

Aus der Klinik für Urologie am Campus Mitte  
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Analytische und klinische Validierung der fünf meistbenutzten  
Testkombinationen zur Bestimmung des totalen und freien  
prostata-spezifischen Antigens: Vergleichsmessungen an 798  
Seren von Patienten mit Prostatakarzinom und benigner  
Prostatahyperplasie

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

vorgelegt der Medizinischen Fakultät  
Charité – Universitätsmedizin Berlin

von

Moritz Klaas

aus Karlsruhe

Gutachter: 1. Prof. Dr. K. Jung  
2. Priv.-Doz. Dr. St. Weikert  
3. Priv.-Doz. Dr. G. Popken

Datum der Promotion: 07.12.2007

Meiner Sunčica

Teile dieser Dissertation sind in folgender Publikation erschienen:

Stephan C<sup>1</sup>, Klaas M<sup>1</sup>, 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

<sup>1</sup>These authors contributed equally to this work.

## Inhaltsverzeichnis

Inhaltsverzeichnis .....	I
Tabellenverzeichnis .....	III
Abbildungsverzeichnis .....	IV
Abkürzungsverzeichnis .....	V
<b>1. Grundlagen .....</b>	<b>1</b>
1.1. Das Prostatakarzinom .....	1
1.2. Das prostata-spezifische Antigen .....	2
1.2.1. Historie .....	3
1.2.2. Biologie und Physiologie des PSA .....	4
1.2.3. Validität und Verbesserungsvorschläge .....	6
1.3. Messgeräte zur Bestimmung des prostata-spezifischen Antigens .....	8
<b>2. Ziele der Arbeit .....</b>	<b>13</b>
<b>3. Material und Methoden .....</b>	<b>14</b>
3.1. Probenmaterial .....	14
3.2. Patientenkollektiv .....	14
3.4. PSA-Bestimmungen .....	15
3.5. Statistische Auswertungen .....	19
<b>4. Ergebnisse .....</b>	<b>20</b>
4.1. Charakterisierung der untersuchten Probanden .....	20
4.2. Qualitätskontrolle .....	22
4.3. Methodenvergleiche der tPSA-, fPSA- und %fPSA-Bestimmungen .....	25
4.3.1. Darstellung nach Passing und Bablok .....	25
4.3.2. Darstellung nach Bland und Altman .....	31
4.3.3. Mountain-Plot .....	35
4.3.4. Äquimolaritätsanalyse .....	38
4.3.5. Methodenvergleiche in der Subgruppe 0-10 µg/l .....	40

4.4.	Bewertung der klinischen Aussagekraft .....	44
4.4.1.	Die ROC-Analyse.....	44
4.4.2.	Diskriminierung anhand festgelegter Grenzwerte .....	51
<b>5.</b>	<b>Diskussion</b> .....	<b>53</b>
<b>6.</b>	<b>Zusammenfassung</b> .....	<b>68</b>
<b>7.</b>	<b>Literaturverzeichnis</b> .....	<b>69</b>
	Erklärung .....	i
	Lebenslauf.....	ii
	Danksagung .....	iii

## Tabellenverzeichnis

Tabelle 1	Charakterisierung der verwendeten PSA-Messsysteme entsprechend den Herstellerangaben.....	17
Tabelle 2	Arbeitsschritte der einzelnen Assays.....	18
Tabelle 3	Verteilung der Patienten in der PCa- und BPH-Gruppe anhand der tPSA-Konzentration.....	21
Tabelle 4	Intraserielle Präzision der fPSA- und tPSA-Bestimmungen mit zwei Serumpools .....	23
Tabelle 5	Interserielle Präzision der fPSA- und tPSA-Bestimmungen mit kommerziellen Kontrollmaterialien.....	24
Tabelle 6	Methodenvergleiche einzelner Tests für tPSA, fPSA %fPSA mit der Referenzmethode entsprechend der nicht-parametrischen Regressionsanalyse nach Passing und Bablok.....	29
Tabelle 7	Berechnete Flächen der ROC-Kurven aller Patientengruppen.....	50
Tabelle 8a	Spezifitäten und Diskriminationspunkte bei festgelegten 90%- und 95%igen Sensitivitäten bei tPSA-Methoden .....	51
Tabelle 8b	Spezifitäten und Diskriminationspunkte bei festgelegten 90%- und 95%igen Sensitivitäten bei %fPSA-Methoden in der Subgruppe tPSA 0-10 µg/l.....	51
Tabelle 9a	Einordnung der Patienten in Erkrankungsgruppen anhand festgelegter Grenzwerte für tPSA (2,5 bzw. 4 µg/l).....	53
Tabelle 9b	Einordnung der Patienten in Erkrankungsgruppen anhand festgelegter Grenzwerte für %fPSA (15, 20 bzw. 25 %) in der Subgruppe tPSA 0-10 µg/l.....	53

## Abbildungsverzeichnis

Abbildung 1	Medianwerte der Methoden in Prozent zur Referenzmethode .....	28
Abbildung 2	Bland und Altman-Schaubilder der tPSA-Werte verschiedener Methoden.....	32
Abbildung 3	Bland und Altman-Schaubilder der fPSA-Werte verschiedener Methoden.....	33
Abbildung 4	Bland und Altman-Schaubilder der %fPSA-Werte verschiedener Methoden.....	34
Abbildung 5	Mountain-Plots für die tPSA-Methoden.....	36
Abbildung 6	Mountain-Plots für die fPSA-Methoden .....	37
Abbildung 7	„Molar-Response“-Graphen für tPSA-Methoden.....	39
Abbildung 8	Medianwerte der Methoden in Prozent zur Referenzmethode der Subgruppe 0-10 µg/l.....	42
Abbildung 9	„Molar-Response“-Graphen für tPSA-Methoden der Subgruppe 0- 10 µg/l.....	44
Abbildung 10	ROC-Kurven für tPSA-Methoden mit Proben aller Patienten sowie der Subgruppe 0-10 µg/l.....	48
Abbildung 11	ROC-Kurven für %fPSA-Methoden mit Proben aller Patienten sowie der Subgruppe 0-10 µg/l.....	49

## Abkürzungsverzeichnis

ACT-PSA	$\alpha_1$ -Antichymotrypsin-PSA-Komplex
AK	Antikörper
BPH	benigne Prostatahyperplasie
BPSA	BPH-assoziiertes PSA
cPSA	komplexiertes PSA
DRU	digital-rektale Untersuchung
fPSA	freies PSA
hK	humanes Kallikrein
iPSA	intaktes PSA
k.A.	keine Angabe
kDa	Kilodalton
K.I.	Konfidenzintervall
PCa	Prostatakarzinom
PSA	prostata-spezifisches Antigen
ROC	Receiver-operation-characteristics
tPSA	Gesamt-PSA
proPSA	Vorläufer-PSA
SD	Standardabweichung
VK	Variationskoeffizient
WHO	Weltgesundheitsorganisation
%fPSA	Quotient fPSA/ tPSA

# 1. Grundlagen

## 1.1. Das Prostatakarzinom

Mit jährlich etwa 218 890 Neuerkrankungen und 27 050 Todesfällen stellt das Prostatakarzinom (PCa) in den USA die mit Abstand am häufigsten vorkommende bösartige Neubildung des Mannes dar und steht in der Krebstodesstatistik an zweiter Stelle nach dem Bronchial-Karzinom (1). In Deutschland ist das PCa neben dem malignen Melanom der Haut die häufigste Krebserkrankung des Mannes (2).

Die Prävalenz des PCa steigt mit zunehmendem Alter deutlich an. Bezüglich der Inzidenz bestehen ethnische und geographische Unterschiede (1). Das Risiko, an einem PCa zu erkranken, wird durch genetische Faktoren und Ernährung beeinflusst (3).

Die Prognose eines Prostatakarzinompatienten hängt entscheidend von Differenzierung und Ausbreitung der Neoplasie bei Diagnosestellung ab. Ein Patient mit einem lokal-begrenzten, nicht-metastasierten Tumor kann durch eine radikale Prostatovesikulotomie oder mittels Bestrahlung geheilt werden. Wird der Tumor erst nach Kapsel-überschreitung entdeckt, fällt die durchschnittliche 5-Jahres-Überlebensrate auf 33 % ab (1). Dem Patienten stehen verschiedene Behandlungsoptionen sowohl mit kurativem als auch mit palliativem Ansatz offen. Neben anderen Faktoren ist vor allem die Lebenserwartung des Betroffenen wichtig für die weitere Therapieplanung (4).

Klinisch wird der Prostatakrebs in der Regel erst in späten Stadien symptomatisch. Um jedoch das PCa bei Männern in einem frühen, auf die Prostata begrenzten Stadium zu diagnostizieren, wurde 1971 ein gesetzliches Früherkennungsprogramm für Männer ab dem 45. Lebensjahr eingeführt. Es sieht eine jährliche digital-rektale Untersuchung (DRU) der Prostata sowie eine klinische Anamnese vor. Da die DRU nur eine relativ geringe diagnostische Sensitivität aufweist und mit ihr zwei Drittel der Karzinome erst nach Kapselüberschreitung erkannt werden, ist diese Methode als alleinige Screeninguntersuchung unzureichend (5). Erst nach Einführung des prostata-

spezifischen Antigens (PSA) zur Früherkennung des PCa vor 20 Jahren konnte eine deutliche Verschiebung hin zu mehr organbegrenzten Tumoren erreicht werden (6). Eine statistisch signifikante Steigerung der relativen 5-Jahres-Überlebensrate von 70 % in den 70er Jahren auf jetzt fast 100 % ist neben den Fortschritten, die in den Bereichen der transrektalen Ultraschalluntersuchung, der Magnet-Resonanz-Tomographie oder der Prostatabiopsie gemacht wurden, höchstwahrscheinlich auf die verbesserte Prostatakarzinomdetektion durch PSA zurückzuführen (1). Heutzutage werden etwa 91 % der Prostatakarzinome in lokalem Stadium diagnostiziert (1). Ob die PSA-Bestimmung die Mortalität des PCa senkt, wird zurzeit in zwei groß angelegten prospektiven Studien untersucht (7,8).

## 1.2. Das prostata-spezifische Antigen

Neben der transrektalen Ultraschalluntersuchung der Prostata, die eine morphologische Beurteilung der Prostata ermöglicht und zur Detektion karzinomähnlicher Areale sowie zur Steuerung der Prostatabiopsie eingesetzt wird, und der DRU ist vor allem das PSA die wichtigste Methode zur Früherkennung und Therapieüberwachung des PCa (9).

Es konnte jedoch gezeigt werden, dass das PSA nicht tumorspezifisch ist. Auch gutartige Prostataerkrankungen oder Manipulationen an der Prostata können einen „pathologischen“ Anstieg der PSA-Konzentration verursachen (10). Andererseits hat ein Teil der Männer mit einem „normalen“ PSA bereits ein Prostatakarzinom (11-13). Infolgedessen ist bei niedrigen PSA-Konzentrationen die Diskriminationsfähigkeit dieses Tumormarkers zwischen benignen und malignen Prostataerkrankungen begrenzt (9). Um möglichst viele PCa-Patienten mit einem frühen Karzinom zu erfassen und gleichzeitig unnötige Prostatabiopsien bei nicht betroffenen Männern zu vermeiden, wurden in den letzten Jahren verschiedene Konzepte entwickelt. Hierbei konnte vor allem durch die zusätzliche Bestimmung der PSA-Untergruppen, wie dem freien PSA (fPSA) und dessen Verhältnis zum Gesamt-PSA (%fPSA) oder dem komplexierten PSA (cPSA), eine Steigerung der diagnostischen Spezifität bei der Prostatakarzinomdetektion erzielt werden (9,14,15). Die Entwicklung %fPSA-orientierter, artifizieller neuronaler Netzwerke stellt einen weiteren Fortschritt bei der Optimierung der Prostatakarzinomdiagnostik dar (16).

Durch die breite Anwendung der PSA-Teste wurden verschiedene Geräte zur Bestimmung von Gesamt-PSA (tPSA) und seiner Untergruppen entwickelt. Allerdings ist unter den Testsystemen eine starke Variabilität der tPSA- und vor allem fPSA-Konzentrationen zu verzeichnen (17). Diese methodenabhängigen Unterschiede können zu Fehlinterpretationen individueller tPSA und %fPSA-Werte führen und infolgedessen falsche klinische Entscheidungen einleiten (18). Durch internationale Standardisierungen wurde in den letzten Jahren versucht, eine bessere Vergleichbarkeit testspezifischer Messergebnisse herbeizuführen (19-22). Vor allem die uneinheitliche Kalibrierung der Testsysteme und die nicht-äquimolare Erkennung von tPSA wurden für die unterschiedlichen Messergebnisse verantwortlich gemacht (19,22-25). In der Folge wurden uniforme Referenzmaterialien entwickelt (26). Die Anpassung der Kalibrierung mit dem so genannten WHO-PSA-Referenzmaterial und die Einführung weiterer äquimolarer Methoden weckten Hoffnungen auf eine bessere Harmonisierung der PSA-Testsysteme (27-31).

Aus der folgenden Darlegung von Details zum PSA und zur Methodenabhängigkeit von PSA-Bestimmungen soll die Zielstellung der vorliegenden Arbeit abgeleitet werden.

### 1.2.1. Historie

Verschiedene Arbeitsgruppen beschrieben zu Beginn der 70er Jahre ein Protein in menschlicher Samenflüssigkeit, welches 1978 auf der Suche nach einem forensischen Marker für Vergewaltigungsdelikte durch Sensabaugh (32) mittels Immunelektrophorese genauer charakterisiert werden konnte (33-35). Kurze Zeit später wurde dieses Protein von Wang et al. (36) aus Prostatagewebe isoliert und als „prostata-spezifisches Antigen“ bezeichnet. Die Entwicklung eines neuen Tests machte dessen Bestimmung im Serum möglich und führten Wang et al. (37) zur Entdeckung eines neuen Tumormarkers (38).

In den folgenden Jahren konnte sich das PSA aufgrund einer besseren Korrelation zum Tumorstadium und einer höheren Sensitivität in der PCa-Diagnostik schnell gegenüber der sauren Phosphatase durchsetzen und wurde zum bedeutendsten Tumormarker der Urologie (39,40). Sein Einsatz bei der Verlaufskontrolle von Prostatakarzinompatienten ist unumstritten (9).

### 1.2.2. Biologie und Physiologie des PSA

Das PSA ist ein ca. 34 kDa schweres, zu 93 % aus Aminosäuren und zu 7 % aus Kohlenhydraten bestehendes Glykoprotein mit einer Primärstruktur aus 237 Aminosäuren (41,42). Wegen der Genlokalisierung auf Chromosom 19q13,3-q13,4, der Funktion als Serinprotease, der Regulierung durch Androgene sowie Homologien der DNA- und Aminosäuresequenzen wird das PSA der Gruppe der humanen Kallikreine (hK) zugeordnet (43-45). Neben PSA (auch als hK3 bezeichnet) könnten in naher Zukunft auch andere in der Prostata gebildete Kallikreine (hK2 und hK4) als Tumormarker in der Prostatakarzinomdiagnostik nützlich sein (46).

Das PSA wird in den Gang- und Drüsenepithelzellen der Vorsteherdrüse gebildet und in das Lumen der Prostata drüse sezerniert. Als Hauptprotein des Seminalplasmas (0,5-2 g/l), dient es der enzymatischen Verflüssigung des gelförmigen Ejakulats, wodurch die Spermienmotilität erhöht wird (47). Bei gesunden Männern verhindert eine intakte Prostata drüsenstruktur den Übertritt von PSA in den Blutstrom. Selbst wenn ein retrograder Übertritt nicht vollständig verhindert werden kann, sind im Normalfall PSA-Konzentrationen im Serum  $10^6$ -fach niedriger als im Sperma. Eine Zerstörung von Prostata gewebe durch Erkrankungen oder Manipulationen der Prostata kann jedoch zur Aufhebung der physiologischen Barriere führen und den Übertritt von PSA in den Blutstrom erleichtern. Erhöhte PSA-Blutspiegel erlangen dadurch diagnostische Bedeutung.

Obwohl verschiedene Studien zeigten, dass das PSA in kleineren Mengen auch in anderen Geweben sowohl beim Mann als auch bei der Frau vorkommt und somit nicht organspezifisch ist, haben diese Funde aus praktischer Sicht keine wesentliche Bedeutung (zusammengefasst in (9)). Weitere Hypothesen, wie etwa diejenige, dass das PSA als Wachstumsfaktor das Prostatakarzinom beeinflusst, konnten bisher noch nicht bestätigt werden (48).

Im Serum tritt das PSA in zwei verschiedenen Formen auf, dem an andere Proteine gebundenen, so genannten komplexierten PSA und dem freien PSA, das nicht an andere Proteine gebunden ist:

1. Der Hauptanteil des PSA besteht aus enzymatisch aktiven Molekülen, die Komplexe mit verschiedenen Protease-Inhibitoren bilden (49,50). Obwohl aktives

PSA bevorzugt an  $\alpha_2$ -Makroglobulin bindet, kann dieser Komplex mit konventionellen PSA-Tests nicht erfasst werden, da sämtliche PSA-Epitope verdeckt sind (51). Zudem wird dieser Komplex schnell hepatisch eliminiert (mit einer Plasma-halbwertszeit von 6,7 Minuten), so dass Serumkonzentrationen ohnehin kaum messbar sind (50). Demgegenüber hat der  $\alpha_1$ -Antichymotrypsin-PSA-Komplex (ACT-PSA) eine Halbwertszeit von 3 Tagen und ist mit einem Anteil von 70-90 % am Gesamt-PSA (tPSA) am bedeutendsten (50,52). Moleküle mit Bindung zu  $\alpha_1$ -Antitrypsin und Inter- $\alpha$ -Trypsin-Inhibitor kommen nur zu 5 % vor.

2. Das freie PSA (fPSA) macht etwa 5-35 % am tPSA aus. Diese enzymatisch inaktive Untergruppe bildet keine Komplexe mit Protease-Inhibitoren, wird glomerulär filtriert und hat eine Halbwertszeit von 12-18 Stunden (50). In dieser Fraktion kommen unterschiedliche Subformen vor: proPSA, BPSA und iPSA (53). Es wird unterschieden zwischen Subformen mit einer durchgehenden und solchen mit einer inneren Spaltung der Aminosäurekette, wobei die erstgenannten als intaktes PSA (iPSA) und die letztgenannten als „nicked“-PSA bezeichnet werden. Innerhalb der iPSA-Varianten lassen sich Vorläuferproteine (proPSA) mit partiellem oder komplett erhaltenem Propeptid bestimmen. Weitere Varianten bezeichnen Isoformen mit unterschiedlichen Glykosylierungsmustern, gekürzte und alternativ gespaltene fPSA-Subformen. Eine fPSA-Subform mit innerer Spaltung der Aminosäurekette wird aufgrund der erhöhten Konzentration in Prostatagewebe der Übergangszone von Patienten mit symptomatischer BPH als BPH-assoziiertes PSA (BPSA) bezeichnet.

Aus noch nicht ganz geklärter Ursache weisen PCa-Patienten im Serum gegenüber gesunden Männern bzw. Männern ohne PCa einen erniedrigten relativen Anteil von fPSA am tPSA bzw. einen erhöhten Anteil von ACT-PSA auf (50,54). Hierdurch erlangt der Gebrauch des Quotienten fPSA/ tPSA (%fPSA) und des gebundenen PSA (cPSA) bei der Differenzierung zwischen PCa-Patienten und Patienten mit einer PSA-Erhöhung aufgrund nicht-maligner Ursache an Bedeutung (55,56).

### 1.2.3. Validität und Verbesserungsvorschläge

Trotz der besseren Früherkennung des PCa dank des PSA verfügt die alleinige Verwendung von PSA in der PCa-Diagnostik nur über eine eingeschränkte Spezifität und einen geringen positiv-prädiktiven Wert (1,9).

