliegenden Arbeit erstmals die Auswirkung der PSA-Test-Variabilität auf die Validität von Nomogrammen untersucht.

Für diese Evaluierung wurden fünf verschiedene verfügbare und anwendbare Nomogramme ausgesucht, die die Wahrscheinlichkeit für das Vorliegen eines PCa bei erstmaliger Prostatabiopsie angeben und welche von drei unterschiedlichen Arbeitsgruppen entwickelt wurden [116-118].

Nachdem das PSA 1994 in den USA in die PCa-Früherkennung eingeführt worden war, wird das PSA-Screening heute kontrovers diskutiert [135]. Es wird eingeschätzt, dass die PSA-Messung in bis zu 50% zu einer Überbehandlung von Tumoren führt, die möglicherweise weniger aggressiv oder gar nicht behandelt werden müssten [10]. Roobol et al. konnten in ihrer aktuellen Analyse der ERSPC allerdings zeigen, dass Männer, die sich dem PSA-Screening unterziehen, ein um 31% geringeres Mortalitätsrisiko besitzen als diejenigen, die auf die PSA-Bestimmung verzichten [11]. Die initiale Analyse ohne Berücksichtigung der „PSA-Kontamination“ und tatsächlichen Teilnahme am Screening von Schröder et al. legte bereits eine signifikante Reduktion der PCa-Mortalität um 20% offen [12].

Eine weitere Studie zeigte dagegen keine geringere Sterblichkeit durch das PSA-Screening [13]. Der hier beobachtete große Anteil von PSA-Tests in der Kontrollgruppe mit mindestens 52% kann möglicherweise zu diesem Ergebnis beigetragen haben [13].

Zur Optimierung der PCa-Früherkennung und zur Vermeidung von unnötigen Biopsien wurden in den letzten 10 Jahren ständig neue Nomogramme zur PCa-Vorhersage entwickelt. Auf die PSA-Test-Variabilität als möglichen Einflussfaktor wurde hierbei jedoch bisher nicht geachtet.

Die mitunter großen analytischen Differenzen zwischen den verschiedenen PSA-Testsystemen, die unterschiedliche tPSA- und fPSA-Messwerte und somit auch

unterschiedliche %fPSA-Werte zur Folge haben, stellen ein bekanntes Problem im klinischen Alltag dar [96, 126, 127].

Die Untersuchung von 157 Blutproben durch Slev et al. ergab, dass die relativen Unterschiede zwischen den PSA-Tests von Bayer, Beckmann und DPC bei einem PSA-Wert von 4 ng/ml mehr als 10% betrugten [127]. Stephan et al. zeigten an einer Kohorte von 596 Patienten, dass sich sowohl die tPSA-Werte stark unterschieden (Mediane 4,6 bis 5,7 µg/l) als auch die %fPSA-Mediane um bis zu 9% (Mediane 15% bis 24% bei PCa-Patienten) variierten [96].

Das Vorhandensein einer PSA-Test-Variabilität konnte in der vorliegenden Arbeit bestätigt werden (Tabelle 3). In der Kohorte aus 780 Patienten unterschieden sich nur die medianen tPSA-Werte, die mit den Testsystemen von Abbott und Bayer ermittelt wurden, nicht voneinander ($P = 0,86$). Zwischen allen anderen tPSA- und %fPSA-Werten existierten signifikante Unterschiede.

Diese Ergebnisse spiegelten sich in den unterschiedlichen medianen Nomogramm-Wahrscheinlichkeiten, die in Tabelle 4 dargestellt sind, wieder.

Verwendete man die tPSA-Werte von Abbott und Bayer als Eingangsvariablen im Nomogramm I, erhielt man die gleichen Wahrscheinlichkeiten. In den Nomogrammen II bis V lagen die Wahrscheinlichkeiten für Abbott und Bayer auch vergleichsweise nah beieinander.

Alle anderen Vergleiche der durch die Nomogramme vorhergesagten Wahrscheinlichkeiten führten zu statistisch signifikanten Unterschieden ($P < 0,0001$). Dabei war auffällig, dass die PSA-Tests von Beckmann, DPC und Roche im Nomogramm I nur zu gering unterschiedlichen Wahrscheinlichkeiten führten, obwohl mit ihnen stark differierende tPSA- und %fPSA-Werte gemessen worden waren. In den %fPSA-basierten Nomogrammen II bis V führten die unterschiedlichen PSA-Werte dieser drei Testsysteme dagegen zu stark variierenden Wahrscheinlichkeiten.

Diese Tatsache lässt vermuten, dass es insbesondere die unterschiedlichen %fPSA-Werte sind, die die große Variabilität in den Nomogramm-Wahrscheinlichkeiten bedingen. Nomogramm I besitzt nur die drei Eingangsvariablen Alter, DRU und tPSA. Die medianen vorhergesagten Wahrscheinlichkeiten lagen in diesem Modell für alle Testsysteme sehr nah beieinander (Bereich: 0,24-0,26). Die tPSA-Unterschiede scheinen also einen weitaus geringeren Einfluss auszuüben.

Unterstützt wird diese Hypothese durch die ähnlichen medianen %fPSA-Werte von Abbott und Bayer (16,2% und 16,4%), die zu ähnlichen PCa-Wahrscheinlichkeiten in den Nomogrammen II bis V führten.

Dieses Ergebnis nimmt Bezug auf die Studie von Kort et al. [126], in der die Serumproben von 70 Patienten hinsichtlich der Unterschiede zwischen den PSA-Testsystemen analysiert wurden. Dort wurde festgestellt, dass die Divergenz zwischen den fPSA-Messwerten größer als zwischen den tPSA-Werten war.

Tabelle 3 zeigt, dass sich die PSA-Werte von Bayer und DPC am deutlichsten unterschieden. Während mit dem Bayer-Test die niedrigsten tPSA-Werte gemessen wurden (Median: 5,5 µg/l), wies der DPC-Test den höchsten medianen tPSA-Wert auf (7,04 µg/l). Die mediane %fPSA-Ratio betrug beim Bayer-Test 16,4%, beim DPC-Test nur 10,6%, was einer Differenz von etwa 6% entsprach.

In verschiedenen Studien wurde gezeigt, dass das Risiko für das Vorliegen eines PCa mit steigendem tPSA-Wert und sinkender %fPSA-Ratio signifikant zunimmt [67, 71].

Während eine retrospektive Auswertung von PCa-Screening-Daten aus dem Jahr 2001 für die BPH-Kohorte mediane %fPSA-Werte von 19% ergab, hatten die PCa-Patienten mit einer Ratio von 9% signifikant niedrigere Werte [136].

Folglich würde man das Karzinomrisiko der hier untersuchten Kohorte von 798 Patienten höher einschätzen, wenn die tPSA- und fPSA-Werte mit dem DPC-Test gemessen wurden, während das PCa-Risiko der gleichen Patienten bei PSA-Messung mit dem Bayer-Test niedriger erscheint.

Diese große Differenz zwischen dem Bayer- und DPC-Test fand sich in den medianen Nomogramm-Wahrscheinlichkeiten wieder. Alle fünf Nomogramme berechneten die höchsten PCa-Wahrscheinlichkeiten bei Anwendung des DPC-Tests. Die niedrigsten Wahrscheinlichkeiten müssten bei direkter Korrelation zu den tPSA- und %fPSA-Werten bei Anwendung der Bayer-Daten in den Nomogrammen ermittelt werden.