Ergebnisse einer Multicenterstudie am Vorgängermodell des heutigen Beckman Access Hybritech Tests legten den tPSA-Grenzwert von 4 µg/l zur Prostatabiopsieindikation nahe (57). Drei Viertel der PCa, die bei einem tPSA-Wert von 4-10 µg/l entdeckt werden, sind lokalisiert und damit potentiell heilbar. Bei Männern mit einem tPSA >10 µg/l fällt dieser Anteil unter 50 %. Andererseits stellt die Erfassung von PCa-Patienten mit einem PSA <10 µg/l eine große Herausforderung dar, weil in diesem Bereich eine hohe Anzahl falsch-positiver Testergebnisse zu vielen unnötigen Biopsien führt. Der positiv-prädiktive Wert liegt in diesem Bereich bei 25 %.

Serumwerte von tPSA und dessen Untergruppen können auch durch andere Ursachen, wie der benignen Prostatahyperplasie (BPH), der Prostatitis, unter Dialysetherapie, nach Ejakulation oder nach Manipulationen an der Prostata erhöht sein (9,58,59). Androgenantagonisten, die bei Prostatabeschwerden verschrieben werden, senken das PSA im Serum (9,60). Zur korrekten Interpretation von PSA-Konzentrationen müssen präanalytische Faktoren berücksichtigt werden (58). Intraindividuell schwanken Serumwerte von PSA um 20 % (61). Im PSA-Bereich 4-10 µg/l, der so genannten „Grauzone“, erweist sich vor allem die Unterscheidung zwischen PCa und BPH als besonders schwierig.

Beweisend für ein Prostatakarzinom ist jedoch ausschließlich der Nachweis von Krebszellen in bioptisch oder operativ entnommenen Gewebeproben. Die ultraschallgesteuerte Prostatabiopsie konnte in den letzten Jahren kontinuierlich verbessert werden, ist aber nicht frei von Risiken (62). Um die Anzahl unnötiger Prostatabiopsien zu verringern und gleichzeitig möglichst wenige Karzinome im Frühstadium zu übersehen, wurden Konzepte zur Steigerung der diagnostischen Leistung des PSA entwickelt.

Verschiedene Studien warnten davor, dass durch die strikte Anwendung des „konventionellen“ Grenzwerts von 4 µg/l eine bedeutende Anzahl organbegrenzter, zum Teil

aggressiver PCa übersehen werden würde (11-13). Allerdings würde die alleinige Absenkung des Grenzwertes mit einer höheren Fehlerquote einhergehen und bis zu drei Viertel der Patienten mit einem PSA unter 4 µg/l unnötigerweise biopsiert werden (63).

Die PSA-Dichte, die PSA-Anstiegsgeschwindigkeit, die PSA-Verdopplungszeit oder die Anwendung alters- und rassenspezifischer PSA-Referenzgrenzen erwiesen sich zwar als hilfreiche Mittel in der PCa-Diagnostik, aber konnten die diagnostische Aussagekraft des tPSA nicht ausreichend verbessern (9,64-67).

Eine deutliche Steigerung der Validität des Tumormarkers wurde durch die Aufnahme der PSA-Isoformen zur Prostatakarzinomdiagnostik erzielt. Im Vergleich zu gesunden Männern und BPH-Patienten ist bei PCa-Patienten der Quotient von fPSA und tPSA (%fPSA) signifikant erniedrigt und der Anteil von ACT-PSA an tPSA erhöht (55). Catalona et al. (68) zeigten in einer prospektiven Studie an 773 Männern, dass durch die zusätzliche Bestimmung von %fPSA in der „Grauzone“ (tPSA-Bereich unter 10 µg/l) bei etwa gleich bleibender Sensitivität die Anzahl unnötiger Biopsien reduziert werden kann. Daher sollten nur bei Patienten mit niedrigen %fPSA-Werten Biopsien entnommen werden. Uneinigkeit besteht jedoch weiterhin bezüglich des empfohlenen Grenzwertes zur Biopsieindikation. Ein Vergleich unterschiedlicher Studien ist unter anderem wegen der Variabilität der verwendeten Methoden kompliziert (siehe Kapitel 1.3) (14).

Des Weiteren wurde ein Test entwickelt, der mit cPSA alle an Protease-Inhibitoren gebundenen PSA-Moleküle, mit Ausnahme des  $\alpha_2$ -Makroglobulin-PSA, erfasst (56). Durch die Bestimmung der Hauptfraktion ergibt sich im Vergleich zum geringer vorkommenden fPSA eine günstigere analytische Ausgangsposition. Zudem ist cPSA stabiler als fPSA (69). Eine prospektive Studie an 831 Männern fand heraus, dass cPSA spezifischer als tPSA ist (70). Im Grauzonenbereich hat das cPSA bei festgelegter 90%iger Sensitivität eine höhere Spezifität als tPSA (13,3 bzw. 8,6 %), aber eine niedrigere als %fPSA (21,5 bzw. 21,9 %). Es ist aber weiterhin unklar, ob tPSA durch cPSA als Screeningtest abgelöst werden sollte.

Schon bei den Vergleichen zwischen der PSA-Messung und der DRU in der PCa-Diagnostik zu Beginn der 90er Jahre konnte beobachtet werden, dass durch die Kombi-

nation beider Methoden eine größere diagnostische Effizienz erreicht werden kann (57). Folglich stellt die Entwicklung von %fPSA orientierten, artifiziellen neuronalen Netzwerken und logistischen Regressionen einen weiteren Fortschritt bei der Optimierung der Prostatakarzinomdiagnostik dar (16,71). Artifizielle neuronale Netzwerken erstellen aus %fPSA und tPSA sowie den Daten der klinischen Untersuchungen (Alter, Prostatavolumen, DRU) ein individuelles Risikoprofil und geben eine Empfehlung über die Notwendigkeit einer Biopsie ab. Dadurch kann die Leistung der %fPSA-Bestimmung weiter verbessert und die Zahl der unnötigen Biopsien im erweiterten Graubereich 2-10 µg/l nochmals gesenkt werden. Neuere Studien zeigen eine verbesserte Früherkennung unter Miteinbeziehung der PSA-Dichte, vor allem im niedrigen PSA-Bereich (72). Auch der Einsatz neuer Tumormarker wie hk2, hk11, proPSA, MIC-1 (macrophage-inhibitory cytokine-1) und MIF (macrophage migration inhibitory factor) in artifiziellen neuronalen Netzwerken kann zusätzlich Biopsien einsparen (73-75).

### **1.3. Messgeräte zur Bestimmung des prostata-spezifischen Antigens**

Mit Zunahme der klinischen Nutzung des PSA stieg die Anzahl der kommerziell erhältlichen PSA-Testmethoden stetig. Der erste PSA-Immunoassay wurde Anfang der 80er Jahre entwickelt (76). Heutzutage sind über 30 verschiedene Testsysteme zur Bestimmung von tPSA und etwa 10 zur Bestimmung von fPSA bzw. cPSA erhältlich (77). Allerdings wurde nachgewiesen, dass Testsysteme zur Bestimmung von tPSA und fPSA bei gleichen Serumproben unterschiedliche Messergebnisse liefern (17).

Schon die beiden ersten kommerziell verfügbaren Methoden, der Pros-Check und der Tandem-R Hybritech, zeigten erhebliche Differenzen (78). Nur in Kenntnis des Referenzbereichs der Bestimmungsmethode, Pros-Check <2,5 µg/l und Tandem-R <4 µg/l, konnten die Ergebnisse korrekt bewertet werden. Die Einführung weiterer Tests mit methodenabhängigen Messergebnissen erschwerte die Interpretation des PSA außerordentlich, zumal sich zunehmend der klinisch etablierte Referenzbereich des Tandem-R durchsetzte (23,57). So wird die Indikation zur Prostatabiopsie bei einem Wechsel des Testsystems vor allem im allgemein üblichen Grenzbereich von 4 µg/l erschwert und kann zu einer ungleichen Einstufung von Männern mit erhöhten PSA-

Konzentrationen und folglich zu einer unterschiedlich hohen Biopsierate führen (23). Studien zum PSA, die an einem Testsystem durchgeführt wurden, können nicht automatisch auf ein anderes übertragen werden. Die mangelnde Vergleichbarkeit methodenabhängiger Messergebnisse führt also im analytischen und klinischen Sinn zu einer Herabstufung des Tumormarkers.

Als mögliche Ursachen für die methodenabhängige Variabilität von PSA-Messungen wurden verschiedenste Faktoren angeführt (23,79). Darunter scheinen neben dem individuellen Aufbau des Testsystems und der unterschiedlichen Epitop-Bindungsplätze der verwendeten PSA-Antikörper in den verschiedenen PSA-Methoden vor allem die nicht-äquimolare Erkennung von freiem und gebundenem PSA und die uneinheitliche Kalibrierung der Testsysteme wichtig zu sein (18,23,80).

Die Äquimolarität einer Methode beschreibt das Verhältnis der gemessenen PSA-Untergruppen zueinander. Eine äquimolare Methode bestimmt tPSA aus der absoluten Konzentration von fPSA und ACT-PSA und ist unabhängig von den Relationen. Eine nicht-äquimolare Methode erfasst in der Regel höhere fPSA-Konzentrationen und kann für Proben mit gleicher tPSA-Konzentration, aber unterschiedlichen Anteilen an fPSA unterschiedliche Ergebnisse liefern (81,82). Nicht-äquimolare Methoden sollen tendenziell zu einer überhöhten tPSA-Bestimmung bei Patienten mit benigner Prostatahyperplasie und Unterbewertung von PSA bei PCa-Patienten neigen (29). 1994 zeigte ein Vergleich zwischen dem Tandem-R und dem nicht-äquimolaren Abbott IMx nur für Proben mit niedrigen fPSA-Anteilen eine gute Übereinstimmung der Ergebnisse (83). Die höhere tPSA-Bestimmung bei steigendem %fPSA des IMx könnte zu einer niedrigeren Spezifität führen. Umgekehrt neigte aber der Tandem-R zu einer niedrigeren Sensitivität. Der Gebrauch äquimolarer Methoden kann die Anzahl der unnötigen Biopsien um bis zu 10 % verringern (29,84,85). Die Gründe für die unterschiedlichen äquimolaren Charakteristiken sind komplex. Antikörper nicht-äquimolarer Methoden sollen auch an Epitope des fPSA binden, die normalerweise durch die Komplexbildung zum ACT-PSA verdeckt werden (54). Bisweilen wurden auch polyklonalen Antikörperseren eine Nicht-Äquimolarität zugeschrieben, weil sie eine größere Affinität zu fPSA besitzen und mit PSA-Formen reagieren die von monoklonalen Antikörpern nicht erkannt werden (25,86,87). In der Regel haben nur wenige Antikörper äquimolare Charakteristiken. In einer Studie von Corey et al. (86)

konnten bei nur 3 von 90 neuen Antikörperkombinationen äquimolare Eigenschaften mit ausreichender Signalstärke nachgewiesen werden. Ein anderer möglicher Unterschied bezieht sich auf die Reaktionskinetiken der Methoden. Bei einer kurzen Inkubationszeit werden PSA-Formen mit einem kleineren Molekulargewicht schneller erfasst (20). Zwar wurde bei Vergleichen oft der äquimolare Hybritech Tandem-R als Referenz herangezogen, doch eine Referenzmethode im eigentlichen Sinn sowie ein internationaler Standard über eine einheitliche Kalibrierung existierte nicht (50,55,88,89). Auch in dieser Arbeit soll das Nachfolgemodell der Access Hybritech als Vergleichsmethode dienen, weil dieser Test die gleichen Antikörper verwendet und große Ähnlichkeiten mit dem Tandem-R besitzt, weit verbreitet ist und zu Beginn der Studie noch als äquimolar bekannt war (18,90).

Um die unzureichende Vergleichbarkeit methodenabhängiger Messergebnisse zu beheben, wurden in den 90er Jahren zwei Konferenzen zur Standardisierung von PSA-Immunoassays organisiert (19-22,91). Die zuvor mehrfach beschriebene Variabilität von tPSA-Konzentrationen zwischen verschiedenen Testsystemen sollte durch die Kalibrierung mit einem Referenzmaterial und durch die Entwicklung äquimolarer Methoden deutlich reduziert werden. Methoden, die nicht weitgehend äquimolar alle genannten PSA-Formen erkennen, messen mit höchster Richtigkeit bei der Zusammensetzung, die der des Kalibrierungsstandards entspricht. In der Folge wurden einheitliche Referenzmaterialien nach dem so genannten WHO-PSA-Standard mit dem Quotienten  $90 \text{ ACT-PSA} / 10 \text{ fPSA}$  entwickelt (26,92). Dieser fPSA-Anteil wird typischerweise bei Serumproben von PCa-Patienten gemessen (93). Nicht-äquimolare Methoden, die bevorzugt fPSA bestimmen, sollten dadurch in diesem Bereich ähnliche Ergebnisse liefern. Die Verwendung von 90:10 Mischungen in Qualitätskontrollen führt zu signifikanter Abnahme der Unterschiede zwischen den Methoden (94). Da aber bei Proben mit höheren fPSA-Anteilen, wie etwa bei gesunden Männern oder BPH-Patienten, weiterhin falsch-hohe Abweichungen resultieren können, kann die einheitliche Kalibrierung die Umstellung der Methoden hin zu äquimolarer Charakteristik nicht ersetzen (94-96).

Es musste festgestellt werden, dass auch für fPSA-Konzentrationen und damit auch für %fPSA-Werte eine erhebliche Methodenabhängigkeit zu verzeichnen ist (18,94). Die aus fPSA und tPSA kalkulierten %fPSA-Werte weichen noch stärker voneinander ab als

die beiden alleine. Vor allem unter %fPSA-Bestimmungen kann die Applikation eines einheitlichen Grenzwertes falsche klinische Schlussfolgerungen nach sich ziehen (18). Einheitliche Grenzwerte bei unterschiedlich hohen %fPSA-Bestimmungen wirken sich auf die Anzahl der für eine Prostatabiopsie in Frage kommenden Patienten aus (19,94,97). Die willkürliche Kombination von fPSA- und tPSA-Konzentrationen unterschiedlicher Methoden zur Bestimmung von %fPSA sollte unterlassen werden, da dies zu einer deutlichen Verringerung der diagnostische Leistung führen kann (98). Das erste internationale WHO-Referenzmedium zur Kalibrierung von fPSA-Methoden wird von den meisten Herstellern noch nicht eingesetzt (30).

In den letzten Jahren konnten Erfolge auf dem Weg zur Harmonisierung der gemessenen PSA-Konzentrationen durch Methoden unterschiedlicher Hersteller gemeldet werden (99). Die Kalibrierungen der einzelnen Methoden durch PSA-Referenzmaterialien der WHO konnten aneinander angeglichen werden (30). Blijenberg et al. (27) und Stamey et al. (26) konnten eine verbesserte Korrelation zwischen den PSA-Methoden demonstrieren, die zuvor nach dem WHO-Standard kalibriert wurden. Es wurden weitere äquimolare Methoden entwickelt (31,99).

Trotz verbesserter Präzision bestehen weiterhin großen Unterschiede zwischen den Methoden (18). Zwar konnte in den Niederlanden der Variationskoeffizient von Anfang der 90er Jahre bis 2000 um 10-45 % verringert werden, ist aber mit 15-20 % noch relativ hoch. Im Jahre 2000 bezifferten sich die Abweichungen zwischen den in Deutschland erhältlichen Tests noch um 30-40 % (18). Tatsächlich haben die meisten Hersteller (darunter auch die hier untersuchten Tests von Abbott, Beckman Coulter, Diagnostic Product Corp. und Roche Diagnostics) in den letzten Jahren den Testaufbau verändert (31,90,100). Bisher konnte nur der Beckman Hybritech-Test (und neuerdings auch Abbott AxSYM und Roche Elecsys 2010) die Zulassung der amerikanischen FDA (Food and Drug Administration) für die Bestimmung von %fPSA erlangen. Zu Beginn dieser Dissertation existierten nur wenige Studien mit entweder einem zu kleinen Probenumfang oder einer limitierten Anzahl an untersuchten Methoden, um die aktuelle Situation richtig abbilden zu können (101-103). Eine Studie aus dem Jahr 2000 an 26 Methoden konnte zwar Verbesserungen in der Standardisierung erkennen, kam aber zu dem Ergebnis, dass nur jede dritte Methode äquimolar misst (101). Mehrere Studien konnten eine Variabilität der tPSA und %fPSA-Ergebnisse zwischen den Methoden

demonstrieren (18,102-104). Hinsichtlich der Auswirkungen methodenabhängiger PSA-Ergebnisse auf die klinische Leistung wurden widersprüchliche Studien veröffentlicht (102,104).

## 2. Ziele der Arbeit

Die Entwicklung neuer äquimolarer Methoden und einheitlicher WHO-PSA-Referenzmaterialien zur Kalibrierung sollten die Variabilität von tPSA- und fPSA-Konzentrationen zwischen verschiedenen kommerziellen Testsystemen verringern und eine bessere Interpretation der PSA-Werte ermöglichen. Vorangegangene Studien konnten die gegenwärtige Situation nur unzureichend abbilden. In einer Studie an 798 Serumproben von PCa- und BPH-Patienten sollen die fünf meistbenutzten Analysesysteme zur Bestimmung von tPSA und fPSA (in einem Fall cPSA) miteinander verglichen und hinsichtlich der Variabilität der PSA-Werte zwischen den Methoden, den äquimolaren Eigenschaften und der diagnostischen Leistung charakterisiert und bewertet werden.

Zusammenfassend sollten folgende Fragen geklärt werden:

- Wie sind die analytischen Zuverlässigkeiten der Tests zu bewerten?
- Konnten die Diskrepanzen methodenabhängiger Messergebnisse verringert werden? Liegt eine ausreichende Austauschbarkeit der Testergebnisse vor?
- Konnten die Hersteller das Ziel der äquimolaren Erfassung von freiem und gebundenem PSA erreichen?
- Wie ist die diagnostische Aussagekraft der einzelnen Tests? Beeinflusst der Gebrauch einheitlicher Grenzwerte die klinische Diskriminationsfähigkeit der Tests?
- Sind die Ziele der Standardisierung von PSA-Messsystemen erreicht worden?

### **3. Material und Methoden**

#### **3.1. Probenmaterial**

Die Bestimmungen erfolgten an Serumproben von insgesamt 798 Patienten, welche in einem Zeitraum von Januar 1999 bis März 2004 abgenommen und von Juni bis August 2004 analysiert wurden. Um mögliche Fehlmessungen zu vermeiden, fanden die Blutentnahmen entweder vor Biopsiegewinnung sowie mindestens sechs Wochen nach Manipulationen an der Prostata, wie der rektalen Tastuntersuchung oder dem transrektalen Ultraschall, statt. Nach der Entnahme wurden die Proben 15 Minuten lang bei 4 °C unter 1600 G zentrifugiert und das aus dem Überstand gewonnene Serum bei -80 °C eingefroren. Die Proben wurden erst am Messtag bei Raumtemperatur wieder aufgetaut und innerhalb der ersten drei Stunden bestimmt.

#### **3.2. Patientenkollektiv**

Diese Studie wurde in Übereinstimmung mit den ethischen Standards der Helsinki Deklaration von 1975 sowie der überarbeiteten Fassung von 1996 durchgeführt. Alle Teilnehmer wurden an der Urologischen Klinik der Charité untersucht und waren mit der Auswertung ihrer Daten einverstanden.

Unter den 798 Patienten befanden sich 468 Patienten mit Prostatakarzinom und 330 Patienten mit benigner Prostatahyperplasie. Bei allen Patienten wurde von Fachärzten der Urologischen Klinik ein klinischer Status einschließlich einer digital-rektalen Tastuntersuchung und einer Prostatavolumenmessung erhoben. Die Probanden waren bezogen auf das Prostataleiden zum Zeitpunkt der Blutentnahme weder medikamentös noch operativ vorbehandelt. Die Diagnosen waren durch histologische Untersuchungen gesichert (Oktantenbiopsien). Bei 377 der PCa-Patienten beziehen sich die Befunde (TNM-Klassifikation; histologische Differenzierung) auf die erhobenen Daten nach Untersuchung des Prostatektomiepräparates. Die histologischen Untersuchungen erfolgten routinemäßig im Institut für Pathologie der Charité in Berlin, CCM (Prof. Dr. M. Dietel). Diese Daten wurden den Krankenakten entnommen.