Das Nomogramm I gab die niedrigsten Wahrscheinlichkeiten mit den Tests von Abbott und Bayer (kein signifikanter Unterschied) an. In den Nomogrammen II bis V waren die niedrigsten Wahrscheinlichkeiten dagegen ausnahmslos dem Abbott-Test zugehörig, auch wenn die Bayer-Wahrscheinlichkeiten relativ nah an denen von

Abbott lagen. Eine mögliche Ursache hierfür könnte sein, dass die medianen tPSA-Werte von Abbott und Bayer identisch waren und der mediane %fPSA-Wert von Abbott nur marginal kleiner war. Möglicherweise spielte auch die beim Bayer-Test erfolgte Berechnung des %fPSA aus cPSA eine Rolle (berechnet wie folgt: $100 - (c/tPSA \times 100)$). 18 der initialen 798 Patienten wiesen einen höheren cPSA- als tPSA-Wert auf. Hinzu kommt, dass der cPSA-Messwert bei vielen Patienten nur marginal kleiner war als der tPSA-Wert, was einer extrem niedrigen berechneten %fPSA-Ratio entsprach. Somit war das PCa-Risiko mit den tPSA- und cPSA-Daten des Bayer-Tests doch höher als mit dem Abbott-Testsystem.

Das Nomogramm IV wies die mit Abstand stärksten Abweichungen in den Nomogramm-Wahrscheinlichkeiten auf: Die Mediane reichten von 0,59 bis 0,76.

Die Variablen, die in diesem Nomogramm berücksichtigt werden, sind Alter, DRU, tPSA und %fPSA (Tabelle 2). Das %fPSA ist die einzige Variable, die verglichen mit dem Nomogramm I, welches sehr dicht beieinander liegende Wahrscheinlichkeiten präsentiert, hinzugekommen ist. Mit den gleichen Eingangsvariablen wie im Nomogramm IV führte das Nomogramm II zu etwas geringeren Unterschieden in den Medianen (0,42-0,54) als Nomogramm IV.

Diese Daten bestätigen die entscheidende Rolle des %fPSA in der Ausgabe der unterschiedlichen PSA-Test abhängigen Nomogramm-Wahrscheinlichkeiten.

Die Kalibrationskurven für die Nomogramme II bis V (Abbildungen 3b-e) zeigten ebenfalls deutliche Unterschiede zwischen den PSA-Testsystemen.

Nomogramm I dagegen wies auch in der Beziehung von festgestellten PCa-Raten bei der Biopsie und vorhergesagten PCa-Wahrscheinlichkeiten kaum Unterschiede zwischen den fünf Testsystemen auf (Abbildung 3a). Dafür präsentierte das Nomogramm, welches das %fPSA nicht berücksichtigt, die niedrigsten ICC-Werte (Bereich: 0,46-0,52).

Die Differenz von der tatsächlichen zu der vorhergesagten PCa-Wahrscheinlichkeit war bei diesem Nomogramm erheblich: Die beobachtete PCa-Rate (80%) war zum Teil doppelt so groß wie die vorausgesagte Wahrscheinlichkeit (40%). Es muss daher angenommen werden, dass mit diesem Nomogramm, je nach Definition des Cut-Off-Wertes, relativ viele Karzinompatienten übersehen würden.

Die Detektionsrate verbesserte sich deutlich, wenn das %fPSA als zusätzlicher Parameter in die Nomogramme eingeschlossen wurde (Abbildungen 3b-e). Die ICC-Werte befanden sich für die Nomogramme II bis V im Bereich von 0,84 bis 0,99.

Zu erwähnen ist, dass sowohl der minimale als auch der maximale ICC dem DPC-Test zugeordnet waren. Dieses zeigt, dass sich der DPC-Test, mit dem in allen Nomogrammen bereits die höchsten PCa-Wahrscheinlichkeiten berechnet wurden, auch in dieser Analyse von den anderen Testsystemen unterscheidet.

Mit steigenden ICC-Werten vergrößerte sich jedoch die Variabilität zwischen den PSA-Testsystemen, wie besonders deutlich im Nomogramm IV zu sehen war (Abbildung 3d). Ob eine bessere Kalibration (höhere ICC-Werte) mit größerer Nomogram-Variabilität im Zusammenhang steht, bleibt jedoch unklar.

Die größten Unterschiede fanden sich zwischen den Kalibrationskurven von DPC und Abbott bzw. Bayer. Auch in diesen Abbildungen wurde das Verhältnis der tPSA- und %fPSA-Werte der einzelnen PSA-Testsysteme reflektiert und somit gezeigt, dass ein direkter Zusammenhang zwischen den Nomogramm-Ergebnissen und den zugrunde liegenden PSA-Werten und damit PSA-Tests besteht.

Bei der Entwicklung von Nomogrammen sind die PSA-Testsysteme und deren möglicher Einfluss bisher nicht berücksichtigt worden. Die meisten Autoren gaben nicht einmal an, welcher PSA-Test für die Messung der PSA-Werte der Studienkohorte verwendet worden war [114-116]. Kawakami et al. [118] wechselten das Testsystem sogar zwischendurch, was eher dafür spricht, dass die Autoren keinen relevanten Zusammenhang zwischen tPSA- und fPSA-Testsystem und Nomogramm-Ergebnis vermuteten.

Interessanterweise war nur bei einem Nomogramm ein Zusammenhang zwischen dem offenbar besten PSA-Test für das Modell und dem PSA-Test, auf dessen Basis das Nomogramm entwickelt worden war, ersichtlich. Nomogramm III wurde anhand einer Patientenpopulation, deren tPSA- und fPSA-Werte mit dem Abbott-Test gemessen worden waren, entwickelt und erreichte den höchsten ICC (0,94) bei Nutzung der Abbott-Daten (Abbildung 3c).

Bei den Nomogrammen I und II wurde das für ihre Entwicklung verwendete PSA-Testsystem nicht angegeben, sodass sich für diese beiden Nomogramme keine Aussage treffen lässt.

Die PSA-Werte der Kohorte, an der die Nomogramme IV und V entwickelt wurden, sind zunächst mit dem PSA-Test von Beckmann, später mit dem von Abbott gemessen worden. Im Nomogramm IV wurde der höchste ICC allerdings bei Gebrauch der Bayer-Test-Daten erreicht, die zweithöchsten ICC-Werte waren dem Abbott- und Beckmann-Test zugeordnet. Im Nomogramm V zeigten die Abbott-Test-Daten dagegen die geringsten ICC-Werte von allen Testsystemen. Der DPC-Test, der nicht an der Entwicklung des Nomogramms beteiligt war, erreichte hier mit einem ICC von 0,99 das beste Ergebnis.

Diese Resultate zeigen, dass der Einfluss der zugrunde liegenden PSA-Testsysteme offenbar kleiner ist als angenommen.

Die Analyse bei verschiedenen Cut-Off-Werten für die Nomogramm-Wahrscheinlichkeiten zeigte, dass sich die Sensitivitäts- und Spezifitätsunterschiede zwischen den fünf Testsystemen bei Anwendung von %fPSA-basierten Nomogrammen deutlich vergrößerten (Abbildung 4a-e). Evident war, dass die tPSA-Unterschiede zwischen den PSA-Tests einen geringeren Einfluss auf die Höhe der Sensitivität und Spezifität ausübten als die %fPSA-Unterschiede.

Die Unterschiede bei den betrachteten Cut-Off-Werten von 0,25, 0,5 und 0,75 waren in den Nomogrammen II bis V zum Teil erheblich. Bei einem Cut-Off-Wert von 0,5 wurde innerhalb des Nomogramms IV ein Spezifitätsunterschied von 30% verzeichnet, wenn anstatt der DPC-Daten die Abbott-Daten angewandt wurden (Abbildung 4d). Wiederum bestanden die größten Unterschiede zwischen den Testsystemen, die die größten Differenzen in den tPSA- und %fPSA-Werten aufwiesen.

Betrachtet man generell den Verlauf der Kurven für die Spezifität und Sensitivität in den Nomogrammen II bis V (Abbildungen 4b-e), so sind bei jedem beliebigen Cut-Off-Wert die größten Unterschiede zwischen dem DPC-Test und den Tests von Abbott bzw. Bayer zu verzeichnen.

Die Sensitivitäts- und Spezifitätskurven, die den Tests von Beckmann und Roche zugeordnet sind, verliefen zwischen den Kurven der DPC- und Abbott/Bayer-Daten. Die in Tabelle 3 dargestellten medianen PSA-Werte, die mit dem Beckmann- und

Roche-Test gemessen wurden, lagen analog zu den Kurven nah beieinander und auch zwischen den Werten, die mit den Tests von DPC und Abbott/Bayer ermittelt wurden. Wiederum wurde ein Zusammenhang zwischen der PSA-Test-Variabilität und den stark unterschiedlichen, PSA-Test-abhängigen Sensitivitäten und Spezifitäten der Nomogramme offensichtlich.

Mittels ROC-Analyse konnten dagegen keine relevanten Unterschiede zwischen den fünf Testsystemen innerhalb der fünf Nomogramme gezeigt werden (Tabelle 5).

Die ROC-Analyse zeichnete die %fPSA-basierten Nomogramme II bis V, welche AUC-Werte zwischen 0,82 und 0,87 erreichten, als exzellente Modelle für die Vorhersage eines PCa aus. Eine isolierte Betrachtung dieser Ergebnisse sollte aufgrund der anderen deutlich divergenten Resultate und Darstellungen in dieser Arbeit nicht vorgenommen werden.

Bei der ROC-Kurven-Analyse handelt es sich um die bevorzugte Methode zur Beurteilung der Diskrimination. Sie wird am häufigsten zur Einschätzung eines neuen Biomarkers oder zur Evaluierung neuer Modelle, die das Risiko für eine bestimmte Erkrankung vorhersagen, eingesetzt [137, 138].

Mit den AUC-Berechnungen kann jedoch ausschließlich beurteilt werden, wie exakt ein Modell ein bestimmtes Ereignis vorhersagt. Die ROC-Analyse kann weder eine Aussage darüber treffen, welches betrachtete Modell eingesetzt oder bevorzugt werden sollte, noch können klinische Konsequenzen mit diesem Verfahren beurteilt werden [139].

Beispielsweise wäre ein neues Modell mit einem großen Spezifitätszuwachs, aber einer geringeren Sensitivität eine insuffiziente Wahl für den klinischen Gebrauch, es würde aber möglicherweise eine höhere AUC ergeben [139].

Utsumi et al. [131] verwendeten für die externe Validierung ihres Nomogramms und den Vergleich mit anderen Nomogrammen allein die ROC-Kurven-Analyse. Sie verglichen alle AUCs miteinander und formulierten daraus das Ergebnis [131].

Um zu beurteilen, ob Vorhersagemodelle akkurat arbeiten und um diese Modelle miteinander zu vergleichen, sollten aber sowohl die Kalibrierung (diese misst, wie gut die vorhergesagten Wahrscheinlichkeiten mit den tatsächlichen Erkrankungsraten

übereinstimmen) als auch die Diskrimination (diese ist ein Maß dafür, wie gut das Modell die „Kranken“ von den „Gesunden“ separiert) herangezogen werden [137].

Zur Beurteilung der Diskriminationsfähigkeit eines Modells stehen dabei nicht nur die populäre ROC-Kurven-Analyse zur Verfügung, sondern z.B. auch die Analyse bei verschiedenen Cut-Off-Werten sowie die Analyse bei 90% und 95% Sensitivität.

Die ROC-Analyse ergab signifikant niedrigere AUCs für die medianen tPSA-Werte verglichen mit den %fPSA-Werten (Tabelle 5). Damit konnte in dieser Studie bestätigt werden, was bereits zahlreiche andere Autoren gezeigt hatten: Mit dem Quotienten %fPSA kann das Risiko für das Vorliegen eines PCa exakter eingeschätzt werden als mit dem Parameter tPSA [50, 65, 66, 68, 75, 86].

Verschiedene Studien hatten gezeigt, dass mit Hilfe des %fPSA die Spezifität des tPSA sowohl im Bereich von 4-10 µg/l [65] als auch für niedrige PSA-Werte < 4 µg/l um 15-30% [76-78] gesteigert werden konnte.

Des Weiteren ergab die ROC-Analyse für die Nomogramme II bis V signifikant höhere AUCs gegenüber dem %fPSA. Dieses Ergebnis steht im Einklang mit den Resultaten einer aktuellen Übersichtsarbeit von Schröder und Kattan [109]. Dort wurden alle zur Verfügung stehenden ANN- und LR-Studien für die Detektion des PCa verglichen und fast immer deutlich höhere AUCs für diese Modelle verglichen mit dem tPSA-Wert nachgewiesen.

Nur das Nomogramm I erreichte keine Verbesserung der diagnostischen Aussagekraft des %fPSA. Dieses kann anhand der Eingangsvariablen des Nomogramms erklärt werden. Während das Nomogramm I lediglich Alter, DRU und tPSA beinhaltet, wurde in alle anderen Nomogramme das %fPSA als zusätzliche Größe eingeschlossen.

Die Analysen der Nomogramm-Wahrscheinlichkeiten bei 90% und 95% Sensitivität zeigten große Unterschiede in den Nomogrammen II bis V je nach gewähltem PSA-Testsystem. Im Nomogramm I ließen sich dagegen erneut keine Auswirkungen der PSA-Test-Variabilität erkennen (Tabelle 7). Die positiven und negativen prädiktiven

Werte zeigten keine PSA-Test-Abhängigkeit und waren bei den gewählten Cut-Off-Werten von 90% und 95% Sensitivität relativ stabil.

Die Ergebnisse der externen Validierung der fünf Nomogramme, wie sie diese Arbeit neben der Untersuchung des Einflusses der PSA-Test-Variabilität auf die Nomogramm-Vorhersage darstellt, können durch Unterschiede in den Patientenpopulationen beeinflusst worden sein.

Die hier untersuchte Kohorte aus 780 Männern wies mit 39% positiven Tastbefunden einen höheren prozentualen Anteil an suspekten DRU-Befunden auf als die Studienpopulationen, die der Nomogramm-Entwicklung dienten [116-118].

Während Kawakami et al. [118] und Chun et al. [117] nur bei 17,5% und 20,3% der Patienten eine positive DRU feststellten, wiesen die drei Kohorten, die Karakiewicz et al. [116] für die Entwicklung der Nomogramme I und II nutzten, in 37,2%, 16,7% und 30,4% suspekten Befunde auf.

Auch die PCa-Prävalenz lag in der Population der hier vorliegenden Studie mit 58,3% vergleichsweise hoch. In den Kohorten für die Entwicklung der fünf Nomogramme war eine Prävalenz von 35,2% bis 41,9% vorhanden.

Die Stärken dieser Arbeit liegen in der relativ großen Studienpopulation und der Tatsache, dass exakt die gleichen Patienten mit allen fünf PSA-Tests gemessen worden sind und so eine gute Vergleichbarkeit der Testsysteme innerhalb der Nomogramme gegeben ist. Vorteilhaft sind darüber hinaus die verschiedenen Methoden zur Kalibrierung und Diskrimination, die nebeneinander zur Anwendung kamen und so eine differenzierte Beantwortung der gegebenen Fragestellung ermöglichten.

Limitationen dieser Arbeit sind im retrospektiven Studiendesign und in der Untersuchung von nur fünf Nomogrammen zu sehen. Weitere Nomogramme für die Vorhersage eines PCa bei der Prostatabiopsie waren entweder nicht online erhältlich [140-142] oder waren anhand zu kleiner Studienpopulationen entwickelt worden [143]. Ein möglicher Ausgangspunkt für weiterführende Untersuchungen ist die Tatsache, dass nur ein Nomogramm eingeschlossen wurde, welches kein %fPSA enthält. Damit kann die Aussage, dass vor allem die %fPSA-Variabilität zu

Unterschieden in den Nomogramm-Wahrscheinlichkeiten führt bzw. dass tPSA-Unterschiede kaum einen Einfluss auf die Nomogramm-Vorhersage haben, nur anhand dieses einen Nomogramms getroffen werden.