### 3.4. PSA-Bestimmungen

Die PSA-Messungen wurden an fünf vollautomatischen Immunoassay-Analysegeräten nach den Anweisungen der Hersteller vorgenommen. Die Messungen erfolgten an Geräten im Institut für Laboratoriumsmedizin und Pathobiochemie (Direktor: Prof. Dr. E. Köttgen). Bei der technischen Durchführung wurde ich dankenswerterweise ganz wesentlich von Frau Janett Reiche unterstützt. Die Analysegeräte sind im Einzelnen: Der AxSYM®- (Abbott Diagnostics GmbH, Wiesbaden, Deutschland), der Access®- (Hybritech®-PSA, Beckman Coulter Inc., Fullerton, USA), der Immulite 2000®-Test (DPC-Biermann GmbH, Bad Nauheim, Deutschland) und der Elecsys 2010®-Test (Roche Diagnostics GmbH, Mannheim, Deutschland) für die Messung von fPSA und tPSA sowie der ADVIA Centaur®- (Bayer Diagnostics, Tarrytown, USA) für cPSA und tPSA.

Prinzipiell arbeiten alle verwendeten Assays nach der „Sandwich-Technik“. Hierbei werden zwei (nicht-konkurrierende) Antikörper (AK), die an verschiedene Epitope des zu bestimmenden PSA-Moleküls binden, in das Analysat gegeben. Der erste AK (genannt „Festphasen-AK“) extrahiert das Antigen aus dem Reagenz. Der Zweite („Licht-Reagenz“) ist mit einer Substanz konjugiert, die eine in Lichtemission endende chemische Reaktion einleitet. Ein Photometer misst das entstandene Lichtsignal und ermittelt anhand zuvor erstellter Kalibrationskurven die PSA-Konzentration in der Probe. Bis auf die Verwendung anderer Antikörper verläuft die Bestimmung des freien PSA prinzipiell gleich. Die Tabellen 1 und 2 zeigen die Charakteristika bzw. Arbeitsschritte der Messsysteme.

Alle Messgeräte wurden mit herstellerspezifischem Kalibrationsmaterial geeicht.

Als Kontrollmaterial für Präzisionsbestimmungen und Richtigkeitskontrollen standen herstellerspezifische Kontrollproben, Kontrollseren von Bio-Rad Laboratories (Richmond, USA) und ein hauseigener Serumpool zur Verfügung. Es wurden intra-serielle und interserielle Messungen vorgenommen.

Da am ADVIA Centaur nicht fPSA sondern cPSA gemessen wird, wurde zum besseren Vergleich der fPSA-Testsysteme aus cPSA des ADVIA Centaur mit folgender Formel

berechnet:  $fPSA = tPSA - cPSA$ . Der prozentuale Anteil des fPSA am tPSA wurde anhand der Formel  $\%fPSA = (fPSA / tPSA) \times 100$  ermittelt.

**Tabelle 1.** Charakterisierung der verwendeten PSA-Messsysteme entsprechend den Herstellerangaben.

Hersteller Analysensystem	Beckman-Coulter Access Hybritech	Abbott AxSYM	Bayer ADVIA Centaur	DPC-Biermann Immulite 2000	Roche Elecsys 2010
Messprinzip <sup>a</sup>	MSA	MEIA	MSA	LIA	MSA
Verwendete AK <sup>b</sup>	m (Maus)/ m (Maus)	m (Maus)/ m (Maus)	p (Ziege)/ m (Maus)	p (Ziege)/ m (Maus)	m (Maus)/ m (Maus)
Äquimolare Messung <sup>c</sup>	ja	ja	ja	k.A.	ja
Kalibratorenzss.	fPSA	90 ACT-/10 fPSA	90 ACT-/10 fPSA	k.A.	90 ACT-/10 fPSA
<b>TPSA</b>					
Probenvolumen [µl]	25	184	35	25	20
Inkubationszeit [s]	15	k.A.	7,5	30	<18 (ges. Test)
Messbereich [µg/l]	0,008-150	0,04-? (k.A.)	0,01-100	0,04.150	0,002-100
Kein „high-dose-hook-effect“ bis [µg/l]	50	k.A.	100	22500	17000
<b>FPsA<sup>d</sup></b>					
Probenvolumen [µl]	25	184	35	25	20
Inkubationszeit [s]	15	k.A.	7,5	60	<18 (ges. Test)
Messbereich [µg/l]	0,005-20	0,02-? (k.A.)	0,03-100	0,02-25	0,01-50
Kein „high-dose-hook-effect“ bis [µg/l]	20	k.A.	8	10000	17000

<sup>a</sup>MEIA= Mikropartikelenzymimmunoassay, MSA= magnetischer Separationassay, LIA= lumeniszierender Immunoassay;

<sup>b</sup>m= monoklonaler Antikörper, p= polyklonaler Antikörper;

<sup>c</sup>Äquimolare Messung von fPSA und ACT-PSA vorhanden;

<sup>d</sup>Bei ADVIA Centaur Bestimmung von cPSA.

**Tabelle 2.** Arbeitsschritte der einzelnen Assays.

<b>Access:</b>	<p>Probe, mit alkalischer Phosphatase konjugierter AK und an paramagnetischen Partikeln gebundene AK werden zusammengegeben&gt; Inkubation&gt; Sandwich-Komplexbildung&gt; Waschschrift unter Magnetfeld&gt; Substratzugabe (Lumi-Phos 530)&gt; Reaktion des Substrats mit alkalischer Phosphatase&gt; Messung des entstandenen Chemielumneszenzsignals und Errechnen der Konzentration.</p>
<b>AxSYM:</b>	<p>Probe, Festphasen-AK und Verdünnung werden gemischt&gt; Inkubation&gt; Antigen-AK-Komplexe binden an Glasfasermembran &gt; Waschvorgang &gt; Substratzugabe (4-Methylumbelliferylphosphat)&gt; Messung des entstandenen Fluoreszenzsignals und Errechnen der Konzentration.</p>
<b>Centaur:</b>	<p>(Bei der cPSA-Bestimmung werden in einem ersten Schritt fPSA-Bindungsstellen durch AK blockiert)&gt; Probe und Acridiumester-markierter AK und an paramagnetische Partikel gebundener AK werden gemischt&gt; Inkubation&gt; Sandwich-Komplexbildung&gt; Waschschrift&gt; Zugabe von Säure und Base&gt; Messung des entstandenen Chemielumneszenzsignals und Errechnen der Konzentration</p>
<b>Immulite:</b>	<p>Probe und an Kugel gebundener AK werden gemischt&gt; Inkubation&gt; erster Waschvorgang&gt; mit alkalischer Phosphatase konjugierter AK bindet an Antigen-AK-Komplex&gt; Inkubation&gt; zweiter Waschvorgang&gt; Sandwich-Komplex-Bildung&gt; chemielumneszierendes Substratmodul wird zugegeben&gt; Messung des gebildeten Chemielumneszenzsignals und Errechnen der Konzentration</p>
<b>Elecsys:</b>	<p>Probe, biotinylierter AK und Ruthenium-Komplex-markierter AK werden gemischt&gt; Inkubation&gt; Zugabe von strepavidin-beschichteten Mikropartikeln&gt; Bindung des Sandwichkomplexes über Biotin-Streptavidin-Wechselwirkung an Festphase&gt; Elektrode fixiert magnetische Partikel&gt; Waschvorgang&gt; Spannung induziert Chemielumneszenzreaktion&gt; Photomultiplier misst und errechnet Konzentration</p>

### 3.5. Statistische Auswertungen

Statistische Berechnungen nahm ich an den Softwareprogrammen MedCalc® Version 9.0.1.0 (MedCalc Software, Belgien), GraphPad Prism® 4.03 (GraphPad Software, USA) und Excel 2002 für Windows vor. Ich wählte den Access-Hybritech Assay als Vergleichsmethode, weil dieser wie auch das Vorgängermodell Tandem-R oft beschrieben wurde und vor Beginn dieser Studie für seine äquimolare Messungen bekannt war (31,90).  $P < 0,05$  wurde als statistisch signifikant gewertet. Bei multiplen Vergleichen passte ich das Signifikanzniveau entsprechend der Bonferroni-Korrektur (105) ( $P < 0,05 / [\text{Zahl der miteinander verglichenen Tests}]$ ) an. Als statistische Tests wurden der Wilcoxon-Test und McNemar-Test angewendet.

Methodenvergleiche habe ich nach den Verfahren von Passing und Bablok (106) sowie Bland und Altman (107) vorgenommen. Mit dem Mountain-Plot stellte ich eine weitere graphische Methode zum analytischen Vergleich dar (108). Zur Beschreibung der Äquimolarität der einzelnen Bestimmungsmethoden benutzte ich Auftragungen und Berechnungen nach der Methode von Semjonow et al. (31). Die ROC-Analyse verwendete ich, um die Validitäten der Testsysteme zu ermitteln und miteinander zu vergleichen (109). Für nähere Erläuterungen der angewandten statistischen Methoden möchte ich auf die entsprechenden Kapitel im Ergebnisteil verweisen.

## 4. Ergebnisse

### 4.1. Charakterisierung der untersuchten Probanden

Die Probanden verteilten sich auf folgende zwei Gruppen: 486 Prostatakarzinom-Patienten mit einem Altersmedian von 63 Jahren (Spannweite: 43-79 Jahre) und 330 Patienten mit benigner Prostatahyperplasie und einem Altersmedian von 66 Jahren (Spannweite: 38-85 Jahre). In den folgenden Kapiteln werde ich sowohl Berechnungen für das gesamte Patientenkollektiv als auch für die beiden Erkrankungsgruppen getrennt durchführen, um auf mögliche erkrankungsbedingte Unterschiede hinzuweisen.

Das mit transrektalem Ultraschall gemessene Prostatavolumen der PCa- und BPH-Patienten betrug im Median 34 bzw. 45,5 cm bei einer Spannweite von 10-110 bzw. 13-180 cm. Die digital-rektale Untersuchung der PCa- und BPH-Patienten ergab in 284 bzw. 30 Fällen einen positiven und in 184 bzw. 300 Fällen einen negativen Tastbefund. In der Gruppe der PCa-Patienten konnte nach histologischer Stadieneinteilung der Großteil der Karzinome einem Stadium T2 der TNM-Klassifikation zugeordnet werden (pT1= 4, T1= 15, pT2= 267, T2= 43, pT3= 105, pT4= 1) und war entsprechend dem Gleason-Score mittelgradig differenziert (pG1= 7, G1= 6, pG2= 250, G2= 65, pG3= 120, G3= 19, Gx= 1). Bei keinem der untersuchten Patienten wurde ein Lymphknotenbefall oder eine Metastasierung festgestellt (N0= 255, M0= 415, Nx= 213, Mx= 53).

Bei über zwei Drittel der insgesamt 798 Patientenserum wurde ein tPSA-Wert zwischen 0 und 10 µg/l gemessen (Tab. 3). In diesem klinisch bedeutsamen tPSA-Bereich war die Verteilung der PSA-Werte in beiden Erkrankungsgruppen nahezu gleich. Dies zeigt die Aufschlüsselung der Proben in den einzelnen PSA-Bereichen in Tabelle 3.

**Tabelle 3.** Verteilung der Patienten in der PCa- und BPH-Gruppe anhand der tPSA-Konzentration.

<b>tPSA [<math>\mu\text{g/l}</math>]<sup>a</sup></b>	<b>BPH-Patienten</b>	<b>PCa-Patienten</b>	<b>Alle Patienten</b>
0-2	77 (23 %)	15 (3 %)	92
0-4	143 (43 %)	68 (15 %)	211
2-4	69 (21 %)	53 (11 %)	122
0-10	282 (85 %)	314 (67 %)	596
2-10	208 (63 %)	299 (64 %)	507
4-10	139 (42 %)	246 (53 %)	385
10-27,1	48 (15 %)	154 (33 %)	202
Gesamt-Anzahl (0-27,1)	330	468	798

<sup>a</sup>Bezug auf die Werte mit dem Access-Hybritech-Test. In Klammern ist der jeweilige prozentuale Anteil der Patienten des jeweiligen PSA-Bereiches an der Gesamtzahl der Patienten in der PCa- bzw. BPH-Gruppe angegeben.

## 4.2. Qualitätskontrolle

Die Ergebnisse der intra- und interseriellen Präzisionsbestimmungen sind in den Tabellen 4 und 5 zusammengefasst. Die Variationskoeffizienten (VK) aller Kontrollen lagen unter 10 % und damit innerhalb der vorgegebenen Toleranzgrenze.

Die intraseriellen Bestimmungen für tPSA und fPSA mit einem hauseigenen Serumpool (Tab. 4) wiesen sowohl beim Vergleich der Ergebnisse mit den verschiedenen Messgeräten untereinander als auch zwischen den als „hoch“ und „niedrig“ definierten PSA-Bereichen keine großen Differenzen auf. Intraserielle Kontrollen mit den so genannten Systemproben der Hersteller wurden aufgrund unzureichender Materialverfügbarkeit nur an den Geräten AxSYM, Centaur und Elecsys vorgenommenen. Die intraseriellen VK-Werte lagen in allen Fällen unterhalb 7 %.

Die mit herstellereigenen Systemproben (Tab. 5) und hauseigenen Serumpools ermittelten Tag-zu-Tag-Präzisionen wiesen bei allen Methoden eine gute analytische Zuverlässigkeit auf (VKs <9 %). Auch hier war die analytische Zuverlässigkeit der tPSA-Bestimmungen mit denen der fPSA-Messungen vergleichbar. Tendenziell war die analytische Zuverlässigkeit in niedrigeren PSA-Bereichen, wie zu erwarten, geringer. Impräzisionskontrollen mit BioRad-Kontrollmaterialien für die tPSA-Bestimmungen zeigten eine vergleichbare Reproduzierbarkeit aller Methoden (VKs <9 %).

**Tabelle 4.** Intraserielle Präzision der fPSA- und tPSA-Bestimmungen mit zwei Serum-pools (n=10).

	<b>Access</b>	<b>AxSYM</b>	<b>Centaur</b>	<b>Immulite</b>	<b>Elecsys</b>
<b>tPSA</b>					
Niedrig					
Mittelwert [ $\mu\text{g/l}$ ]	2,33	1,99	2,04	2,38	2,31
SD [ $\mu\text{g/l}$ ]	0,05	0,04	0,10	0,12	0,03
VK [%]	1,9	1,8	4,8	5,2	1,4
Hoch					
Mittelwert [ $\mu\text{g/l}$ ]	17,0	16,2	15,3	17,8	16,7
SD [ $\mu\text{g/l}$ ]	0,26	0,66	0,73	0,71	0,37
VK [%]	1,5	4,1	4,8	4,0	2,2
<b>fPSA<sup>a</sup></b>					
Niedrig					
Mittelwert [ $\mu\text{g/l}$ ]	0,43	0,49	1,64	0,34	0,43
SD [ $\mu\text{g/l}$ ]	0,03	0,02	0,07	0,01	0,01
VK [%]	6,9	4,3	4,3	3,7	2,9
Hoch					
Mittelwert [ $\mu\text{g/l}$ ]	3,79	3,99	11,8	3,18	3,81
SD [ $\mu\text{g/l}$ ]	0,10	0,13	0,68	0,16	0,04
VK [%]	2,7	3,2	5,8	5,0	4,5

<sup>a</sup>Centaur: Angabe in cPSA statt fPSA.

**Tabelle 5.** Interserielle Präzision der fPSA- und tPSA-Bestimmungen mit kommerziellen Kontrollmaterialien.

Gerät	Test	Kontrollmaterial	Mittelwert [ $\mu\text{g/l}$ ]	SD [ $\mu\text{g/l}$ ]	VK [%]
<b>Access</b> (n= 17)	tPSA	PSA QC I	0,96	0,04	4,24
		II	14,8	0,46	3,12
		III	89,6	2,03	2,26
	fPSA	1	0,97	0,03	3,48
		2	12,3	0,45	3,66
<b>AxSYM</b> (n= 41)	tPSA	L	0,50	0,03	5,84
		M	4,08	0,22	5,43
		H	24,9	1,39	5,57
	fPSA	L	0,42	0,02	5,33
		M	1,02	0,05	5,21
		H	7,17	0,50	6,91
<b>Centaur<sup>a</sup></b> (n= 40)	tPSA	Ligand plus 1	0,39	0,02	5,98
		2	1,54	0,07	4,36
		3	13,2	0,65	4,93
	cPSA	1	2,73	0,14	5,14
		2	13,8	0,64	4,59
		3	69,1	3,86	5,59
<b>Immulite</b> (n= 20)	tPSA	TMCO I	2,57	0,20	2,64
		II	7,38	0,32	3,18
		III	30,6	1,79	3,76
	fPSA	TMCO I	1,11	0,10	6,61
		II	3,29	0,26	2,90
		III	14,2	1,03	6,69
<b>Elecsys</b> (n= 45)	tPSA	Preci control I	5,01	0,27	5,40
		II	47,6	3,44	7,23
	fPSA	I	1,07	0,05	4,48
		II	10,1	0,57	5,56

<sup>a</sup>Centaur: Angabe in cPSA statt fPSA.

### **4.3. Methodenvergleiche der tPSA-, fPSA- und %fPSA-Bestimmungen**

Im Kapitel 3.5 habe ich dargelegt, dass für den Vergleich zwischen den verschiedenen PSA-Bestimmungen die Messungen mit den PSA-Methoden der Fa. Beckman-Coulter als Bezug dienen. Die folgenden Auswertungen beziehen sich also auf die Access-Methode als Referenzmethode.

Die Methodenvergleiche erfolgten, indem in einer ersten Auswertung alle untersuchten Patienten eingeschlossen wurden. Damit wurde der tPSA-Bereich bis 27,1 µg/l ausgewertet. Dies wird in den folgenden Abschnitten 4.3.1 bis 4.3.4 erläutert. In einer zweiten Auswertung wurde speziell eine Subgruppenanalyse der Patienten mit tPSA-Werten bis 10 µg/l vorgenommen. Damit sollten eventuelle Besonderheiten dieses PSA-Bereiches berücksichtigt werden. Dieser Aspekt wird im Abschnitt 4.3.5 beschrieben.

#### **4.3.1. Darstellung nach Passing und Bablok**

Bei der parameterfreien Regressionsanalyse nach Passing und Bablok (106) wird davon ausgegangen, dass die Ergebnisse beider Messmethoden fehlerbehaftet sind. Die klassische lineare Regressionsanalyse setzt hingegen voraus, dass der als X-Variable eingetragene Wert weitgehend fehlerfrei ist und normalverteilte Werte vorliegen. Diese Bedingungen sind bei vergleichenden Messungen, die auf ähnlichen Prinzipien beruhen und unterschiedliche Wertebereiche umfassen, nicht erfüllt, so dass von Passing und Bablok (106) für Methodenvergleiche ein parameterfreies Auswerteverfahren entwickelt wurde. Wie bei der konventionellen Regressionsanalyse werden die Ergebnisse beider Methoden gegeneinander aufgetragen und die Medianwerte, die Steigung und der Schnittpunkt der Geraden mit der Ordinate errechnet. Sind die 95%igen Konfidenzintervalle der Steigung oder des Schnittpunkts unterschiedlich von 1 bzw. 0, kann von einem signifikanten Unterschied ausgegangen werden.