In der vorliegenden Arbeit konnte die Annahme, dass die PSA-Test-Variabilität einen Einfluss auf die Vorhersage eines PCa mittels Nomogrammen ausübt, bestätigt werden. Diese Aussage ist aber auf %fPSA-Unterschiede und %fPSA-basierte Nomogramme beschränkt. Es konnte nicht nachgewiesen werden, dass Nomogramme besonders gut mit demjenigen PSA-Test funktionieren, auf dessen Basis sie entwickelt wurden.

Diese Resultate werfen die Frage auf, ob Nomogramme, die der Vorhersage des Biopsieergebnisses dienen, wirklich für alle Populationen und alle PSA-Testsysteme anwendbar sind. Momentan ist bei den frei erhältlichen Nomogrammen kein Kommentar zu finden, dass die ermittelten Nomogramm-Wahrscheinlichkeiten von den verwendeten PSA-Testsystemen abhängen.

Für die Nutzung von ANN-Modellen für die PCa-Detektion wurde bereits ein ANN entwickelt, das den jeweiligen tPSA- und fPSA-Test berücksichtigt [129].

Die Ergebnisse dieser Studie sollen die Aufmerksamkeit auf die Unterschiede zwischen den PSA-Testsystemen und deren Auswirkungen auf die Nomogramm-Wahrscheinlichkeiten lenken. Bei einem stetig ansteigenden Gebrauch von multivariaten Modellen darf bei routinemäßiger Validierung der Diskriminationsfähigkeit mittels ROC-Analyse die Notwendigkeit einer exakten Kalibration nicht außer Acht gelassen werden.

In weiteren externen Validierungen mit prospektivem Studiendesign und anderen Kohorten sollten sowohl Nomogramme als auch ANN-Modelle, die der PCa-Vorhersage dienen, hinsichtlich ihrer PSA-Test-Abhängigkeit untersucht werden.

Bei Nutzung anderer PCa-Nomogramme, die beispielsweise dem Staging dienen, den postoperativen Verlauf oder ein biochemisches Rezidiv voraussagen, ist ein derart starker Einfluss der PSA-Test-Variabilität eher weniger wahrscheinlich. Dennoch sollte auch dieser Aspekt in zukünftigen Studien untersucht werden.

6. Zusammenfassung

Die Prostatastanzbiopsie, deren Indikation mit Hilfe des PSA-Wertes und des DRU-Befundes gestellt wird, führt nur in etwa 25% zu einem positiven Ergebnis.

Nomogramme, mit denen die Wahrscheinlichkeit für das Vorliegen eines PCa bei der Biopsie ermittelt werden kann, stellen nichtinvasive Verfahren zur verbesserten PCa-Früherkennung und zur Reduzierung der unnötigen Biopsien dar.

Zu den Eingangsvariablen der Nomogramme gehören neben den klinischen Parametern die tPSA- und teilweise auch die %fPSA-Werte. Für die Messung dieser Parameter existieren eine Reihe verschiedener Testsysteme, zwischen denen eine deutliche Variabilität besteht.

In der vorliegenden Arbeit wurde der Einfluss der PSA-Test-Variabilität auf die Vorhersage eines PCa mittels Nomogrammen untersucht.

Für diese Evaluierung wurden die Daten von 798 Patienten, deren PSA-Werte mit fünf häufig verwendeten PSA-Testsystemen gemessen worden waren, mit fünf verschiedenen Nomogrammen analysiert. Zusätzlich erfolgte in dieser Arbeit die externe Validierung der fünf Nomogramme.

Zur Beurteilung der Diskriminationsfähigkeit der Nomogramme wurde die ROC-Analyse sowie die Analyse bei verschiedenen Cut-Off-Werten für die Nomogramm-Wahrscheinlichkeiten und bei 90% und 95% Sensitivität durchgeführt.

Die Kalibrierung der Nomogramme erfolgte durch den Vergleich der vorhergesagten PCa-Wahrscheinlichkeiten und der beobachteten PCa-Raten bei der Biopsie.

Die Auswertung der medianen Nomogramm-Wahrscheinlichkeiten zeigte erhebliche Unterschiede zwischen den PSA-Testsystemen: Das mediane PCa-Risiko schwankte innerhalb eines Nomogramms je nach verwendetem PSA-Test zwischen 0,59 und 0,76. Auch die Analyse bei verschiedenen Cut-Off-Werten ergab Spezifitätsunterschiede von 30% und mehr zwischen den Testsystemen.

Mittels ROC-Kurven-Analyse wurden vergleichbare AUCs für die verschiedenen PSA-Tests gefunden. Die Analyse bei 90% und 95% Sensitivität zeigte dagegen wieder deutliche Unterschiede in den Nomogramm-Wahrscheinlichkeiten.

In dieser Arbeit konnte zum ersten Mal gezeigt werden, dass die PSA-Test-Variabilität einen Einfluss auf die Vorhersage eines PCa mittels Nomogrammen ausübt. Dies gilt insbesondere für %fPSA-Unterschiede und %fPSA-basierte Nomogramme. Diese Ergebnisse sollten in weiteren, prospektiven Studien validiert werden.

7. Literaturverzeichnis

1. Batzler WU, Giersiepen K, Hentschel S. Cancer in Germany 2003-2004 - Incidence and Trends. Gesundheitsberichterstattung für Deutschland 2008;70-3.
2. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer Statistics, 2009. *CA Cancer J Clin* 2009;59:225-49.
3. Rodriguez C, Freedland SJ, Deka A, Jacobs EJ, McCullough ML, Patel AV, Thun MJ, Calle EE. Body mass index, weight change, and risk of prostate cancer in the Cancer Prevention Study II Nutrition Cohort. *Cancer Epidemiol Biomarkers Prev* 2007;16:63-9.
4. Thompson I, Tangen C, Tolcher A, Crawford E, Eisenberger M, Moinpour C. Association of African-American ethnic background with survival in men with metastatic prostate cancer. *J Natl Cancer Inst* 2001;93:219-25.
5. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, Lee C, Montie JE, Shah RB, Pienta KJ, Rubin MA, Chinnaiyan AM. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 2005;310:644-8.
6. Demichelis F, Fall K, Perner S, Andren O, Schmidt F, Setlur SR, Hoshida Y, Mosquera JM, Pawitan Y, Lee C, Adami HO, Mucci LA, Kantoff PW, Andersson SO, Chinnaiyan AM, Johansson JE, Rubin MA. TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. *Oncogene* 2007;26:4596-9.
7. Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X, Morris DS, Menon A, Jing X, Cao Q, Han B, Yu J, Wang L, Montie JE, Rubin MA, Pienta KJ, Roulston D, Shah RB, Varambally S, Mehra R, Chinnaiyan AM. Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature* 2007;448:595-9.
8. Bestmann B, Loetters C, Diemer T, Weidner W, Küchler T, Rohde V. Prostate-specific symptoms of prostate cancer in a German general population. *Prostate Cancer and Prostatic Diseases* 2007;10:52-9.
9. Han M, Partin AW, Piantadosi S, Epstein JI, Walsh PC. Era specific biochemical recurrence-free survival following radical prostatectomy for clinically localized prostate cancer. *J Urol* 2001;166:416-9.
10. Draisma G, Boer R, Otto SJ, van dC, I, Damhuis RA, Schroder FH, De Koning HJ. Lead times and overdiagnosis due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. *J Natl Cancer Inst* 2003;95:868-78.
11. Roobol MJ, Kerkhof M, Schröder FH. Prostate Cancer Mortality Reduction by Prostate-Specific Antigen-Based Screening Adjusted for Nonattendance and

- Contamination in the European Randomised Study of Screening for Prostate Cancer (ERSPC). Epub ahead of print. *Eur Urol* 2009.
12. Schröder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med* 2009;360:1320-8.
 13. Andriole GL, Grubb RL, Buys SS. Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med* 2009;360:1310-9.
 14. Catalona WJ, Smith DS, Ratliff TL, Dodds KM, Coplen DE, Yuan JJ, Petros JA, Andriole GL. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *N Engl J Med* 1991;324:1156-61.
 15. Stamey TA, Yang N, Hay AR, McNeal JE, Freiha FS, Redwine E. Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate. *N Engl J Med* 1987;317:909-16.
 16. Ablin RJ, Soanes WA, Bronson P, Witebsky E. Precipitating antigens of the normal human prostate. *J Reprod Fertil* 1970;22:573-4.
 17. Hara M, Koyanagi Y, Inoue T, Fukuyama T. Some physico-chemical characteristics of " -seminoprotein", an antigenic component specific for human seminal plasma. Forensic immunological study of body fluids and secretion. *Nippon Hoigaku Zasshi* 1971;25:322-4.
 18. Li TS, Beling CG. Isolation and characterization of two specific antigens of human seminal plasma. *Fertil Steril* 1973;24:134-44.
 19. Sensabaugh GF. Isolation and characterization of a semen-specific protein from human seminal plasma: a potential new marker for semen identification. *J Forensic Sci* 1978;23:106-15.
 20. Wang MC, Valenzuela LA, Murphy GP, Chu TM. Purification of a human prostate specific antigen. *Invest Urol* 1979;17:159-63.
 21. Papsidero LD, Wang MC, Valenzuela LA, Murphy GP, Chu TM. A prostate antigen in sera of prostatic cancer patients. *Cancer Res* 1980;40:2428-32.
 22. Wang MC, Papsidero LD, Kuriyama M, Valenzuela LA, Murphy GP, Chu TM. Prostate antigen: a new potential marker for prostatic cancer. *Prostate* 1981;2:89-96.
 23. Lundwall A, Lilja H. Molecular cloning of human prostate specific antigen cDNA. *FEBS Lett* 1987;214:317-22.
 24. Riegman PH, Vlietstra RJ, van der Korput JA, Romijn JC, Trapman J. Characterization of the prostate-specific antigen gene: a novel human kallikrein-like gene. *Biochem Biophys Res Commun* 1989;159:95-102.

25. Yousef GM, Diamandis EP. The new human tissue kallikrein gene family: structure, function, and association to disease. *Endocr Rev* 2001;22:184-204.
26. Kumar A, Mikolajczyk SD, Goel AS, Millar LS, Saedi MS. Expression of pro form of prostate-specific antigen by mammalian cells and its conversion to mature, active form by human kallikrein 2. *Cancer Res* 1997;57:3111-4.
27. Watt KW, Lee PJ, M'Timkulu T, Chan WP, Loo R. Human prostate-specific antigen: structural and functional similarity with serine proteases. *Proc Natl Acad Sci U S A* 1986;83:3166-70.
28. Ward AM, Catto JW, Hamdy FC. Prostate specific antigen: biology, biochemistry and available commercial assays. *Ann Clin Biochem* 2001;38:633-51.
29. Lilja H, Oldbring J, Rannevik G, Laurell CB. Seminal vesicle-secreted proteins and their reactions during gelation and liquefaction of human semen. *J Clin Invest* 1987;80:281-5.
30. Luboldt HJ. Leitlinie zur PSA-Bestimmung in der Prostatakarzinomdiagnostik (Früherkennung des Prostatakarzinoms). *Urologe A* 2002;41:509-13.
31. Stenman UH. Prostate-specific antigen, clinical use and staging: an overview. *Br J Urol* 1997;79 Suppl 1:53-60.
32. Stephan C, Jung K, Lein M, Diamandis EP. PSA and other tissue kallikreins for prostate cancer detection. *Eur J Cancer* 2007;43:1918-26.
33. Benson MC, Whang IS, Pantuck A, Ring K, Kaplan SA, Olsson CA, Cooner WH. Prostate specific antigen density: a means of distinguishing benign prostatic hypertrophy and prostate cancer. *J Urol* 1992;147:815-6.
34. Oesterling JE. Prostate specific antigen: a critical assessment of the most useful tumor marker for adenocarcinoma of the prostate. *J Urol* 1991;145:907-23.
35. Hankey BF, Feuer EJ, Clegg LX, Hayes RB, Legler JM, Prorok PC. Cancer surveillance series: interpreting trends in prostate cancer- part I: Evidence of the effects of screening in recent prostate cancer incidence, mortality, and survival rates. *J Natl Cancer Inst* 1999;91:1017-24.
36. Miller DC, Hafez KS, Stewart A, Montie JE, Wei JT. Prostate carcinoma presentation, diagnosis, and staging: an update from the National Cancer Data Base. *Cancer* 2003;98:1169-78.
37. Catalona WJ, Richie JP, Ahmann FR, Hudson MA, Scardino PT, Flanigan RC, deKernion JB, Ratliff TL, Kavoussi LR, Dalkin BL. Comparison of digital rectal examination and serum prostate specific antigen in the early detection of prostate cancer: results of a multicenter clinical trial of 6,630 men. *J Urol* 1994;151:1283-90.

38. Brawer MK, Chetner MP, Beatie J, Buchner DM, Vessella RL, Lange PH. Screening for prostatic carcinoma with prostate specific antigen. *J Urol* 1992;147:841-5.
39. Thompson IM, Pauler DK, Goodman PJ, Tangen CM, Lucia MS, Parnes HL, Minasian LM, Ford LG, Lippman SM, Crawford ED, Crowley JJ, Coltman CA, Jr. Prevalence of prostate cancer among men with a prostate-specific antigen level ≤ 4.0 ng per milliliter. *N Engl J Med* 2004;350:2239-46.
40. Schröder FH, van der Crujisen-Koeter I, De Koning HJ, Vis AN, Hoedemaker RF, Kranse R. Prostate cancer detection at low prostate specific antigen. *J Urol* 2000;163:806-12.
41. Christensson A, Björk T, Nilsson O, Dahlen U, Matikainen MT, Cockett AT, Abrahamsson PA, Lilja H. Serum prostate specific antigen complexed to alpha 1-antichymotrypsin as an indicator of prostate cancer. *J Urol* 1993;150:100-5.
42. Lilja H, Christensson A, Dahlen U, Matikainen MT, Nilsson O, Pettersson K, Lovgren T. Prostate-specific antigen in serum occurs predominantly in complex with alpha 1-antichymotrypsin. *Clin Chem* 1991;37:1618-25.
43. Stenman UH, Leinonen J, Alfthan H, Rannikko S, Tuhkanen K, Alfthan O. A complex between prostate-specific antigen and alpha 1-antichymotrypsin is the major form of prostate-specific antigen in serum of patients with prostatic cancer: assay of the complex improves clinical sensitivity for cancer. *Cancer Res* 1991;51:222-6.
44. Allard WJ, Zhou Z, Yeung KK. Novel immunoassay for the measurement of complexed prostate-specific antigen in serum. *Clin Chem* 1998;44:1216-23.
45. Chen Z, Komatsu K, Prestigiacomo A, Stamey TA. Addition of purified prostate specific antigen to serum from female subjects: studies on the relative inhibition by alpha 2-macroglobulin and alpha 1-antichymotrypsin. *J Urol* 1996;156:1357-63.
46. Leinonen J, Zhang WM, Stenman UH. Complex formation between PSA isoenzymes and protease inhibitors. *J Urol* 1996;155:1099-103.
47. Zhang WM, Finne P, Leinonen J, Vesalainen S, Nordling S, Rannikko S, Stenman UH. Characterization and immunological determination of the complex between prostate-specific antigen and alpha2-macroglobulin. *Clin Chem* 1998;44:2471-9.
48. Zhang WM, Finne P, Leinonen J, Salo J, Stenman UH. Determination of prostate-specific antigen complexed to alpha(2)-macroglobulin in serum increases the specificity of free to total PSA for prostate cancer. *Urology* 2000;56:267-72.
49. Christensson A, Lilja H. Complex formation between protein C inhibitor and prostate-specific antigen in vitro and in human semen. *Eur J Biochem* 1994;220:45-53.