Für eine erste Übersicht der Vergleichsmessungen habe ich die prozentualen Unterschiede zwischen der Access-Methode und den anderen PSA-Methoden errechnet. Hierbei wurde der jeweils mit der Access-Methode gemessene Wert 100 % gesetzt und mit den Ergebnissen der anderen Methoden verglichen. Dies ist in der Abbildung 1

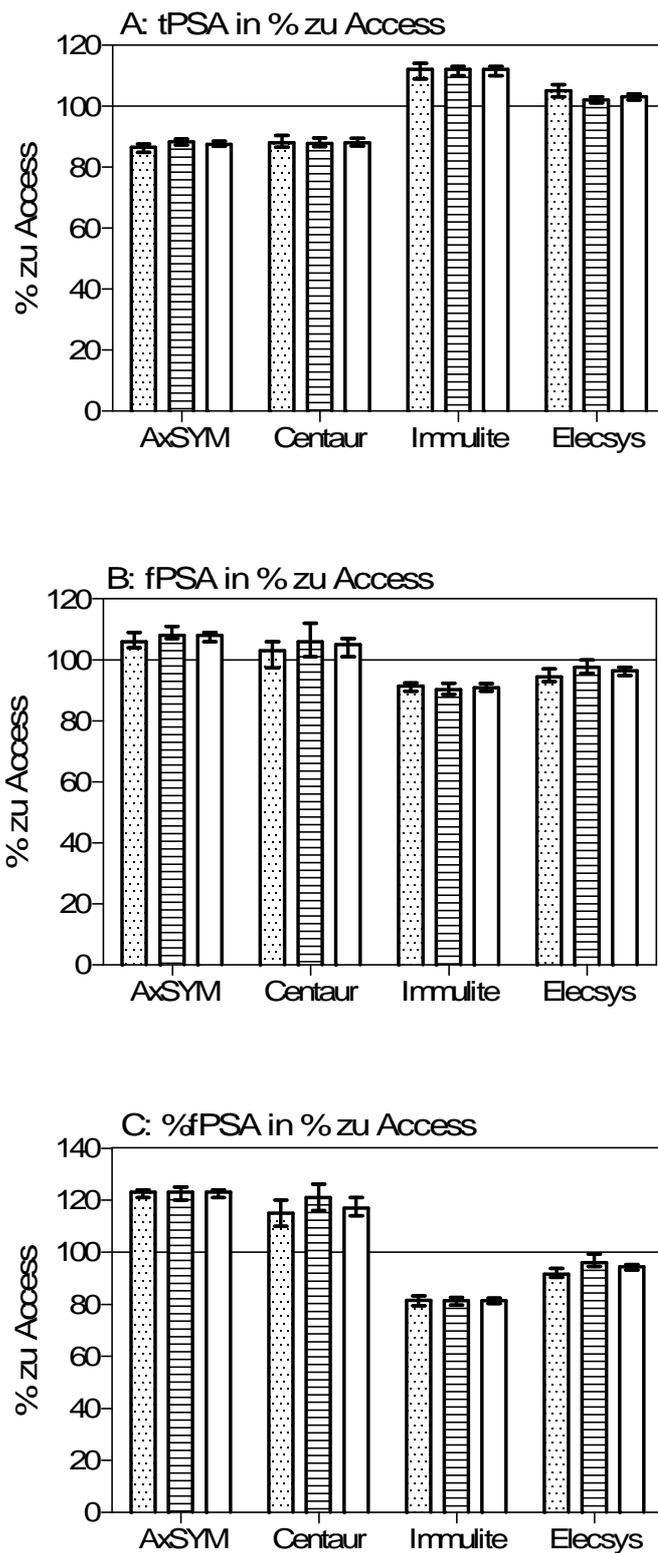
zusammengestellt. Aufgrund der prozentualen Darstellung werden die Unterschiede besser deutlich als sie im Vergleich der absoluten Werte erkennbar sind.

Die absoluten Werte sind in der Tabelle 6 zusammengefasst. Der Vergleich der Medianwerte verschiedener Methoden mit der Vergleichsmethode Access-Hybritech zeigt, dass die Messwerte zum Teil erheblich zwischen den Bestimmungsmethoden differieren. Da gemäß des Bonferroni-Postulats mit der Zahl der miteinander verglichenen Tests das Risiko wächst, Ergebnisse fälschlicherweise als signifikant zu deklarieren, entschied ich mich zur Anwendung der Bonferroni-Korrektur (105). Nach der Formel  $P < 0,05/k$  wird das Signifikanzniveau der Anzahl an Tests  $k$  angepasst. Daher sinkt das Signifikanzniveau in den folgenden Tests auf  $P < 0,0001$ . Der Wilcoxon-Test für parameterfreie, gepaarte Daten ergab, mit Ausnahme des Vergleichs zum fPSA-Centaur unter BPH-Patienten und fPSA-Elecsys innerhalb der PCa-Gruppe ( $P = 0,41$  bzw.  $P = 0,11$ ), signifikante Unterschiede zwischen allen t-, f- und %fPSA-Bestimmungen und dem Access. Insgesamt waren 78 % der Steigungen und 58 % der Schnittpunkte der Regressionsgeraden mit der Y-Achse ungleich 1 bzw. 0. Somit unterscheiden sich die untersuchten Methoden mehrheitlich signifikant zur Referenzmethode.

Die Ergebnisse lassen sich wie folgt zusammenfassen:

- fPSA- und tPSA-Konzentrationen der verschiedenen Methoden weichen in umgekehrter Richtung von der Referenzmethode ab. AxSYM und Centaur messen höhere fPSA- und niedrigere tPSA-Konzentrationen als der Access-Test. Bei Immulite und Elecsys verhält es sich umgekehrt.
- tPSA-Konzentrationen von AxSYM, Centaur und Immulite weichen im Schnitt um 9,3 (Immulite) bis 13,3 % (Centaur) von der Referenzmethode ab. Unter %fPSA-Medianwerten macht sich der Unterschied zwischen den Methoden noch deutlicher bemerkbar: Gegenüber der Referenzmethode weist AxSYM durchschnittlich 23 % höhere und Immulite 18,6 % niedrigere Werte auf.
- Mit dem Elecsys bestimmte PSA-Konzentrationen haben, mit Abweichungen von 2,3 % für tPSA-Konzentrationen und 5,3 % für %fPSA-Werte, die höchste Übereinstimmung mit dem Access-Test.
- Differenzen können sich auch innerhalb einer Methode zwischen den Erkrankungsgruppen bemerkbar machen. Steigungen des Vergleichs zwischen den

berechneten fPSA- und %fPSA-Werte des Centaur und des Access der PCa- und BPH-Patientengruppen sind signifikant verschieden. Konfidenzintervalle der Medianwerte des %fPSA- Elecsys für BPH- und PCa-Patienten überschneiden sich nicht.



**Abbildung 1.** Medianwerte der Methoden in Prozent zur Referenzmethode (Access).

Säulen: gepunktet – BPH-Patienten, gestrichelt – PCa-Patienten, leer – alle Patienten.

Aufgetragen sind die Medianwerte und die entsprechenden 95%igen Konfidenzintervalle.

**Tabelle 6.** Methodenvergleiche einzelner Tests für tPSA, fPSA %fPSA mit der Referenzmethode Access entsprechend der nicht-parametrischen Regressionsanalyse nach Passing und Bablok.<sup>a</sup>

		Access	AxSYM	Centaur	Immulite	Elecsys
<b>tPSA [<math>\mu\text{g/l}</math>]</b>						
Alle (n= 798)	Median	6,63 (6,23; 6,94)	5,68 <sup>b</sup> (5,37; 5,96)	5,56 <sup>b</sup> (5,25; 5,94)	7,20 <sup>b</sup> (6,78; 7,67)	6,63 <sup>b</sup> (6,20; 7,02)
	Steigung	1,00	0,87 <sup>c</sup> (0,86; 0,88)	0,85 <sup>c</sup> (0,84; 0,86)	1,08 <sup>c</sup> (1,06; 1,10)	0,99 (0,98; 1,00)
	Schnittpunkt	0,00	0,05 (-0,01; +0,11)	0,16 <sup>d</sup> (+0,09; +0,22)	0,17 <sup>d</sup> (+0,07; +0,27)	0,2 <sup>d</sup> (+0,13; +0,25)
BPH (n= 330)	Median	4,69 (4,11; 5,28)	4,01 <sup>b</sup> (3,41; 4,57)	4,10 <sup>b</sup> (3,44; 4,58)	5,23 <sup>b</sup> (4,34; 5,69)	4,85 <sup>b</sup> (4,13; 5,53)
	Steigung	1,00	0,84 <sup>c</sup> (0,83; 0,86)	0,84 <sup>c</sup> (0,82; 0,86)	1,07 <sup>c</sup> (1,04; 1,10)	1,00 (0,99; 1,02)
	Schnittpunkt	0,00	0,08 <sup>d</sup> (+0,02; +0,16)	0,15 <sup>d</sup> (+0,08; +0,22)	0,16 <sup>d</sup> (+0,05; +0,25)	0,17 <sup>d</sup> (+0,10; +0,23)
PCa (n= 468)	Median	7,76 (7,28; 8,17)	6,68 <sup>b</sup> (6,29; 7,05)	6,66 <sup>b</sup> (6,25; 7,18)	8,55 <sup>b</sup> (8,00; 9,23)	7,75 <sup>b</sup> (7,33; 8,29)
	Steigung	1,00	0,87 <sup>c</sup> (0,86; 0,89)	0,85 <sup>c</sup> (0,83; 0,87)	1,07 (1,04; 1,10)	0,99 (0,97; 1,01)
	Schnittpunkt	0,00	0,07 (-0,04; +0,17)	0,22 <sup>d</sup> (+0,08; +0,35)	0,29 <sup>d</sup> (+0,09; +0,48)	0,22 <sup>d</sup> (+0,10; +0,33)
<b>fPSA [<math>\mu\text{g/l}</math>]</b>						
Alle (n= 798)	Median	0,77 (0,71; 0,80)	0,82 <sup>b</sup> (0,77; 0,86)	0,77 <sup>b</sup> (0,72; 0,83)	0,67 <sup>b</sup> (0,64; 0,71)	0,74 <sup>b</sup> (0,69; 0,78)
	Steigung	1,00	1,05 <sup>c</sup> (1,02; 1,07)	1,18 <sup>c</sup> (1,13; 1,25)	0,94 <sup>c</sup> (0,92; 0,96)	0,96 <sup>c</sup> (0,94; 0,99)
	Schnittpunkt	0,00	0,02 <sup>d</sup> (+0,00; +0,04)	-0,09 <sup>d</sup> (-0,13; -0,06)	-0,02 <sup>d</sup> (-0,03; -0,01)	0,00 (-0,02; +0,01)
BPH (n= 330)	Median	0,755 (0,69; 0,82)	0,81 <sup>b</sup> (0,72; 0,88)	0,73 (0,68; 0,81)	0,65 <sup>b</sup> (0,60; 0,74)	0,72 <sup>b</sup> (0,65; 0,79)
	Steigung	1,00	1,04 <sup>c</sup> (1,00; 1,07)	1,05 (0,99; 1,12)	0,93 <sup>c</sup> (0,91; 0,96)	0,93 <sup>c</sup> (0,90; 0,96)
	Schnittpunkt	0,00	0,02 (-0,01; +0,04)	-0,02 (-0,06; +0,02)	-0,01 (-0,03; +0,00)	0,01 (-0,01; +0,03)
PCa (n= 468)	Median	0,77 (0,71; 0,82)	0,82 <sup>b</sup> (0,77; 0,87)	0,79 <sup>b</sup> (0,73; 0,86)	0,68 <sup>b</sup> (0,64; 0,72)	0,76 (0,69; 0,81)
	Steigung	1,00	1,05 <sup>c</sup> (1,02; 1,09)	1,33 <sup>c</sup> (1,24; 1,44)	0,94 <sup>c</sup> (0,92; 0,97)	1,00 (0,96; 1,03)
	Schnittpunkt	0,00	0,02 (-0,00; +0,04)	-0,18 <sup>d</sup> (-0,27; -0,11)	-0,02 <sup>d</sup> (-0,04; -0,01)	-0,01 (-0,04; +0,01)

Fortsetzung auf der nächsten Seite -

Fortsetzung von **Tabelle 6.**

		Access	AxSYM	Centaur	Immulite	Elecsys
<b>%fPSA [%]</b>						
Alle (n= 798)	Median	12,8 (12,1; 13,5)	15,9 <sup>b</sup> (15,2; 17,0)	16 <sup>b</sup> (15,1; 17,0)	10,4 <sup>b</sup> (9,81; 11,0)	12,3 <sup>b</sup> (11,8; 13,0)
	Steigung	1,00	1,24 <sup>c</sup> (1,21; 1,27)	1,32 <sup>c</sup> (1,26; 1,38)	0,82 <sup>c</sup> (0,80; 0,84)	0,90 <sup>c</sup> (0,88; 0,92)
	Schnittpunkt	0,00	-0,15 (-0,50; +0,20)	-1,84 <sup>d</sup> (-2,70; -1,10)	-0,06 (-0,34; +0,17)	0,54 <sup>d</sup> (+0,27; +0,82)
BPH (n= 330)	Median	18,3 (17,3; 19,1)	22,3 <sup>b</sup> (21,2; 23,9)	22,1 <sup>b</sup> (21,2; 24,0)	14,5 <sup>b</sup> (13,8; 15,4)	16,8 <sup>b</sup> (15,7; 17,8)
	Steigung	1,00	1,23 <sup>c</sup> (1,18; 1,28)	1,22 <sup>c</sup> (1,14; 1,31)	0,80 <sup>c</sup> (0,76; 0,84)	0,86 <sup>c</sup> (0,82; 0,90)
	Schnittpunkt	0,00	-0,09 (-0,95; +0,58)	-1,53 (-3,5; +0,07)	0,28 (-0,32; +0,87)	1,00 <sup>d</sup> (+0,33; +1,58)
PCa (n= 468)	Median	10,3 (9,58; 10,7)	12,4 <sup>b</sup> (11,9; 13,4)	12,6 <sup>b</sup> (11,9; 13,6)	8,02 <sup>b</sup> (7,55; 8,66)	10,1 <sup>b</sup> (9,44; 10,7)
	Steigung	1,00	1,26 <sup>c</sup> (1,22; 1,31)	1,61 <sup>c</sup> (1,48; 1,76)	0,85 <sup>c</sup> (0,82; 0,88)	0,98 (0,94; 1,01)
	Schnittpunkt	0,00	-0,40 (-0,83; +0,03)	-4,31 <sup>d</sup> (-5,8; -2,9)	-0,36 <sup>d</sup> (-0,7; -0,04)	-0,19 (-0,51; +0,17)

<sup>a</sup>95%ige Konfidenzintervalle in Klammern; aus Übersichtsgründen wurden Semikola statt Bindestriche verwendet. fPSA und %fPSA der Centaur-Tests wurden berechnet (siehe Kapitel 3.4).

<sup>b</sup>Nach dem Wilcoxon-Test signifikant unterschiedlich zu den Werten des Access-Tests (nach Bonferroni-Korrektur (105),  $P < 0,0001$ ).

<sup>c</sup>Steigungen der Regressionsgeraden, die einen signifikanten Unterschied zwischen den miteinander verglichenen Methoden anzeigen (K.I. ungleich 1).

<sup>d</sup>Schnittpunkte der Geraden mit der Y-Achse, die einen signifikanten Unterschied zwischen den miteinander verglichenen Methoden anzeigen (K.I. ungleich 0).

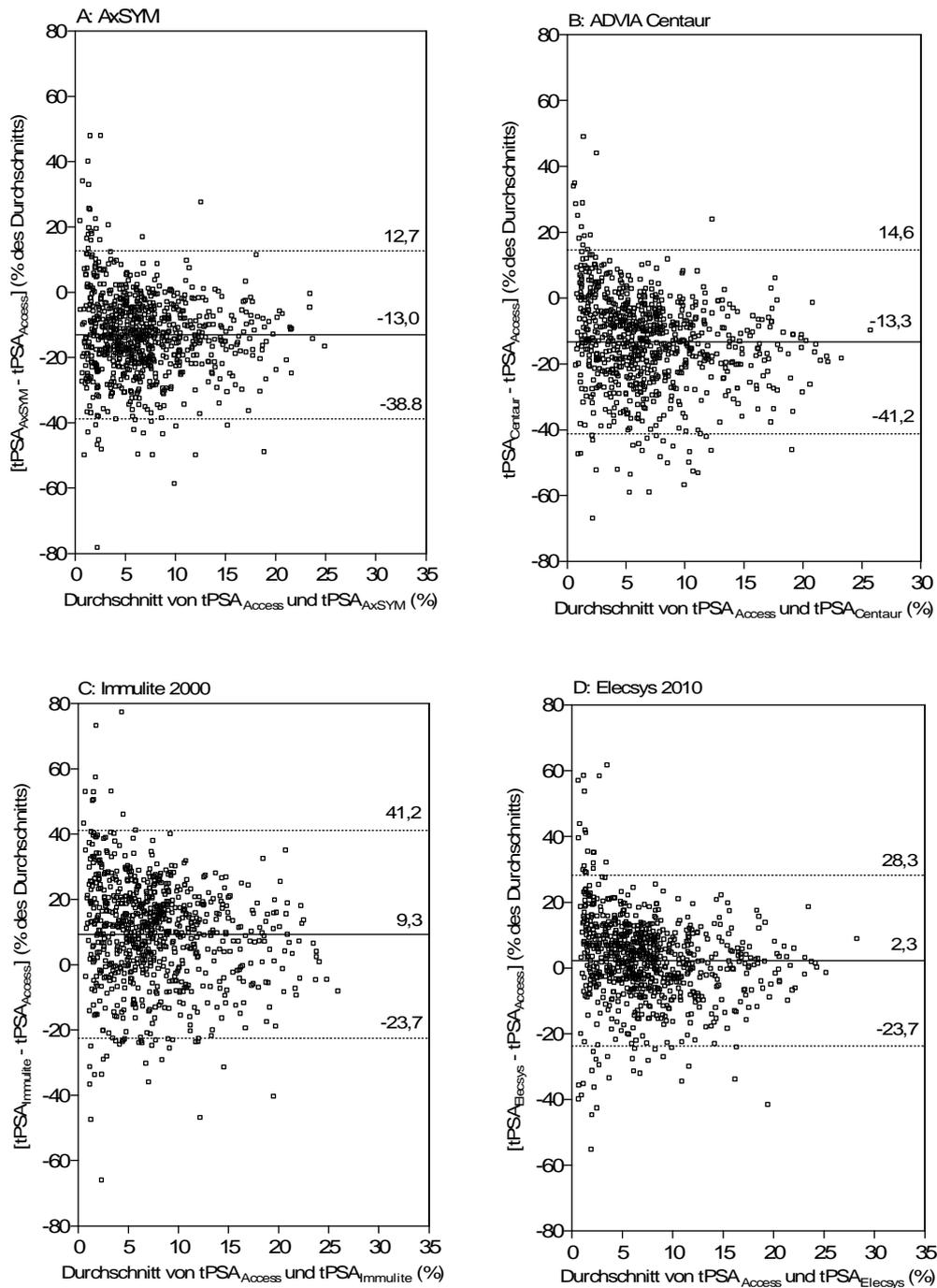
### 4.3.2. Darstellung nach Bland und Altman

Die Bland-Altman-Darstellung ergibt sich aus der Auftragung der Differenzen zwischen den Werten, die mit zwei zu vergleichenden Methoden ermittelt werden, gegenüber dem Mittelwert aus den Einzelwerten der beiden Methoden. Die Differenzen auf der Ordinate können als absolute oder prozentuale Werte angegeben werden. Ich habe entsprechend der Empfehlung von Dewitte et al. (110) die prozentuale Auftragung gewählt. In der Abbildung markieren horizontale Linien den Mittelwert der Differenzen sowie die zugehörigen 1,96fachen Standardabweichungen (SD).