50. Polascik TJ, Oesterling JE, Partin AW. Prostate specific antigen: a decade of discovery-what we have learned and where we are going. *J Urol* 1999;162:293-306.
51. Lein M, Kwiatkowski M, Semjonow A, Luboldt HJ, Hammerer P, Stephan C, Klevecka V, Taymoorian K, Schnorr D, Recker F, Loening SA, Jung K. A multicenter clinical trial on the use of complexed prostate specific antigen in low prostate specific antigen concentrations. *J Urol* 2003;170:1175-9.
52. Stephan C, Schnorr D, Loening SA, Jung K. Re: 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. Roddam AW, Duffy MJ, Hamdy FC, et al. *Eur Urol* 2005;48:386-99. *Eur Urol* 2005;48:1059-60.
53. Mikolajczyk SD, Marks LS, Partin AW, Rittenhouse HG. Free prostate-specific antigen in serum is becoming more complex. *Urology* 2002;59:797-802.
54. Mikolajczyk SD, Grauer LS, Millar LS, Hill TM, Kumar A, Rittenhouse HG, Wolfert RL, Saedi MS. A precursor form of PSA (pPSA) is a component of the free PSA in prostate cancer serum. *Urology* 1997;50:710-4.
55. Peter J, Unverzagt C, Krogh TN, Vorm O, Hoesel W. Identification of precursor forms of free prostate-specific antigen in serum of prostate cancer patients by immunosorption and mass spectrometry. *Cancer Res* 2001;61:957-62.
56. Mikolajczyk SD, Marker KM, Millar LS, Kumar A, Saedi MS, Payne JK, Evans CL, Gasior CL, Linton HJ, Carpenter P, Rittenhouse HG. A truncated precursor form of prostate-specific antigen is a more specific serum marker of prostate cancer. *Cancer Res* 2001;61:6958-63.
57. Sokoll LJ, Wang Y, Feng Z, Kagan J, Partin AW, Sanda MG, Thompson IM, Chan DW. [-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.
58. Stephan C, Kahrs AM, Cammann H, Lein M, Schrader M, Deger S, Miller K, Jung K. A [-2]proPSA-based artificial neural network significantly improves differentiation between prostate cancer and benign prostatic diseases. *Prostate* 2009;69:198-207.
59. Mikolajczyk SD, Millar LS, Wang TJ, Rittenhouse HG, Wolfert RL, Marks LS, Song W, Wheeler TM, Slawin KM. "BPSA," a specific molecular form of free prostate-specific antigen, is found predominantly in the transition zone of patients with nodular benign prostatic hyperplasia. *Urology* 2000;55:41-5.
60. Canto EI, Singh H, Shariat SF, Lamb DJ, Mikolajczyk SD, Linton HJ, Rittenhouse HG, Kadmon D, Miles BJ, Slawin KM. Serum BPSA outperforms both total PSA and free PSA as a predictor of prostatic enlargement in men without prostate cancer. *Urology* 2004;63:905-10.

61. Steuber T, Nurmikko P, Haese A, Pettersson K, Graefen M, Hammerer P, Huland H, Lilja H. Discrimination of benign from malignant prostatic disease by selective measurements of single chain, intact free prostate specific antigen. *J Urol* 2002;168:1917-22.
62. Haese A, Noldus J, Steuber T, Huland H, Lilja H. [Isoforms of free prostate-specific antigen] Isoformen des freien prostataspezifischen Antigens. *Urologe A* 2004;43:675-9.
63. Bangma CH, Rietbergen JB, Kranse R, Blijenberg BG, Petterson K, Schroder FH. The free-to-total prostate specific antigen ratio improves the specificity of prostate specific antigen in screening for prostate cancer in the general population. *J Urol* 1997;157:2191-6.
64. Catalona WJ, Smith DS, Ornstein DK. Prostate cancer detection in men with serum PSA concentrations of 2.6 to 4.0 ng/mL and benign prostate examination. Enhancement of specificity with free PSA measurements. *JAMA* 1997;277:1452-5.
65. Catalona WJ, Partin AW, Slawin KM, Brawer MK, Flanigan RC, Patel A, Richie JP, deKernion JB, Walsh PC, Scardino PT, Lange PH, Subong EN, Parson RE, Gasior GH, Loveland KG, Southwick PC. Use of the percentage of free prostate-specific antigen to enhance differentiation of prostate cancer from benign prostatic disease: a prospective multicenter clinical trial. *JAMA* 1998;279:1542-7.
66. Jung K, Stephan C, Lein M, Henke W, Schnorr D, Brux B, Schürenkämper P, Loening SA. Analytical performance and clinical validity of two free prostate-specific antigen assays compared. *Clin Chem* 1996;42:1026-33.
67. Partin AW, Catalona WJ, Southwick PC, Subong EN, Gasior GH, Chan DW. Analysis of percent free prostate-specific antigen (PSA) for prostate cancer detection: influence of total PSA, prostate volume, and age. *Urology* 1996;48:55-61.
68. Van Cangh PJ, De Nayer P, De Vischer L, Sauvage P, Tombal B, Lorge F, Wese FX, Opsomer R. Free to total prostate-specific antigen (PSA) ratio improves the discrimination between prostate cancer and benign prostatic hyperplasia (BPH) in the diagnostic gray zone of 1.8 to 10 ng/mL total PSA. *Urology* 1996;48:67-70.
69. Stephan C, Jung K, Lein M, Sinha P, Schnorr D, Loening SA. Molecular forms of prostate-specific antigen and human kallikrein 2 as promising tools for early diagnosis of prostate cancer. *Cancer Epidemiol Biomarkers Prev* 2000;9:1133-47.
70. Stephan C, Jung K, Diamandis EP, Rittenhouse HG, Lein M, Loening SA. Prostate specific antigen, its molecular forms, and other kallikrein markers for detection of prostate cancer. *Urology* 2002;59:2-8.

71. Catalona WJ, Smith DS, Wolfert RL, Wang TJ, Rittenhouse HG, Ratliff TL, Nadler RB. Evaluation of percentage of free serum prostate-specific antigen to improve specificity of prostate cancer screening. *JAMA* 1995;274:1214-20.
72. Stenman UH, Leinonen J, Zhang WM, Finne P. Prostate-specific antigen. *Semin Cancer Biol* 1999;9:83-93.
73. Lein M, Stephan C, Jung K, Schnorr D, Loening SA. [Molecular forms of prostate-specific antigen and human kallikrein 2 as possible indicators in prostatic carcinoma diagnosis]. *Urologe A* 2000;39:313-23.
74. Catalona WJ. Clinical utility of measurements of free and total prostate-specific antigen (PSA): a review. *Prostate Suppl* 1996;7:64-9.
75. Woodrum DL, Brawer MK, Partin AW, Catalona WJ, Southwick PC. Interpretation of free prostate specific antigen clinical research studies for the detection of prostate cancer. *J Urol* 1998;159:5-12.
76. Catalona WJ, Partin AW, Finlay JA, Chan DW, Rittenhouse HG, Wolfert RL, Woodrum DL. Use of percentage of free prostate-specific antigen to identify men at high risk of prostate cancer when PSA levels are 2.51 to 4 ng/mL and digital rectal examination is not suspicious for prostate cancer: an alternative model. *Urology* 1999;54:220-4.
77. Jung K, Stephan C, Elgeti U, Lein M, Brux B, Kristiansen G, Rudolph B, Hauptmann S, Schnorr D, Loening SA. Molecular forms of prostate-specific antigen in serum with concentrations of total prostate-specific antigen <4 µg/l - are they useful tools for early detection and screening of prostate cancer? *Int J Cancer* 2001;93:759-65.
78. Roehl KA, Antenor JA, Catalona WJ. Robustness of free prostate specific antigen measurements to reduce unnecessary biopsies in the 2.6 to 4.0 ng./ml. range. *J Urol* 2002;168:922-5.
79. Anonymous. Sachstandserhebung "Krebs-Früherkennung" der Deutschen Krebshilfe. www.krebshilfe.de 2007.
80. Catalona WJ, Smith DS, Ratliff TL, Basler JW. Detection of organ-confined prostate cancer is increased through prostate-specific antigen-based screening. *JAMA* 1993;270:948-54.
81. Vis AN, Hoedemaeker RF, Roobol M, Der Kwast TH, Schroder FH. Tumor characteristics in screening for prostate cancer with and without rectal examination as an initial screening test at low PSA (0.0-3.9 ng/ml). *Prostate* 2001;47:252-61.
82. Song JM, Kim CB, Chung HC, Kane RL. Prostate specific antigen, digital rectal examination and transrectal ultrasonography: a meta-analysis for this diagnostic triad of prostate cancer in symptomatic korean men. *Yonsei Med J* 2005;46:414-24.