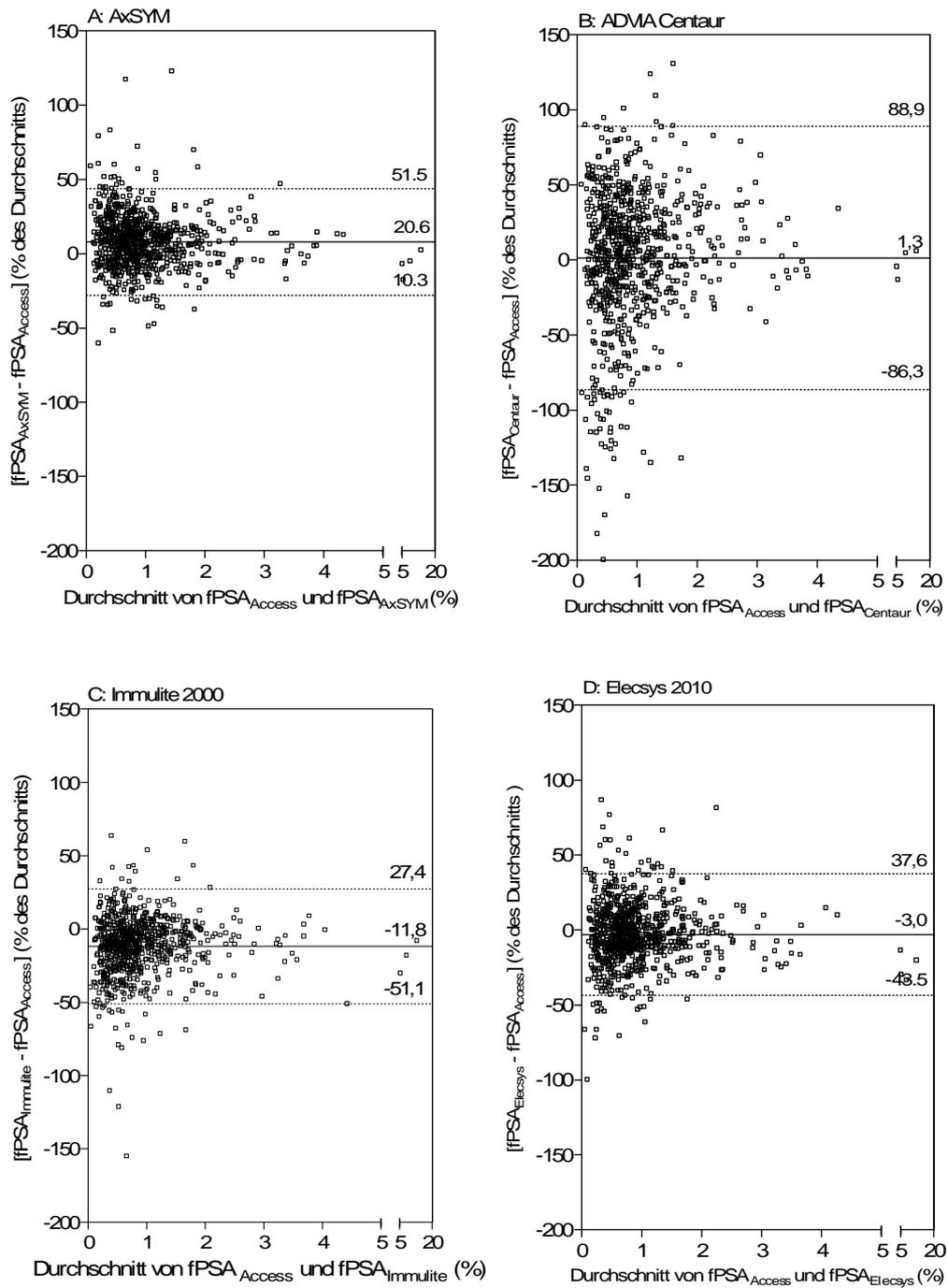
Die in Kapitel 4.3.1 beschriebenen prozentualen Abweichungen der Messergebnisse gegenüber Access können in den folgenden Abbildungen an der Höhe der durchgezogenen horizontalen Linien abgelesen werden (Abb. 2, 3 und 4).

Abbildung 2 zeigt die Vergleiche der tPSA-Methoden mit der Referenzmethode. Hinsichtlich der Streuungen zeigen sich keine großen Differenzen. Der Elecsys-Test zeigt die kleinste Abweichung vom Access.

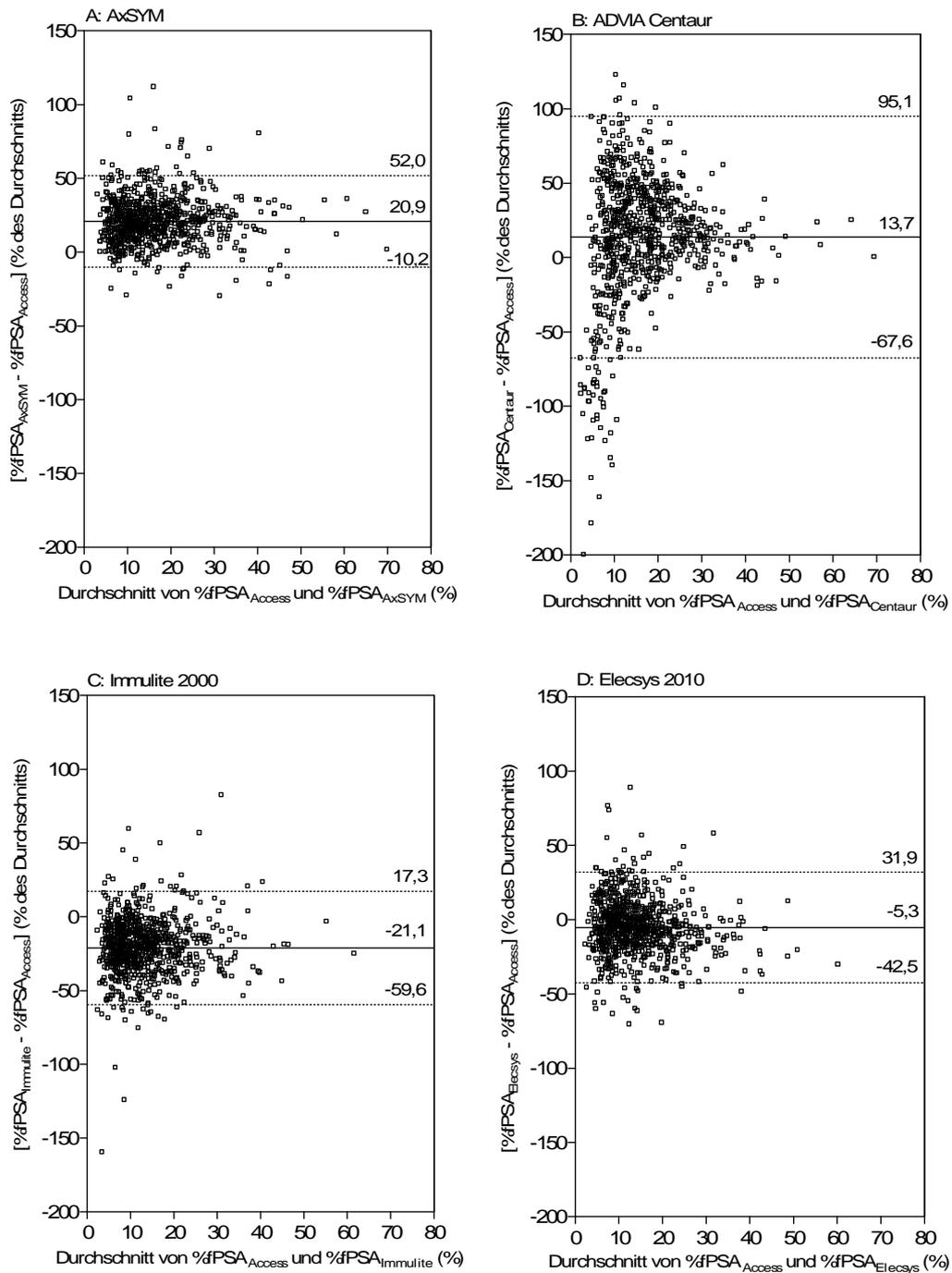
Die Abbildungen 3 und 4 zeigen Vergleiche zwischen den fPSA- und %fPSA-Werten. Bei dem Vergleich der fPSA- und %fPSA-Werte zwischen Centaur und Access wiesen 17 Proben y-Werte unter -200 bzw. über 150 auf. Bei diesen Proben wurden cPSA-Konzentrationen gemessen, die höher als die tPSA-Werte waren und bei der Berechnung von fPSA negative Werte ergaben. Diese Werte wurden bei der weiteren Berechnung nicht berücksichtigt. Trotzdem imponierten die fPSA- und %fPSA-Graphen von Centaur gegenüber den anderen Punktwolken durch ihre große Streuung.



**Abbildung 2.** Bland und Altman-Schaubilder der tPSA-Werte verschiedener Methoden. Prozentuale Unterschiede zwischen tPSA-Werten von AxSYM, Centaur, Immulite oder Elecsys und dem Access Hybritech Test wurden gegen den Mittelwert der beiden Tests entsprechend der Methode nach Bland und Altman (107) aufgetragen. Die durchgezogene Linie zeigt den Mittelwert der Differenzen, die gestrichelten Linien verdeutlichen die obere und untere 1,96fache Standardabweichung.



**Abbildung 3.** Bland und Altman-Schaubilder der fPSA-Werte verschiedener Methoden. Die durchgezogene Linie zeigt den Mittelwert der Differenzen, die gestrichelten Linien die obere und untere 1,96-fache Standardabweichung.



**Abbildung 4.** Bland und Altman-Schaubilder der %fPSA-Werte verschiedener Methoden.

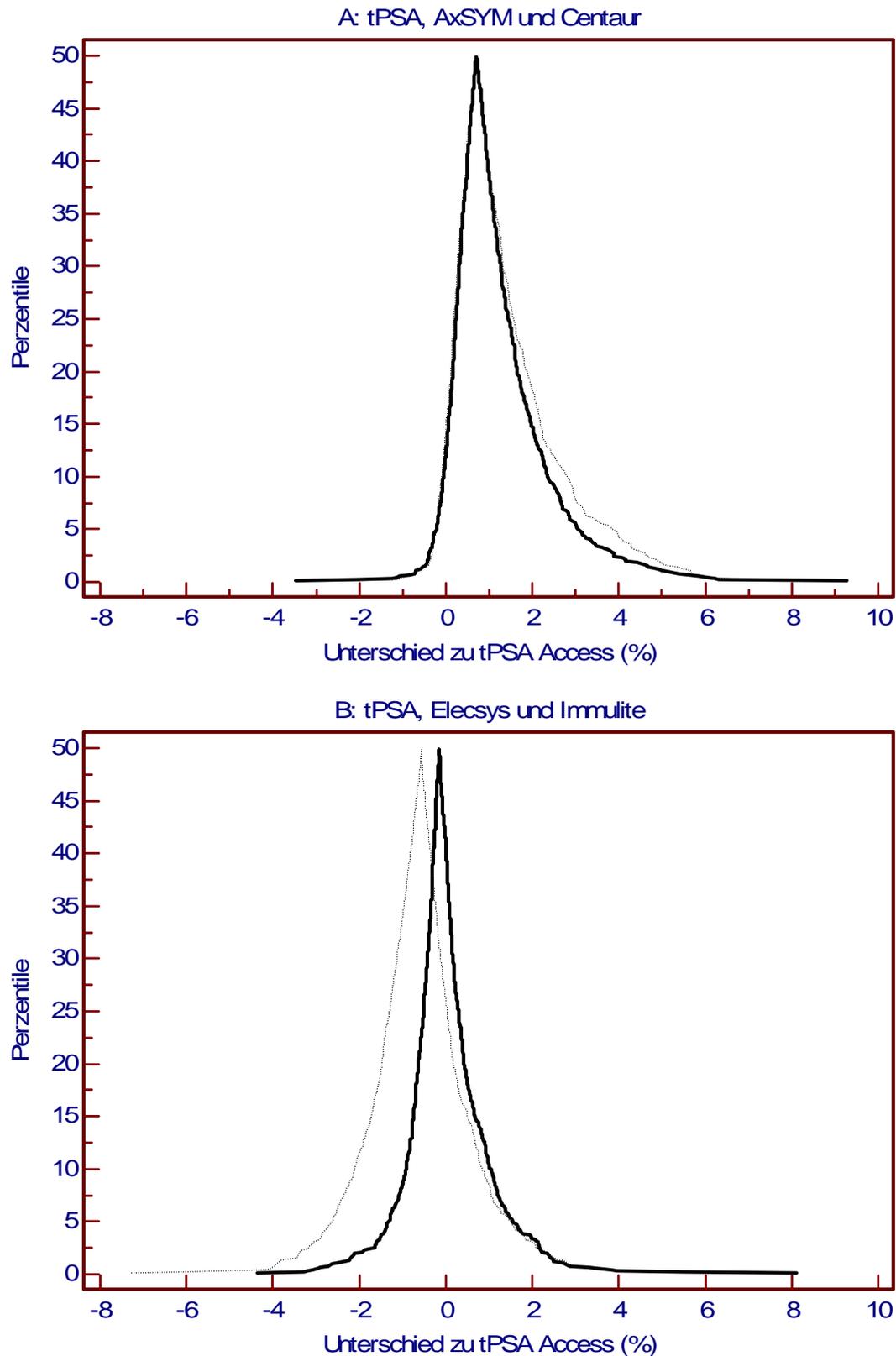
Die durchgezogene Linie zeigt den Mittelwert der Differenzen, die gestrichelten Linien die obere und untere 1,96-fache Standardabweichung.

### 4.3.3. Mountain-Plot

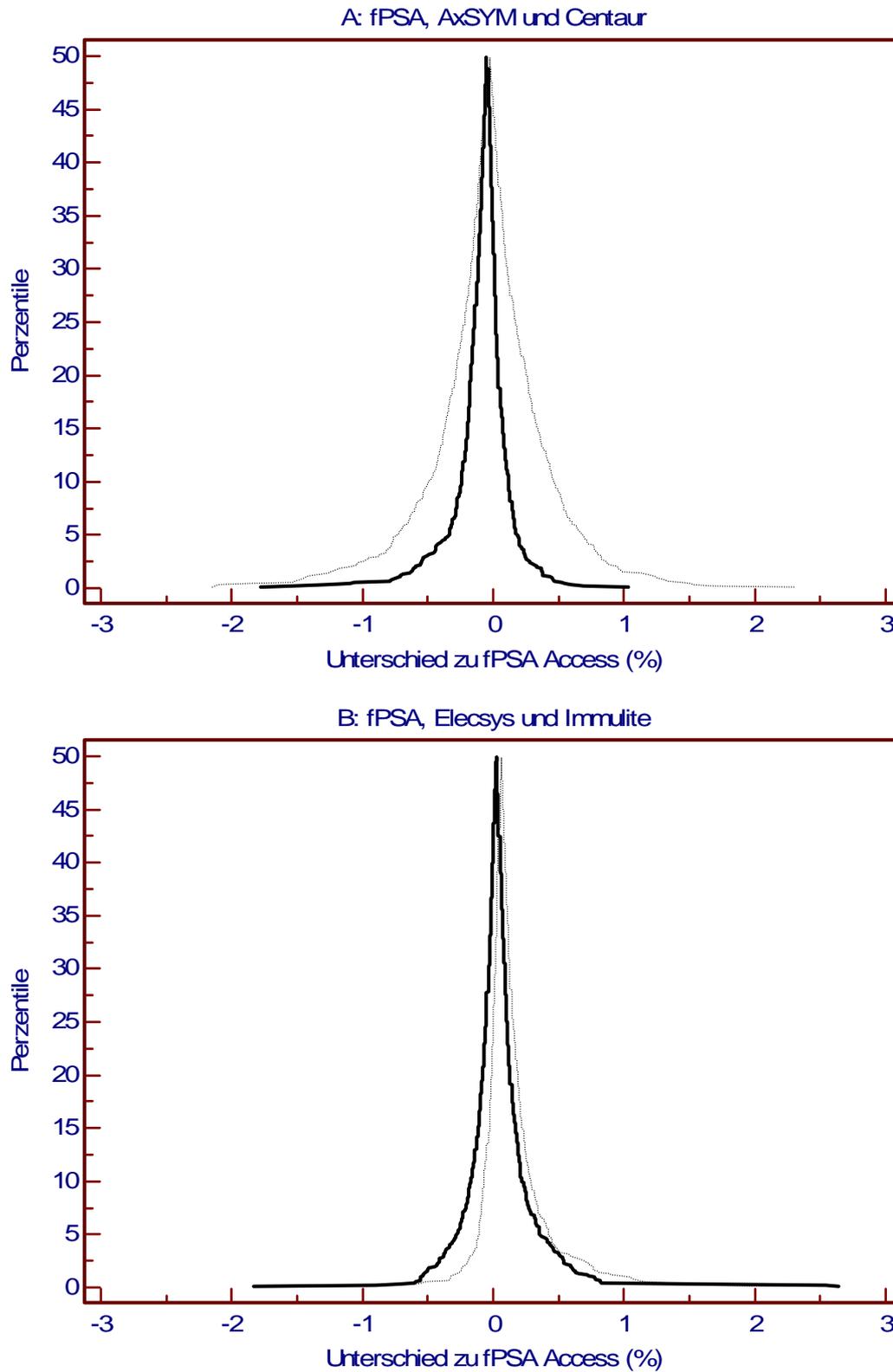
Ergänzend zu der in Kapitel 4.3.2 durchgeführten Bland-Altman-Darstellung habe ich den Mountain-Plot (folded empirical cumulative distribution plot) gewählt (108). Mit dem Mountain-Plot wird die Übereinstimmung von bis zu zwei Messmethoden mit der Referenzmethode graphisch dargestellt. Dazu werden die Differenzen der dargestellten Messmethoden zur Referenzmethode errechnet und auf der X-Achse aufgetragen. Auf der Y-Achse werden die Perzentilen der Differenzen angegeben, so dass optisch ein Berg entsteht. Je enger und spitzer der Berg ist, desto kleiner ist die Variation der Differenzen einer Messmethode, und je näher die Spitze bei 0 liegt, desto kleiner ist die gesamte Messabweichung einer Messmethode von der Referenzmethode. Im Gegensatz zu vielen anderen graphischen Methodenvergleichen können mit dem Mountain-Plot schnell seltener vorkommende große Fehler detektiert werden.

Bei den Vergleichen der tPSA-Methoden mit der Referenzmethode (Abb. 5) zeigen sich ähnliche Graphen. Die Spitze der Graphen von AxSYM und Centaur (beide +0,5 %) beweisen gegenüber dem Access-Test höher gemessene tPSA-Konzentrationen. Bei Immulite (-0,5 %) und Elecsys (-0,1 %) verhält es sich umgekehrt, wobei tPSA-Werte des Elecsys noch die größte Ähnlichkeit zum Access aufweisen. Die langen Ausläufer aller vier Graphen reflektieren die großen Unterschiede der Methoden zur Referenzmethode. Die Häufigkeit extremer Unterschiede, welche an der Breite des 95%igen Konfidenzintervalls abzuschätzen ist, stellte sich bei den Vergleichen zur Referenzmethode in folgender Reihenfolge dar: Centaur (-0,5 %; +5 %) > Immulite (-3 %; +2 %) > AxSYM (-0,5 %; +4 %) > Elecsys (-1,7 %; +2 %).

Gegenüber den tPSA-Methoden, fallen die Unterschiede der fPSA-Methoden (Abb. 6) geringer aus. Die Spitzen aller Graphen liegen näher an Null. Der Vergleich zwischen Centaur und Access zeigt die breiteste Streuung der Differenzen, was auch beim Vergleich der Konfidenzintervalle deutlich wird: Centaur (-1 %; +1 %) > AxSYM (-0,5 %; +0,5 %) > Immulite (-0,4 %; +0,5 %) > Elecsys (-0,2 %; +0,5 %).



**Abbildung 5.** Mountain-Plots für die tPSA-Methoden: A) von AxSYM (*durchgezogene Linie*) und Centaur (*gestrichelte Linie*); sowie B) von Elecsys (*durchgezogen*) und Immulite (*gestrichelt*).

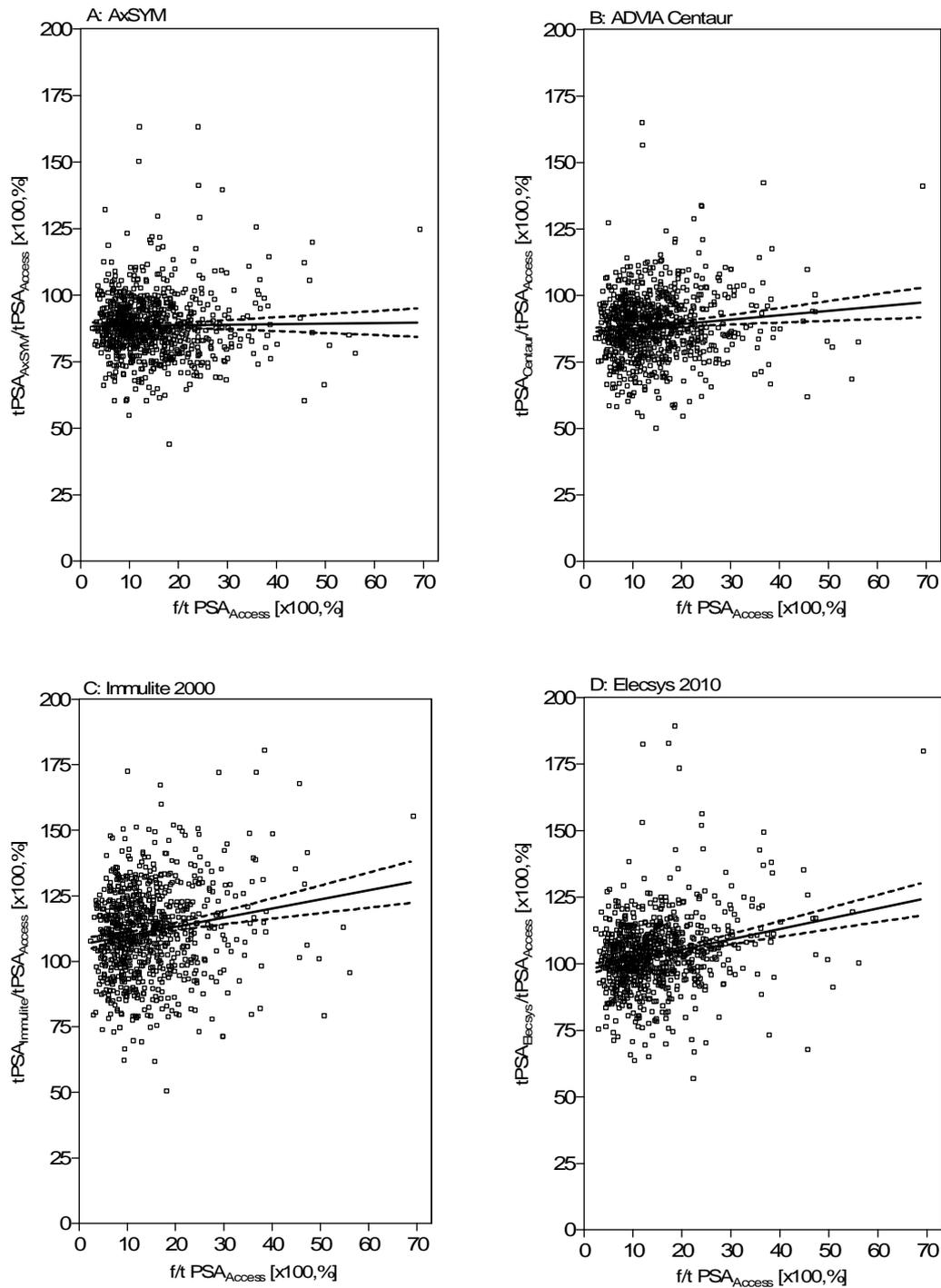


**Abbildung 6.** Mountain-Plots für die fPSA-Methoden: A) von AxSYM (*durchgezogene Linie*) und Centaur (*gestrichelte Linie*); sowie B) von Elecsys (*durchgezogen*) und Immulite (*gestrichelt*).

#### 4.3.4. Äquimolaritätsanalyse

Zur Charakterisierung der äquimolaren Eigenschaften der tPSA-Methoden wurden so genannte „Molar-Response“-Graphen erstellt (31). In diesen Schaubildern werden die prozentualen Unterschiede von tPSA-Konzentrationen des untersuchten Methoden zur Vergleichsmethode (Access) gegen die dazugehörigen %fPSA-Werte (Access) aufgetragen und eine Regressionsgerade gebildet. Bei der Analyse des Diagramms kommt der Steigung der Regressionsgeraden eine besondere Bedeutung zu. Bei Methoden mit äquimolaren Eigenschaften bleibt das Verhältnis des tPSA zur äquimolaren Referenzmethode unabhängig vom Quotienten  $f/tPSA$  gleich, so dass die Steigung der Regressionsgeraden nicht signifikant von Null abweicht. Geraden mit Steigungen ungleich Null weisen auf eine von dem Quotienten %fPSA abhängige bevorzugte Messung einer Untergruppe hin. Ein positiver Wert bedeutet eine bevorzugte Messung von fPSA, ein negativer eine Bevorzugung des gebundenen Anteils.

Abbildung 7 zeigt, dass nur mit der AxSYM-Methode freies und gebundenes PSA äquimolar erfasst wird. Im Gegensatz dazu haben die Geraden von Centaur, Immulite und Elecsys eine positive Steigung mit einer signifikanten Abweichung von Null.



**Abbildung 7.** „Molar-response“-Graphen für tPSA-Methoden.

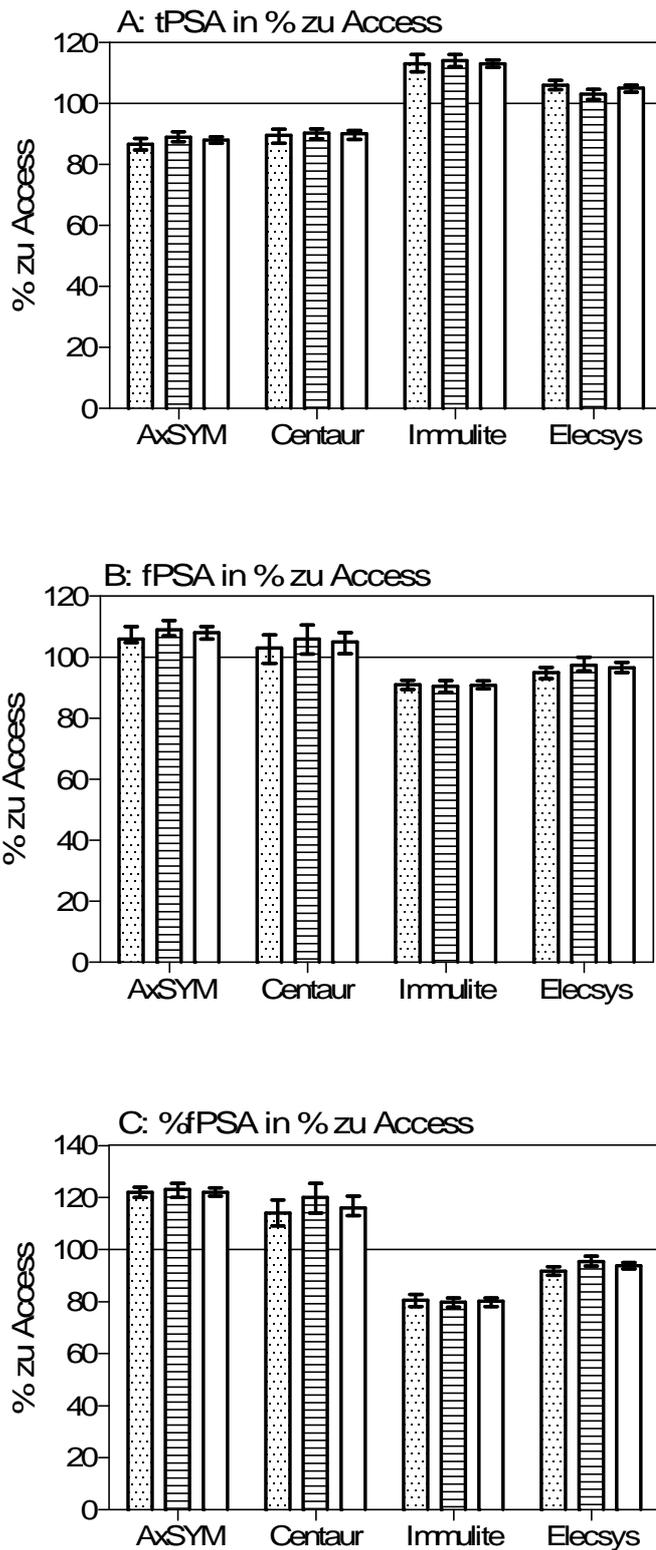
Die Gleichungen der Regressionsgeraden (95%ige Konfidenzintervalle der Schnittpunkte und Steigungen in Klammern) sowie die Abweichungen der Steigung von Null mit den dazugehörigen P-Werten sind: A) AxSYM,  $y=0,02$  (-0,08 bis 0,12)  $x + 88,2$  (86,5 bis 89,9),  $P=0,67$ ; B) Centaur,  $y=0,16$  (0,06 bis 0,27)  $x + 86,0$  (84,2 bis 87,7),  $P=0,0018$ ; C) Immulite,  $y=0,35$  (0,20 bis 0,49)  $x + 106$  (103 bis 108),  $P < 0,0001$ ; D) Elecsys,  $y=0,39$  (0,28 bis 0,50)  $x + 97,6$  (95,7 bis 99,5),  $P < 0,0001$ .

#### 4.3.5. Methodenvergleiche in der Subgruppe 0-10 µg/l

Wie bereits am Anfang des Kapitels ausgeführt, habe ich zusätzlich zu den Methodenvergleichen mit allen Patienten noch eine Subgruppenanalyse für die Patienten mit tPSA-Werten bis 10 µg/l durchgeführt. Damit sollten Aussagen für den so genannten PSA-Graubereich bis 10 µg/l erschlossen werden, der für den Urologen besonders wichtig ist.

Für die Subgruppe der Patientenproben mit einem tPSA von 0-10 µg/l ergaben die Methodenvergleiche ebenfalls große Differenzen zwischen den Methoden. Die in Abbildung 8 dargestellten prozentualen Abweichungen der PSA-Messungen zwischen den verschiedenen Methoden sind mit denen des gesamten Patientenkollektivs, wie ich sie in Abbildung 1 gezeigt hatte, nahezu identisch. fPSA- und tPSA-Werte weichen in umgekehrter Richtung von der Referenzmethode ab.

Berechnungen nach Passing und Bablok (nicht dargestellt) zeigten deutliche Unterschiede der Steigungen der Regressionsgeraden und der Schnittpunkte der Geraden mit der Y-Achse: 95%ige Konfidenzintervalle der Steigungen und Schnittpunkte waren zu 83 % bzw. 47 % ungleich 1 bzw. 0. Wie bei der Analyse des gesamten Patientenkollektivs zeigten auch hier die Medianwerte der einzelnen Methoden deutliche Abweichungen zum Access-Test. Medianwerte der tPSA-Konzentrationen aller 596 Patienten reichten bei von 4,6 µg/l (AxSYM) bis 5,7 µg/l (Immulite), %fPSA-Medianwerte von 11,4 % (Immulite) bis 17,9 % (AxSYM). Mit Ausnahme der fPSA-Centaur-Bestimmung unter BPH-Patienten und fPSA-Elecsys bei PCa-Patienten ( $P= 0,32$  bzw.  $P= 0,13$ ), waren nach dem Wilcoxon-Test alle Medianwerte der t-, f- und %fPSA-Bestimmungen signifikant unterschiedlich zum Access-Test ( $P < 0,0001$ ). In dieser Subgruppe konnte ebenfalls beobachtet werden, dass sich auch innerhalb einer Methode zwischen den Erkrankungsgruppen Differenzen bemerkbar machen können. Bei den Regressionsgeraden der fPSA- und %fPSA-Vergleiche mit dem Centaur sind Steigungen zwischen PCa- und BPH-Patienten signifikant verschieden, Schnittpunkte der Geraden mit der Ordinate weichen nur in der PCa-Gruppe von 0 ab. Die Konfidenzintervalle der Medianwerte des %fPSA-Elecsys für BPH- und PCa-Patienten überschneiden sich nicht.



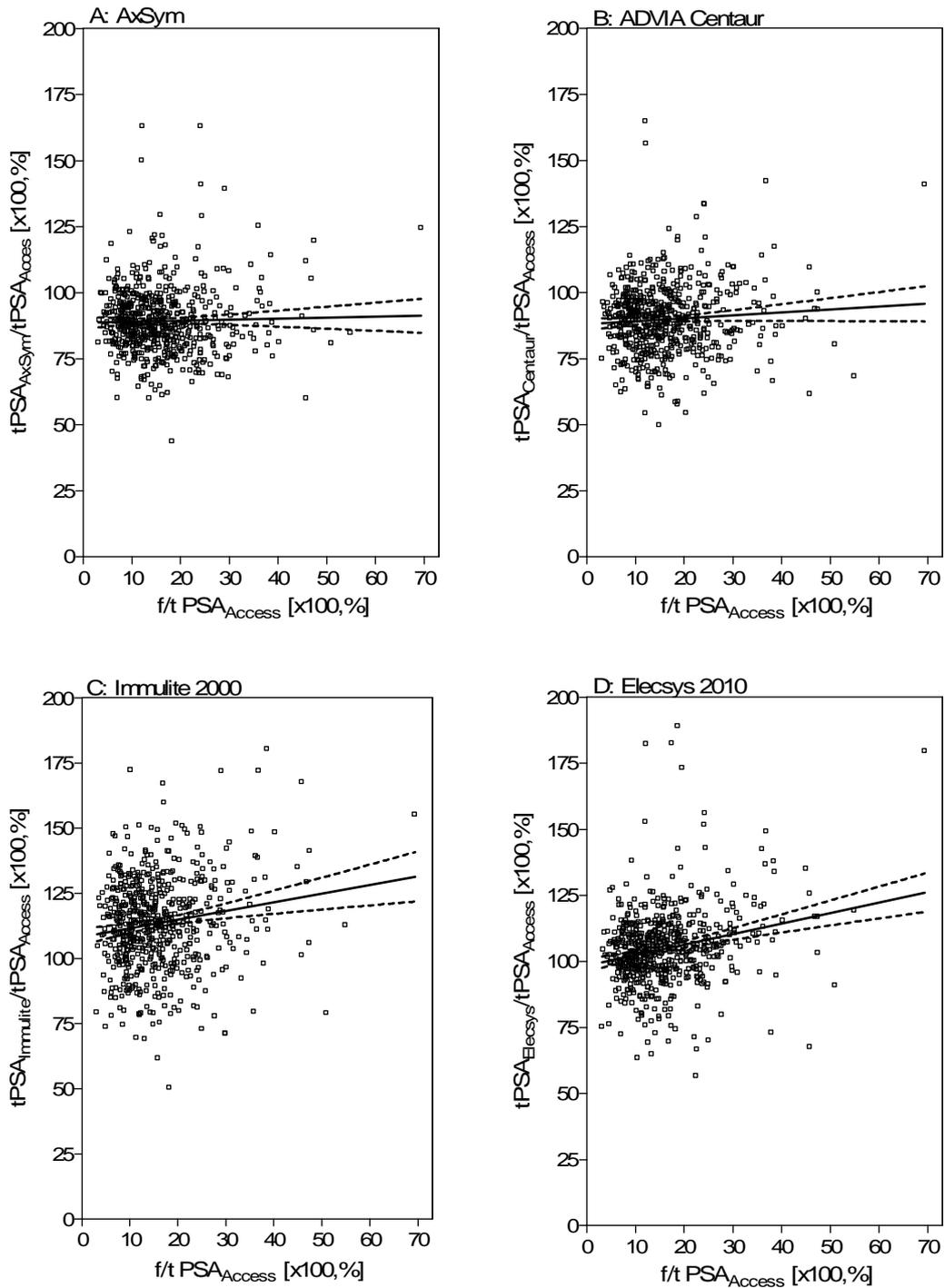
**Abbildung 8.** Medianwerte der Methoden in Prozent zur Referenzmethode (Access) der Subgruppe 0-10 µg/l.

Säulenformat: gepunktet – BPH-Patienten, gestrichelt – PCa-Patienten, leer – alle Patienten. Aufgetragen sind die Medianwerte und entsprechende 95%ige Konfidenzintervalle.

Bland-Altman-Schaubilder dieser Subgruppe (nicht dargestellt) unterscheiden sich ebenfalls kaum von denen des gesamten Patientenkollektivs (Abb. 2, 3 und 4). Bei den tPSA-Methoden ergeben sich hinsichtlich der 1,96fachen Standardabweichungen keine großen Unterschiede zwischen den Methoden. Medianwerte des tPSA-Elecsys-Tests weisen die geringsten Unterschiede zum Access auf. Wie bei der Analyse der gesamten Proben wurden bei den fPSA- und %fPSA-Centaur-Graphen 17 Ausreißer mit Y-Werten unter -200 oder über 150 und errechneten negativen fPSA-Werten nicht berücksichtigt. Dennoch macht sich auch hier bei dem Centaur-Test eine große Streuung der Werte bemerkbar. Bei %fPSA-Vergleichen sind obere und untere 1,96fache Standardabweichungen des Centaur (+88,6 %; -64,3 %) deutlich größer als bei AxSYM (+51,5 %; -10,3 %), Immulite (+14,2 %; -59,4 %) und Elecsys (+30,4 %; -43,8 %).

Mountain-Plots (nicht dargestellt) aller fPSA- und tPSA-Methoden zeigten etwas geringere Unterschiede zur Referenzmethode, als bei der Analyse des gesamten Patientenkollektivs beobachtet wurden. Bei allen Vergleichen waren die „Berge“ etwas schmaler und hatten kürzere Ausläufer aber gleich abweichende Spitzen. Die Konfidenzintervalle stellten sich folgendermaßen dar: Centaur (-0,5 %; +3 %) und Immulite (-2,5 %; +1 %) > AxSYM (-0,5 %; +2,5 %) > Elecsys (-1,2 %; +1 %) für tPSA; und Centaur (-0,7 %; +0,7 %) > AxSYM (-0,3 %; +0,3 %) > Elecsys (-0,2 %; +0,2 %) > Immulite (-0,1 %; +0,25 %) für fPSA.

Die so genannten „Molar-Response“-Graphen (Abb. 9) zeigen, dass in der Subgruppe 0-10 µg/l AxSYM und Centaur freies und gebundenes PSA äquimolar messen, wohingegen tPSA-Methoden von Immulite und Elecsys nicht-äquimolare Eigenschaften aufwiesen. In einem weiteren Schritt habe ich in dieser Subgruppe nur Proben mit einem %fPSA <25 % analysiert (nicht dargestellt). Innerhalb dieses Bereichs wurde die Nichtäquimolarität der Immulite- und die Elecsys-Methode nicht mehr bemerkt. Bei diesen Methoden zeigten die Steigungen der Regressionsgeraden keine signifikante Abweichungen von Null (P= 0,11 und P= 0,10 für Immulite und Elecsys).



**Abbildung 9.** „Molar-Response“-Graphen für tPSA-Methoden der Subgruppe 0-10 µg/l. Die Gleichungen der Regressionsgeraden (95%ige Konfidenzintervalle der Schnittpunkte und Steigungen in Klammern) sowie die Abweichungen der Steigung von Null mit den dazugehörigen P-Werten sind: A) AxSYM,  $y=0.04 (-0.08-0.15) x + 88.7 (86.5-90.9)$ ,  $P=0.55$ ; B) Centaur,  $y=0.11 (-0.01-0.24) x + 88.0 (85.7-90.2)$ ,  $P=0.08$ ; C) Immulite,  $y=0.33 (0.16-0.51) x + 108 (105-111)$ ,  $P=0.0002$ ; D) Elecsys,  $y=0.40 (0.26-0.53) x + 98.4 (95.9-100)$ ,  $P < 0.0001$ .