83. Cooner WH, Mosley BR, Rutherford CL, Jr., Beard JH, Pond HS, Terry WJ, Igel TC, Kidd DD. Prostate cancer detection in a clinical urological practice by ultrasonography, digital rectal examination and prostate specific antigen. *J Urol* 1990;143:1146-52.
84. Sedelaar JP, De La Rosette JJ, Beerlage HP, Wijkstra H, Debruyne FM, Aarnink RG. Transrectal ultrasound imaging of the prostate: review and perspectives of recent developments. *Prostate Cancer and Prostatic Diseases* 1999;2:241-52.
85. Stephan C, Jung K, Lein M, Schnorr D, Loening SA. The ratio of free to total prostate specific antigen in serum is correlated to the prostate volume. *Int J Cancer* 1996;67:461-2.
86. Stephan C, Lein M, Jung K, Schnorr D, Loening SA. The influence of prostate volume on the ratio of free to total prostate specific antigen in serum of patients with prostate carcinoma and benign prostate hyperplasia. *Cancer* 1997;79:104-9.
87. Oesterling JE, Jacobsen SJ, Chute CG. Serum prostate-specific antigen in a community-based population of healthy men: establishment of age-specific reference ranges. *JAMA* 1993;270:860-6.
88. McConnell JD, Barry MJ, Bruskewitz RC. Benign prostatic hyperplasia: diagnosis and treatment. Agency for Health Care Policy and Research. *Clin Pract Guidel Quick Ref Guide Clin* 1994;8:1-17.
89. Carter HB, Pearson JD, Metter EJ, Brant LJ, Chan DW, Andres R, Fozard JL, Walsh PC. Longitudinal evaluation of prostate-specific antigen levels in men with and without prostate disease. *JAMA* 1992;267:2215-20.
90. Berger AP, Deibl M, Strasak A, Bektic J, Pelzer AE, Klocker H, Steiner H, Fritsche G, Bartsch G, Horninger W. Large-scale study of clinical impact of PSA velocity: long-term PSA kinetics as method of differentiating men with from those without prostate cancer. *Urology* 2007;69:134-8.
91. Loeb S, Roehl KA, Catalona WJ, Nadler RB. Prostate specific antigen velocity threshold for predicting prostate cancer in young men. *J Urol* 2007;177:899-902.
92. Schroder FH, Roobol MJ, van der Kwast TH, Kranse R, Bangma CH. Does PSA velocity predict prostate cancer in pre-screened populations? *Eur Urol* 2006;49:460-5.
93. Thompson IM, Ankerst DP, Chi C, Goodman PJ, Tangen CM, Lucia MS, Feng Z, Parnes HL, Coltman CA, Jr. Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. *J Natl Cancer Inst* 2006;98:529-34.
94. Komatsu K, Wehner N, Prestigiacomo AF, Chen Z, Stamey TA. Physiologic (intraindividual) variation of serum prostate-specific antigen in 814 men from a screening population. *Urology* 1996;47:343-6.

95. Nixon RG, Wener MH, Smith KM, Parson RE, Strobel SA, Brawer MK. Biological variation of prostate specific antigen levels in serum: an evaluation of day-to-day physiological fluctuations in a well-defined cohort of 24 patients. *J Urol* 1997;157:2183-90.
96. Stephan C, Klaas M, Muller C, Schnorr D, Loening SA, Jung K. Interchangeability of measurements of total and free prostate-specific antigen in serum with 5 frequently used assay combinations: an update. *Clin Chem* 2006;52:59-64.
97. Stephan C, Stroebel G, Heinau M, Lenz A, Roemer A, Lein M, Schnorr D, Loening SA, Jung K. The ratio of prostate-specific antigen (PSA) to prostate volume (PSA density) as a parameter to improve the detection of prostate carcinoma in PSA values in the range of < 4 ng/mL. *Cancer* 2005;104:993-1003.
98. Thon WF, Gadban F, Truss MC, Kuczyk M, Hartmann U, Jonas U. Prostate-specific antigen density- a reliable parameter for the detection of prostate cancer? *World J Urol* 1996;14:53-8.
99. Stamey TA, Kabalin JN, McNeal JE, Johnstone IM, Freiha F, Redwine EA, Yang N. Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate. II. Radical prostatectomy treated patients. *J Urol* 1989;141:1076-83.
100. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev* 1958;65:386-408.
101. Dybowski R, Gant V. Artificial neural networks in pathology and medical laboratories. *Lancet* 1995;346:1203-7.
102. 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.
103. Babaian RJ, Fritsche H, Ayala A, Bhadkamkar V, Johnston DA, Naccarato W, Zhang Z. 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.
104. Djavan B, Remzi M, Zlotta A, Seitz C, Snow P, Marberger M. Novel artificial neural network for early detection of prostate cancer. *J Clin Oncol* 2002;20:921-9.
105. Finne P, Finne R, Auvinen A, Juusela H, Aro J, Maattanen L, Hakama M, Rannikko S, Tammela TL, Stenman U. Predicting the outcome of prostate biopsy in screen-positive men by a multilayer perceptron network. *Urology* 2000;56:418-22.
106. Kalra P, Togami J, Bansal BSG, Partin AW, Brawer MK, Babaian RJ, Ross LS, Niederberger CS. A neurocomputational model for prostate carcinoma detection. *Cancer* 2003;98:1849-54.

107. Stephan C, Cammann H, Semjonow A, Diamandis EP, Wymenga LFA, Lein M, Sinha P, Loening SA, Jung K. Multicenter evaluation of an artificial neural network to increase prostate cancer detection rate and reduce unnecessary biopsies. *Clin Chem* 2002;48:1279-87.
108. Stephan C, Cammann H, Meyer HA, Lein M, Jung K. PSA and new biomarkers within multivariate models to improve early detection of prostate cancer. *Cancer Lett* 2007;249:18-29.
109. Schroder F, Kattan MW. The Comparability of Models for Predicting the Risk of a Positive Prostate Biopsy with Prostate-Specific Antigen Alone: A Systematic Review. *Eur Urol* 2008;54:274-90.
110. Stephan C, Jung K, Cammann H, Vogel B, Brux B, Kristiansen G, Rudolph B, Hauptmann S, Lein M, Schnorr D, Sinha P, Loening SA. An artificial neural network considerably improves the diagnostic power of percent free prostate-specific antigen in prostate cancer diagnosis: Results of a 5-year investigation. *Int J Cancer* 2002;99:466-73.
111. Chun FK, Karakiewicz PI, Huland H, Graefen M. Role of nomograms for prostate cancer in 2007. *World J Urol* 2007;25:131-42.
112. Shariat SF, Capitanio U, Jeldres C, Karakiewicz PI. Can nomograms be superior to other predictive tools? *BJU Int* 2009;103:492-5.
113. Shariat S, Karakiewicz PI, Margulis V, Kattan MW. Inventory of prostate cancer predictive tools. *Curr Opin Urol* 2008;18:279-96.
114. Eastham JA, May R, Robertson JL. 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;45:709-13.
115. Garzotto M, Hudson RG, Peters L, Hsieh YC, Barrera E, Mori M, Beer TM, Klein T. Predictive modeling for the presence of prostate carcinoma using clinical, laboratory, and ultrasound parameters in patients with prostate specific antigen levels ≤ 10 ng/mL. *Cancer* 2003;98:1417-22.
116. Karakiewicz PI, Benayoun S, Kattan MW, Perrotte P, Valiquette L, Scardino PT, Cagiannos I, Heinzer H, Tanguay S, Aprikian AG, Huland H, Graefen M. 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.
117. Chun FK, Briganti A, Graefen M. Development and external validation of an extended 10-core biopsy nomogram. *Eur Urol* 2007;52:436-44.
118. Kawakami S, Numao N, Okubo Y, Koga F, Yamamoto S, Saito K, Fujii Y, Yonese J, Masuda H, Kihara K, Fukui I. Development, validation, and head-to-head comparison of logistic regression-based nomograms and artificial neural