#### 4.4. Bewertung der klinischen Aussagekraft

##### 4.4.1. Die ROC-Analyse

Die Receiver-operation-characteristics (ROC)-Analyse, dessen Anwendung ausführlich von Zweig und Campbell ([109](#)) beschrieben wurde, gilt heutzutage als Standardmethode, um Laborteste oder andere diagnostische Messverfahren hinsichtlich ihrer klinischen Validität zu vergleichen ([111](#)).

Die ROC-Kurve stellt eine Vielzahl an Sensitivitäts- und Spezifitätsberechnungen unabhängig von den Absolutwerten der bestimmten Parameter dar und erlaubt eine umfassende Beurteilung über die Diskriminationsfähigkeit diagnostischer Tests, zwischen erkrankt und nicht erkrankt zu unterscheiden. Dazu werden alle ermittelten Sensitivitäten und Spezifitäten in ein Koordinatensystem mit 1 - Spezifität als Abszisse und Sensitivität als Ordinate eingetragen und anschließend zu einer Kurve miteinander verbunden. Die Fläche unter der Kurve („area under curve“) beinhaltet alle Sensitivitäten und Spezifitäten und wird als ein zusammenfassendes Maß für die diagnostische Aussagekraft eines Tests herangezogen. Eine Fläche von 0,5 gilt als nicht aussagekräftig. Je näher sich die Fläche 1 nähert, desto größer wird die diagnostische Aussagekraft eines Tests.

Die in den Abbildungen 10 und 11 dargestellten ROC-Kurven für tPSA und %fPSA zeigen insgesamt eine große Übereinstimmung. Nur für %fPSA-Centaur besteht bei beiden untersuchten Patientenkollektiven eine Einbuchtung zu Beginn der Kurve (Abb. 11), was einer geringeren Spezifität bei niedrig gewählter Sensitivität entspricht.

Die errechneten Flächen unter den Kurven der verschiedenen Bestimmungsmethoden, die für alle Patienten und für die Untergruppe 0-10 µg/l ermittelt wurden, sind in Tabelle 7 dargestellt. Wegen der Anzahl der miteinander verglichenen Tests und dem damit verbundenen Risiko, Ergebnisse fälschlicherweise als signifikant zu deklarieren, habe ich in den folgenden Untersuchungen das Signifikanzniveau entsprechend der Bonferroni-Korrektur ([105](#)) auf  $P < 0,0001$  gesenkt. Unabhängig von der verwendeten PSA-Methodik und der untersuchten Untergruppe sind (wie zu erwarten) die Flächen

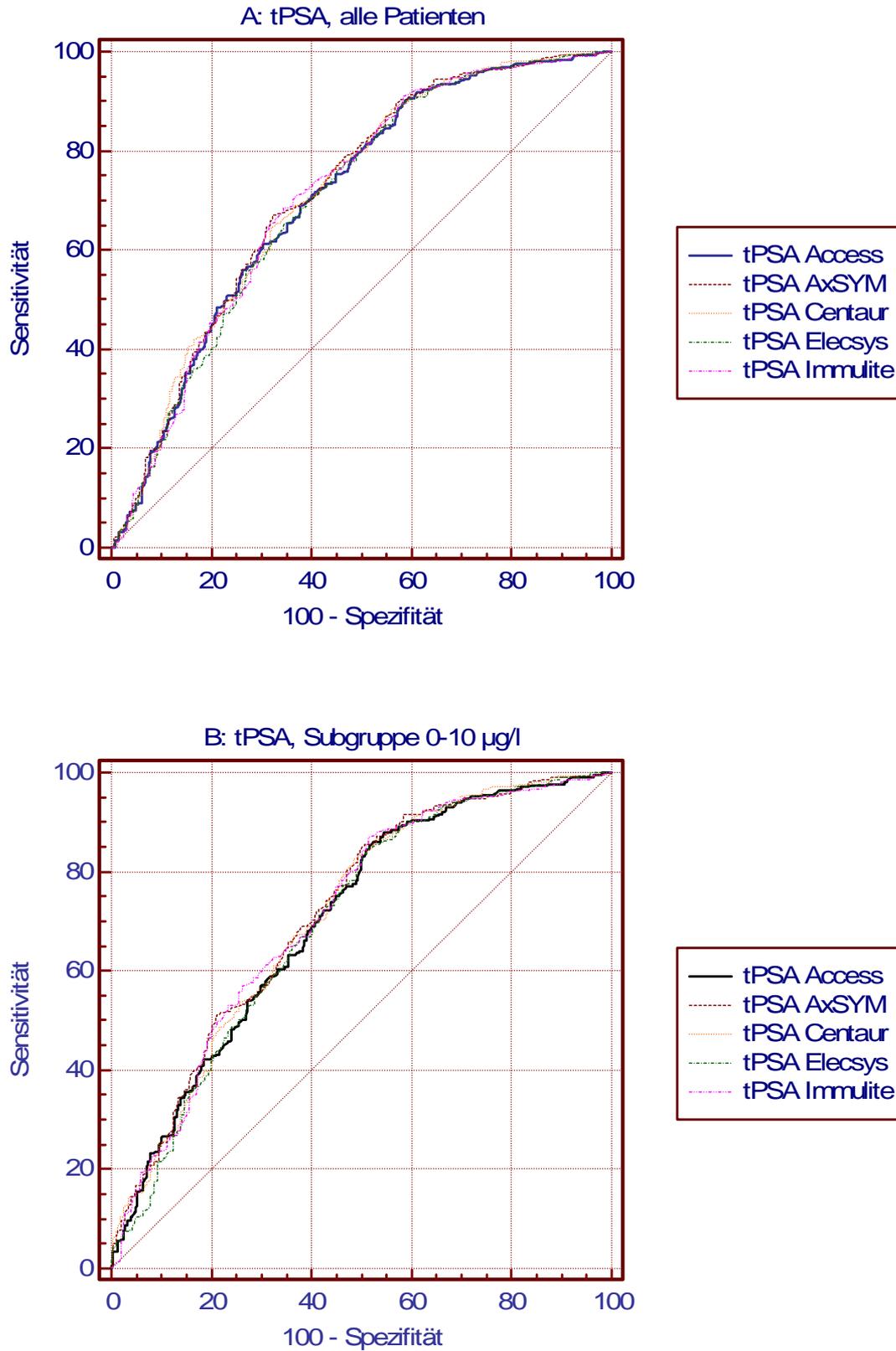
von %fPSA am größten und fPSA mit Flächen von circa 0,5 (nicht dargestellt) für die Diskriminierung von PCa und BPH nicht geeignet. Unter den tPSA-Methoden ist AxSYM mit einer Fläche von 0,72 am größten und Elecsys mit einer Fläche von 0,70 am kleinsten. Auch wenn nach der Bonferroni-Korrektur nicht signifikant, ist der Unterschied zwischen diesen beiden Flächen am größten ( $P= 0,001$ ). In der Subgruppe bestehen keine signifikanten Unterschiede zwischen den Flächen der verschiedenen tPSA-Methoden ( $P= 0,09-0,96$ ). Bei den %fPSA-Methoden nehmen die Flächen in der Reihenfolge Centaur, Elecsys, AxSYM, Immulite, Access zu, mit signifikanten Unterschieden zwischen Centaur und allen anderen Methoden sowie Unterschieden zwischen Elecsys und Access. In der Subgruppe betragen die Flächen zwischen 0,77 (Centaur) und 0,81 (Access). Bis auf den Vergleich zu Elecsys waren alle Flächen signifikant unterschiedlich zu Centaur. Mit einer Fläche von 0,79 ist die Fläche von Elecsys signifikant kleiner als die Fläche von Access (0,81).

Um innerhalb der ROC-Analyse relevante Bereiche detaillierter zu beschreiben und die klinische Validität weiter zu evaluieren, können Subgruppen gebildet, Diskriminationspunkte oder häufig benutzte Kriterien wie die 95%ige Spezifität bzw. die 95%ige Sensitivität berechnet und miteinander verglichen werden. Im folgenden Schritt habe ich daher Spezifitäten und Grenzwerte der einzelnen Methoden zu festgelegten 90%- und 95%igen Sensitivitäten berechnet und miteinander verglichen (Tab. 8). Um die Testbedingungen besser den klinischen Gegebenheiten anzupassen, wurde die Berechnung der Diskriminationspunkte für die verschiedenen tPSA-Methoden unter Einschluss aller Patienten vorgenommen (Tab. 8a) und anschließend %fPSA-Methoden in dem dafür vorgesehenen Bereich 0-10  $\mu\text{g/l}$  evaluiert (Tab. 8b).

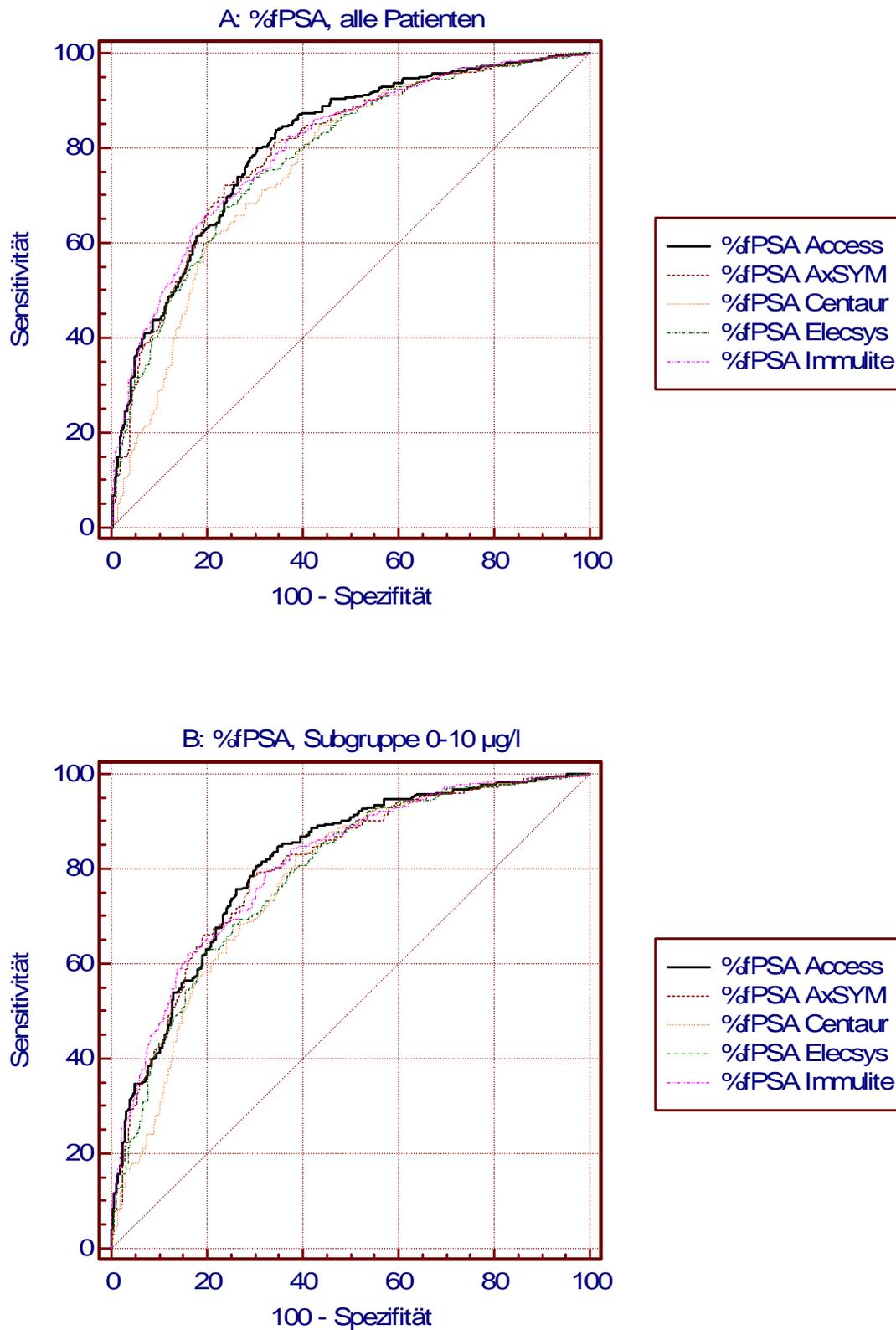
Ein Vergleich der tPSA-Methoden (Tab. 8a) zeigt komplett unterschiedliche Grenzwerte bei festgelegten Sensitivitäten: die tPSA-Diskriminationspunkte reichen von 3,11  $\mu\text{g/l}$  (Centaur) bis 3,95  $\mu\text{g/l}$  (Immulite) bei 90%iger Sensitivität und von 2,21 (AxSYM) bis 2,79  $\mu\text{g/l}$  (Immulite) bei 95%iger Sensitivität. Die Spezifitäten der verschiedenen Methoden bei 90- und 95%iger Sensitivität waren nicht signifikant unterschiedlich ( $P= 0,13-1,0$ ).

Auch unter den %fPSA-Methoden (Tab. 8b) resultierten bei einem Vergleich der Validitäten zu festgelegter 90%iger und 95%iger Sensitivität absolut unterschiedliche Werte.

Grenzwerte schwankten zwischen 15,4 % (Immulite) und 24,2 % (AxSYM) bei fixer 90%iger Sensitivität und zwischen 18,1 % (Immulite) und 27,9 % (AxSYM) bei 95%iger Sensitivität. Bei festgelegter 90%iger Sensitivität hatte der Access-Test eine signifikant höhere Spezifität als AxSYM, Centaur und Immulite. Die anderen Vergleiche zeigten keine signifikanten Unterschiede. Bei 95%iger Sensitivität gab es wenige Unterschiede zwischen den Spezifitäten der einzelnen Methoden, bis auf die signifikant besseren Spezifitäten von Access und AxSYM gegenüber dem Elecsys.



**Abbildung 10.** ROC-Kurven für tPSA-Methoden mit Proben aller Patienten (A) sowie der Subgruppe 0-10 µg/l (B).



**Abbildung 11.** ROC-Kurven für %fPSA-Methoden mit Proben aller Patienten (A) sowie der Subgruppe 0-10 µg/l (B).

**Tabelle 7.** Berechnete Flächen der ROC-Kurven aller Patientengruppen (95%ige Konfidenzintervalle in Klammern; Patientengruppen wurden nach tPSA-Access gebildet).

	Access	AxSYM	Centaur	Immulite	Elecsys
<b>tPSA</b>					
alle	0,71 (0,67 - 0,74)	0,72 (0,68 - 0,75)	0,71 (0,68 - 0,75)	0,71 (0,68 - 0,74)	0,70 (0,67 - 0,73)
0-10 µg/l	0,70 (0,67 - 0,74)	0,72 (0,68 - 0,75)	0,71 (0,67 - 0,75)	0,71 (0,68 - 0,75)	0,70 (0,66 - 0,74)
<b>%fPSA</b>					
alle	0,81 <sup>c,e</sup> (0,78 - 0,84)	0,80 <sup>c</sup> (0,77 - 0,82)	0,76 <sup>a,b,d,e</sup> (0,73 - 0,79)	0,80 <sup>c</sup> (0,77 - 0,83)	0,78 <sup>a,c</sup> (0,75 - 0,81)
0-10 µg/l	0,81 <sup>c,e</sup> (0,78 - 0,84)	0,80 <sup>c</sup> (0,77 - 0,83)	0,77 <sup>a,b,d</sup> (0,74 - 0,81)	0,81 <sup>c</sup> (0,77 - 0,84)	0,79 <sup>a</sup> (0,75 - 0,82)

Nach Bonferroni-Korrektur (105) zeigt der paarweise Vergleich signifikante Unterschiede ( $P < 0,0001$ ) zu : <sup>a</sup>Access, <sup>b</sup>AxSYM, <sup>c</sup>Centaur, <sup>d</sup>Immulite, <sup>e</sup>Elecsys.

**Tabelle 8a.** Spezifitäten und Diskriminationspunkte (Cutoffs) bei festgelegten 90%- und 95%igen Sensitivitäten bei tPSA-Methoden (95%ige Konfidenzintervalle in Klammern).

Methode	90% Sensitivität		95% Sensitivität	
	Cutoff [ $\mu\text{g/l}$ ]	Spezifität [%]	Cutoff [ $\mu\text{g/l}$ ]	Spezifität [%]
Access	3,46 (2,98 – 3,76)	41,2 (35,9 – 46,7)	2,45 (1,97 – 2,91)	28,2 (23,4 – 33,4)
AxSYM	3,18 (2,72 – 3,41)	42,4 (37,0 – 48,0)	2,21 (1,77 – 2,70)	31,5 (26,5 – 36,8)
Centaur	3,11 (2,61 – 3,44)	40,9 (35,6 – 46,6)	2,25 (1,91 – 2,60)	28,8 (24,0 – 34,0)
Immulite	3,95 (3,36 – 4,23)	41,8 (36,4 – 47,7)	2,79 (2,20 – 3,24)	29,7 (24,8 – 34,9)
Elecsys	3,68 (3,19 – 4,0)	41,2 (35,9 – 46,7)	2,54 (2,07 – 3,07)	29,4 (24,5 – 34,6)

**Tabelle 8b.** Spezifitäten und Diskriminationspunkte (Cutoffs) bei festgelegten 90%- und 95%igen Sensitivitäten bei %fPSA-Methoden in der Subgruppe tPSA 0-10  $\mu\text{g/l}$  (95%ige Konfidenzintervalle in Klammern).

Methode	90% Sensitivität		95% Sensitivität	
	Cutoff [%]	Spezifität [%]	Cutoff [%]	Spezifität [%]
Access	18,7 (17,2 – 20,0)	52,1 <sup>b,c,d</sup> (47,1-57,2)	21,1 (19,4 – 24,4)	37,2 <sup>e</sup> (32,5 – 42,2)
AxSYM	24,2 (22,0 – 25,7)	46,8 <sup>a</sup> (41,8 – 51,9)	27,9 (25,1 – 31,7)	35,5 <sup>e</sup> (30,7– 40,4)
Centaur	24,1 (21,8 – 25,8)	47,5 <sup>a</sup> (42,5 – 52,6)	27,6 (25,1 – 29,6)	35,1 (30,4 – 40,0)
Immulite	15,4 (14,5 – 17,1)	48,6 <sup>a</sup> (43,5 – 53,6)	18,1 (16,7 – 19,7)	34,0 (29,4 – 39,0)
Elecsys	17,8 (17,0 – 19,1)	48,9 (43,9 – 54,0)	20,9 (18,7 – 21,7)	31,9 <sup>a,b</sup> (27,3 – 36,8)

Nach Bonferroni-Korrektur (105) zeigt der paarweise Vergleich signifikante Unterschiede ( $P < 0,0001$ ) zu : <sup>a</sup>Access, <sup>b</sup>AxSYM, <sup>c</sup>Centaur, <sup>d</sup>Immulite, <sup>e</sup>Elecsys.

#### 4.4.2. Diskriminierung anhand festgelegter Grenzwerte

Die Daten in den Tabellen 9a und 9b verdeutlichen, dass die Wahl fixer Diskriminationspunkte für tPSA (2,5 bzw. 4 µg/l) und %fPSA (15, 20 bzw. 25 %) zur unterschiedlichen Einordnung der Patienten in die jeweiligen Gruppen „Prostatakarzinom“ bzw. „BPH“ führen. Entsprechend den klinischen Gegebenheiten habe ich mich auch bei dieser Analyse dazu entschieden, tPSA-Grenzwerte innerhalb des gesamten Patientenkollektivs (Tab. 9a) zu untersuchen und %fPSA-Methoden zur Differenzierung der Patienten in der Subgruppe tPSA 0-10 µg/l (Tab. 9b) vorzunehmen.