- network models predicting prostate cancer on initial extended biopsy. *Eur Urol* 2008;54:601-11.
119. Chun FK, Graefen M, Briganti A, Gallina A, Hopp J, Kattan MW, Huland H, Karakiewicz PI. Initial biopsy outcome prediction--head-to-head comparison of a logistic regression-based nomogram versus artificial neural network. *Eur Urol* 2007;51:1236-40.
 120. Finne P, Finne R, Bangma C, Hugosson J, Hakama M, Auvinen A, Stenman UH. Algorithms based on prostate-specific antigen (PSA), free PSA, digital rectal examination and prostate volume reduce false-positive PSA results in prostate cancer screening. *Int J Cancer* 2004;111:310-5.
 121. Kattan MW. Comparison of Cox regression with other methods for determining prediction models and nomograms. *J Urol* 2003;170:6-9.
 122. Terrin N, Schmid CH, Griffith JL, D'Agostino RB, Selker HP. External validity of predictive models: a comparison of logistic regression, classification trees, and neural networks. *J Clin Epidemiol* 2003;56:721-9.
 123. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. *J Biomed Inform* 2002;35:352-9.
 124. Stephan C, Meyer HA, Cammann H, Lein M, Loening SA, Jung K. Re: Felix K.-H. Chun, Markus Graefen, Alberto Briganti, Andrea Gallina, Julia Hopp, Michael W. Kattan, Hartwig Huland and Pierre I. Karakiewicz. Initial biopsy outcome prediction--head-to-head comparison of a logistic regression-based nomogram versus artificial neural network. *Eur Urol* 2007;51:1236-43. *Eur Urol* 2007;51:1446-7.
 125. Sargent DJ. Comparison of artificial neural networks with other statistical approaches. *Cancer* 2001;91:1636-42.
 126. Kort SA, Martens F, Vanpoucke H, van Duijnhoven HL, Blankenstein MA. Comparison of 6 automated assays for total and free prostate-specific antigen with special reference to their reactivity toward the WHO 96/670 reference preparation. *Clin Chem* 2006;52:1568-74.
 127. Slev PR, La'ulu SL, Roberts WL. Intermethod differences in results for total PSA, free PSA, and percentage of free PSA. *Am J Clin Pathol* 2008;129:952-8.
 128. Stephan C, Xu C, Cammann H, Graefen M, Haese A, Huland H, Semjonow A, Diamandis EP, Remzi M, Djavan B, Wildhagen MF, Blijenberg BG, Finne P, Stenman UH, Jung K, Meyer HA. Assay-specific artificial neural networks for five different PSA assays and populations with PSA 2-10 ng/ml in 4,480 men. *World J Urol* 2007;25:95-103.
 129. Stephan C, Cammann H, Meyer HA, Muller C, Deger S, Lein M, Jung K. An artificial neural network for five different assay systems of prostate-specific antigen in prostate cancer diagnostics. *BJU Int* 2008;102:799-805.

130. Chun FK, Karakiewicz PI, Briganti A. Significance of PSA inter-assay variability on clinical prostate cancer detection-aids. [Abstract]. *J Urol* 2008;179:722.
131. Utsumi T, Kawamura K, Suzuki H. External validation and head-to-head comparison of Japanese and Western prostate biopsy nomograms using Japanese data sets. *Int J Urol* 2009;16:416-9.
132. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine [Review]. *Clin Chem* 1993;39:561-77.
133. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361-87.
134. Lin L. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989;45:255-68.
135. Lilja H, Ulmert D, Vickers AJ. Prostate-specific antigen and prostate cancer: prediction, detection and monitoring. *Nat Rev Cancer* 2008;8:268-78.
136. Miele ME. Percent free PSA as an additional measure in a prostate cancer screen. *Clin Lab Sci* 2001;14:102-7.
137. Cook N. Use and Misuse of the Receiver Operating Characteristic Curve in Risk Prediction. *Circulation* 2007;115:928-35.
138. Pencina MJ, D'Agostino RBSr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27:157-72.
139. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006;26:565-74.
140. Benecchi L, Pieri AM, Melissari M. A novel nomogram to predict the probability of prostate cancer on repeat biopsy. *J Urol* 2008;180:146-9.
141. Kawamura K, Suzuki H, Kamiya N. Development of a new nomogram for predicting the probability of a positive initial prostate biopsy in Japanese patients with serum PSA levels less than 10 ng/ml. *Int J Urol* 2008;15:598-603.
142. Nam RK, Toi A, Klotz LH. Assessing individual risk for prostate cancer. *J Clin Oncol* 2007;25:3582-8.
143. Rochester MA, Pashayan N, Matthews F. Development and validation of risk score for predicting positive repeat prostate biopsy in patients with a previous negative biopsy in a UK population. *BMC Urol* 2009;9:7.

Eidesstattliche Erklärung

Ich, Kerstin Siemßen, erkläre an Eides statt, dass ich die vorgelegte Dissertation mit dem Titel

„Evaluierung des Einflusses der PSA-Test-Variabilität auf die Vorhersage eines Prostatakarzinoms mittels Nomogrammen“

in der Klinik für Urologie der Medizinischen Fakultät Charité – Universitätsmedizin Berlin mit Unterstützung durch Herrn PD Dr. med. Carsten Stephan und Herrn Dr. ing. Henning Cammann ohne sonstige Hilfe selbst verfasst habe.

Bei der Erstellung der Dissertation habe ich keine anderen als die aufgeführten Quellen und Hilfsmittel benutzt. Die vorliegende Dissertation beruht auf den Ergebnissen eigenständiger wissenschaftlicher Arbeit und stellt auch in Teilen keine Kopie anderer Arbeiten dar.

Ich habe an keiner anderen Stelle ein Promotionsverfahren beantragt.

Berlin, den 04.01.2012

Kerstin Siemßen

Lebenslauf

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Danksagung

An erster Stelle danke ich meinem Doktorvater Herrn PD Dr. med. Carsten Stephan für die Überlassung des Themas und das damit in mich gesetzte Vertrauen.

Ich danke ihm sehr für seine Unterstützung, sein großes Engagement und die hervorragende Betreuung dieser Arbeit.

Herrn Dr. ing. Henning Cammann möchte ich für die Beratung in mathematisch-statistischen Fragen meinen Dank aussprechen.

Ich danke meinen Eltern sowie Sabina und Frank für all die Unterstützung.

Ein besonderer Dank gilt Sven - für sein Interesse an meiner Arbeit, seine Unterstützung und Motivation.

Berlin, den 04.01.2012

Kerstin Siemßen