Bei der Verwendung fester Diskriminationspunkte für tPSA (Tab. 9a) führen die verschiedenen Methoden zu signifikant unterschiedlichen Differenzierungen von PCa und BPH. Hätte man für das gesamte Patientenkollektiv den in der Praxis gebräuchlichen tPSA-Wert von 4 µg/l als alleiniges Entscheidungskriterium für das Prostatakarzinom genommen, wäre je nach benutzter Methode eine unterschiedliche Anzahl an Patienten falsch-negativ bzw. falsch-positiv klassifiziert worden. Wäre beispielsweise der Test an Immulite oder Centaur vorgenommen worden, wären 46 oder 93 der insgesamt 486 Prostatakarzinompatienten nicht richtig erkannt worden, während 192 bzw. 168 BPH-Patienten falsch-positiv eingestuft worden wären.

Beim Gebrauch fester Grenzwerte für %fPSA (Tab. 9b) unterscheidet sich ebenfalls der Anteil der fälschlicherweise als positiv bzw. negativ klassifizierten BPH- bzw. PCa-Patienten signifikant. Bei einem Grenzwert von 15 % würden beispielsweise zwischen 65 (AxSYM) und 179 (Immulite) Männer unnötigerweise biopsiert werden. Im Gegenzug würden mit den Immulite-Werten fast 100 PCa-Patienten weniger übersehen als beim AxSYM und Centaur. Umgerechnet würde dies bedeuten, dass der Immulite-Test bei einem Grenzwert von 15 % eine 30 % höhere Sensitivität und 31-34 % niedrigere Spezifität als zwei andere Methoden aufweist. Diese Unterschiede treten in geringerem Maße auch bei höheren Grenzwerten auf.

**Tabelle 9a.** Einordnung der Patienten in Erkrankungsgruppen anhand festgelegter Grenzwerte für tPSA (2,5 bzw. 4 µg/l).

	Access	AxSYM	Centaur	Immulite	Elecsys
Ausgewählter Grenzwert	Anzahl der eingestuften Patienten				
<b>tPSA</b>					
BPH >4 µg/l <sup>1</sup>	187 (56,6 %) <sup>b,c</sup>	165 (50 %) <sup>a,d,e</sup>	168 (50,9 %) <sup>a,d,e</sup>	192 (58,2 %) <sup>b,c</sup>	188 (57,0 %) <sup>a,c</sup>
BPH >2,5 µg/l	234 (70,9 %) <sup>b,c</sup>	214 (64,8 %) <sup>a,d,e</sup>	217 (65,8 %) <sup>a,d,e</sup>	241 (73,0 %) <sup>b,c</sup>	234 (70,9 %) <sup>b,c</sup>
PCa <4 µg/l <sup>2</sup>	68 (14,5 %) <sup>b,c,d</sup>	86 (18,4 %) <sup>a,d,e</sup>	93 (19,9 %) <sup>a,d,e</sup>	46 (9,83 %) <sup>a,b,c,e</sup>	59 (12,6 %) <sup>b,c,d</sup>
PCa <2,5 µg/l	26 (5,56 %) <sup>d</sup>	26 (5,56 %) <sup>d</sup>	31 (6,62 %) <sup>d,e</sup>	17 (3,63 %) <sup>a,b,c</sup>	22 (4,70%) <sup>c</sup>

<sup>1</sup>Anzahl der 330 BPH-Patienten mit einer tPSA-Konzentration >4 bzw. 2,5 µg/l.

<sup>2</sup>Anzahl der 468 PCa-Patienten mit einer tPSA-Konzentration <4 bzw. 2,5 µg/l.

Nach dem McNemar-Test für gepaarte Daten (nach Bonferroni-Korrektur (105), P <0,0001) waren die einzelnen Test signifikant unterschiedlich zu: <sup>a</sup>Access, <sup>b</sup>AxSYM, <sup>c</sup>Centaur, <sup>d</sup>Immulite, <sup>e</sup>Elecsys.

**Tabelle 9b.** Einordnung der Patienten in Erkrankungsgruppen anhand festgelegter Grenzwerte für %fPSA (15, 20 bzw. 25 %) in der Subgruppe tPSA 0-10 µg/l.

	Access	AxSYM	Centaur	Immulite	Elecsys
Ausgewählter Grenzwert	Anzahl der eingestuften Patienten				
<b>%fPSA</b>					
BPH <25 % <sup>1</sup>	217 (77,0 %) <sup>b,c,d,e</sup>	161 (57,1 %) <sup>a,d,e</sup>	156 (55,3 %) <sup>a,d,e</sup>	243 (86,2 %) <sup>a,b,c</sup>	233 (82,6 %) <sup>a,b,c</sup>
BPH <20 %	155 (55,0 %) <sup>b,c,d,e</sup>	102 (36,2 %) <sup>a,d,e</sup>	105 (37,2 %) <sup>a,d,e</sup>	203 (72,0 %) <sup>a,b,c,e</sup>	179 (63,5 %) <sup>a,b,c,d</sup>
BPH <15 %	79 (28,0 %) <sup>b,c,d,e</sup>	44 (15,6 %) <sup>a,d,e</sup>	54 (19,1 %) <sup>a,d,e</sup>	141 (50,0 %) <sup>a,b,c,e</sup>	101 (35,8 %) <sup>a,b,c,d</sup>
PCa >25 % <sup>2</sup>	9 (2,87 %) <sup>b,c</sup>	29 (9,24 %) <sup>a,d,e</sup>	24 (7,64 %) <sup>a,d,e</sup>	4 (1,27%) <sup>b,c</sup>	7 (2,23 %) <sup>b,c</sup>
PCa >20 %	21 (6,69 %) <sup>b,c,d</sup>	58 (18,5 %) <sup>a,d,e</sup>	65 (20,7 %) <sup>a,d,e</sup>	7 (2,23 %) <sup>a,b,c,e</sup>	17 (5,41 %) <sup>b,c,d</sup>
PCa >15 %	75 (23,9 %) <sup>b,c,d</sup>	131 (41,7 %) <sup>a,d,e</sup>	132 (42,0 %) <sup>a,d,e</sup>	35 (11,1 %) <sup>a,b,c,e</sup>	74 (23,6 %) <sup>b,c,d</sup>

<sup>1</sup>Anzahl der 282 BPH-Patienten mit einer %fPSA-Konzentration <25, 20 bzw. 15 %.

<sup>2</sup>Anzahl der 314 PCa-Patienten mit einer %fPSA-Konzentration >25, 20 bzw. 15 %.

Nach dem McNemar-Test für gepaarte Daten (nach Bonferroni-Korrektur (105), P <0,0001) waren die einzelnen Test signifikant unterschiedlich zu: <sup>a</sup>Access, <sup>b</sup>AxSYM, <sup>c</sup>Centaur, <sup>d</sup>Immulite, <sup>e</sup>Elecsys.

## 5. Diskussion

Seit der Entdeckung des PSA hat sich die Früherkennung des Prostatakarzinoms entscheidend verbessert (1). Die begrenzte diagnostische Aussagekraft des Tumormarkers im niedrigen PSA-Bereich konnte vor allem durch die zusätzliche Bestimmung der PSA-Untergruppen gesteigert werden (50,54). In der Vergangenheit konnte gezeigt werden, dass verschiedene Testsysteme zur Bestimmung von tPSA und fPSA unterschiedliche Ergebnisse liefern (17). Die Variabilität der methodenabhängigen Messergebnisse und mögliche Fehlinterpretationen sollten durch den Einsatz einheitlicher Referenzmaterialien und äquimolarer Methoden minimiert werden (26,30,31). Nur wenige Studien untersuchten die Unterschiede und möglichen Fehlinterpretationen, die durch den Gebrauch unterschiedlicher Methoden zur Bestimmung von %fPSA entstehen können (102,103,112). Zu Beginn dieser Studie lagen widersprüchliche Untersuchungsergebnisse zum Stand der Harmonisierung methodenabhängiger PSA-Ergebnisse vor (18,102-104).

Wie im Aufbau des Ergebnisteils möchte ich zuerst die analytischen Ergebnisse interpretieren und danach auf deren klinische Konsequenzen eingehen. Zuletzt werde ich versuchen, die Ursachen der methodenabhängigen PSA-Bestimmungen und deren Rolle bei zukünftigen Standardisierungsbemühungen zu benennen.

### **Analytische Zuverlässigkeit der Bestimmungen**

Bei über zwei Drittel der untersuchten Patientenseren wurde eine tPSA-Konzentration unter 10 µg/l gemessen (Tab. 3). In diesem klinisch relevanten tPSA-Bereich war die Verteilung der PSA-Werte in beiden Erkrankungsgruppen nahezu gleich. Die intra- und interseriellen Qualitätskontrollen ergaben für alle fünf Methoden eine hohe Zuverlässigkeit (Tab. 4 und 5). Man kann also davon ausgehen, dass die hier vorliegenden Daten eine zuverlässige Beurteilung der Sachlage erlauben.

### **Methodenvergleiche – Nachweis von testspezifischen Unterschieden der tPSA-, fPSA- und %fPSA-Werte**

Die Methodenvergleiche, berechnet in Form direkter Vergleiche der Medianwerte bzw. als Regressionsgeraden nach Passing und Bablok (Tab. 6) oder in der Darstellung nach

Bland und Altman (Abb. 2,3 und 4) sowie die so genannten Mountain-Plots (Abb. 5 und 6) zeigten zwischen den Methoden diskrepante Messergebnisse. Die tPSA-Konzentrationen zwischen den verschiedenen Methoden variierten deutlich. AxSYM und Centaur messen im Schnitt etwa 9 % niedrigere und Immulite etwa 13 % höhere tPSA-Konzentrationen als Access Hybritech. Elecsys misst dagegen nur knapp 2 % höhere tPSA-Konzentrationen als Access. Die größten Abweichungen der tPSA-Konzentrationen bestehen zwischen den Immulite- und AxSYM-Werten: Immulite bestimmt 22 % höhere tPSA-Konzentrationen als AxSYM. Spiegelbildlich dazu sind von AxSYM und Centaur bestimmte fPSA- und %fPSA-Werte höher als mit den Immulite- bzw. Elecsys-Testen. Die größte Abweichung der fPSA-Konzentrationen besteht auch hier zwischen dem AxSYM- und Immulite-Test. Die Medianwerte von fPSA-AxSYM sind 18 % höher als vom fPSA-Immulite. Aus tPSA- und fPSA-Konzentrationen errechnete %fPSA-Werte weichen noch deutlicher voneinander ab. Immulite misst 35 % niedrigere Werte als Centaur. AxSYM und Centaur bestimmen im Schnitt etwa 23 % höhere und Immulite 18 % niedrigere Werte als Access. Auch unter %fPSA-Vergleichen kann Elecsys mit einer Abweichung von nur 5 % die größte Übereinstimmung mit dem Access vorweisen. Andererseits zeigten auch hier die Steigungen der Regressionsgeraden und der Schnittpunkt der Geraden mit der Y-Achse signifikante Unterschiede einzelner Teste zum Access. Die mit dem Centaur und Elecsys zusätzlich beobachteten Unterschiede der Steigung zwischen PCa- und BPH-Patienten sind eindeutig. Methodenvergleiche nach Passing und Bablok innerhalb der klinisch relevanten Subgruppe mit einem tPSA unter 10 µg/l (Abb. 8) präsentierten ähnliche Differenzen zwischen den Methoden zur Bestimmung von tPSA und fPSA. Die größten Differenzen der Medianwerte innerhalb der Subgruppe 0-10 µg/l resultierten beim Vergleich zwischen AxSYM und Immulite: tPSA- und %fPSA-Bestimmungen von AxSYM waren im Schnitt 19 % niedriger bzw. 36 % höher. Auch hier zeigten Vergleiche der Medianwerte zwischen Access und Elecsys auf den ersten Blick Ähnlichkeiten. Bei näherem Hinsehen machten sich aber bei diesem Vergleich ebenfalls signifikante Unterschiede bemerkbar, unter anderem auch zwischen den Erkrankungsgruppen. Bland-Altman-Darstellungen und Mountain-Plots veranschaulichten besonders deutlich die beschriebenen Unterschiede zwischen den Methoden. Die beobachteten Auffälligkeiten zwischen den fPSA- bzw. %fPSA-Werten am Centaur im Vergleich zu den anderen Methoden sind nicht ohne weiteres zu erklären. Der fPSA-Centaur-Wert ergibt sich aus der Berechnung von fPSA-Centaur aus cPSA und tPSA. Trotz der oben genannten hohen Zuverlässigkeit

der cPSA- und tPSA-Bestimmungen am Centaur resultierten in einigen Fällen durch Wiederholungsmessungen bestätigte cPSA-Werte mit einer höheren Konzentration als tPSA. Dies führt bei der Berechnung von fPSA zu negativen Werten. Aber auch nach Herausnahme der negativen fPSA-Centaur-Werte fallen die fPSA- und %fPSA-Graphen von Centaur gegenüber den anderen durch ihre große Streuung auf (Abb. 3,4 und 6). Obwohl eine Fehlerfortpflanzung bei dieser Art der Berechnung des fPSA als Ursache der erhöhten Variationsbreite zu berücksichtigen ist, muss eventuell eine beeinträchtigte analytische Spezifität der cPSA-Bestimmung angenommen werden.

Folglich zeigen die Ergebnisse der Vergleiche der verschiedenen Bestimmungen, dass eindeutig testspezifische Unterschiede zwischen den Bestimmungen bestehen. Sie entsprechen zum Teil Ergebnissen, die nach 2000 bekannt wurden. Eine 2004 veröffentlichte Screeningstudie untersuchte Serumproben von 2304 Männern an den tPSA-Tests Access und Centaur und kam zu dem Ergebnis, dass der Access-Test 23 % höhere tPSA-Konzentrationen bestimmt als der Centaur (103). Eine weitere Studie beobachtete 10-17 % höhere tPSA-Konzentrationen der Access-Methode gegenüber dem Centaur (104). Die am Access gemessenen Werte in dieser Arbeit liegen um 19 % höher als die am Centaur erhaltenen tPSA-Werte. Butch et al. (100) demonstrierten im Jahre 2002 ebenfalls, dass fPSA-Konzentrationen zwischen den Messergebnissen von Access und Elecsys kleine aber statistisch signifikante Unterschiede bestehen.

Nach einer im Jahre 2006 von Kort et al. (113) veröffentlichten Studie konnte durch die Einführung des ersten internationalen WHO-PSA-Referenzmaterials vor allem unter den tPSA-Tests eine bessere Vergleichbarkeit erzielt werden. In dieser Studie wurden 70 Patientenseren mit tPSA-Konzentrationen bis 50 µg/l sowie nach WHO-Referenzmaterial zusammengestellte Proben an sechs Methoden analysiert – unter denen sich auch die hier untersuchten Access, AxSYM, Centaur und Immulite befanden. Bei Methodenvergleichen wurde der Abbott ARCHITECT Test als Referenzmethode verwendet, der sich zum AxSYM zwar im Detektionssystem (Chemielumineszenz statt Fluoreszenz), der Mikropartikelgröße und der Inkubationszeit unterscheidet, aber die gleichen Antikörper benutzt und nach dem WHO-PSA-Referenzmaterial kalibriert wird. Wie bei dem Methodenvergleich dieser Arbeit nahmen in dieser Studie tPSA-Konzentrationen in folgender Reihenfolge zu und spiegelbildlich dazu fPSA-Konzentrationen ab: Centaur, AxSYM, Access, Immulite. %fPSA-Werte von

Immulite waren niedriger als die von anderen Methoden. Schaubilder nach Bland und Altman zeigten beträchtliche Unterschiede zwischen den Methoden: Medianwerte der tPSA-Methoden wichen um 0,5-1,0 µg/l von der Referenzmethode ab, fPSA-Medianwerte um 0,12-0,40 µg/l. Bei den fPSA-Graphen fiel auch hier der Vergleich mit den errechneten fPSA-Werten des Centaur durch die große Streuung auf, wobei die stärksten Ausreißer vor allem im tPSA-Bereich über 10 µg/l zu finden waren. Methodenvergleiche nach Passing und Bablok zeigten signifikante Unterschiede zwischen den %fPSA-Methoden. Bis auf den AxSYM-Test waren alle %fPSA-Methoden bezüglich der Steigung der Regressionsgeraden oder des Schnittpunktes mit der Y-Achse signifikant verschieden vom ARCHITECT. Eine Rekalibration mit dem WHO-PSA-Referenzmaterial führte bei dem Access-Test zu einer Abweichung der Regressionsgeraden von >20 %, während bei den anderen Methoden die Abweichung <7 % blieb.

Eine mögliche Erklärung für die große Diskrepanz zwischen den verschiedenen Methoden könnte sein, dass nur drei Methoden nach dem WHO-Standard kalibriert wurden (Tab. 1). tPSA-Tests von Beckman und DPC, die nicht nach dem WHO-Standard kalibriert wurden, messen tendenziell etwas höhere tPSA-Konzentrationen (Tab. 6). Andererseits bestanden auch zwischen den gleich-kalibrierten Methoden Differenzen. So misst Elecsys höhere tPSA-Konzentrationen als AxSYM und Centaur. Die einheitliche Kalibrierung der Methoden ist dabei zwar wichtig, aber bei weitem nicht der einzige Einflussfaktor (79). Bei der Analyse der %fPSA-Werte wurde die gleiche Situation vorgefunden. „Nicht-WHO-Methoden“ zeigten eher niedrigere Werte (Immulite und auch Access) mit Median-%fPSA-Werten von 10,4 und 12,8 %, da durch die Quotientenbildung der höhere tPSA-Wert einen niedrigeren %fPSA-Wert ergibt. Dass die WHO-Kalibrierung allein keine Lösung der fehlenden Vergleichbarkeit von %fPSA-Werten sein kann, zeigen die in der vorliegenden Arbeit aufgezeigten Ergebnisse der beiden WHO-kalibrierten Methoden AxSYM und Elecsys: Mediane und Diskriminationspunkte bei 90- und 95%iger Sensitivität (Tab. 8b) waren deutlich unterschiedlich voneinander. Zwischen allen Methoden bestanden signifikante Unterschiede. Der Gebrauch fester Grenzwerte führte noch zu größeren Diskrepanzen zwischen den Methoden als bei tPSA, was sich durch die Anzahl der Fehlinterpretationen beweisen lässt (Tab. 9b). Dies ist aus klinischer Sicht bedeutend, da vor allem das %fPSA zur Verringerung unnötiger Biopsien führen soll (68). In dieser Arbeit variierten %fPSA-Grenzwerte bei festgelegter 95%iger Sensitivität in der Subgruppe zwischen 18 und 28

% (Tab. 8b), so dass nur methodenabhängige Grenzwerte vergleichbare Ergebnisse erzielen würden. Da diese Grenzwerte nicht an einer Screeningpopulation bestimmt worden sind, können sie nur zu vergleichenden Zwecken, nicht jedoch als Grenzwerte für die Praxis dienen.

### **Untersuchungen zur äquimolaren PSA-Messung mit verschiedenen Testsystemen**

Die so genannten „Molar-Response“-Graphen für tPSA-Methoden (Abb. 7) bescheinigten nur dem AxSYM-Test eine äquimolare Messung von freiem und gebundenem PSA. TPSA-Methoden von Centaur, Immulite und Elecsys zeigten im Allgemeinen nicht-äquimolare Eigenschaften mit Steigungen der Geraden ungleich Null. Zwei Studien beschrieben entgegen meinen Beobachtungen eine Äquimolarität für Immulite und Elecsys (27,102). In den beiden genannten Studien wurde jedoch auf die Erstellung detaillierter „Molar-Response“-Graphen verzichtet, sondern nur Steigungen der Regressionsgeraden ermittelter und errechneter Werte zwischen der PCa- und BPH-Gruppe miteinander verglichen. In meiner Arbeit führte eine weitere Eingrenzung der Patientenseren zu einer besseren Äquimolarität (Abb. 9). Der Centaur-Test zeigte im tPSA-Bereich bis 10 µg/l eine äquimolare Erfassung von fPSA und ACT-PSA. Bei der Verwendung von Serumproben mit einem %fPSA <25 % ergaben sich auch für die tPSA-Methoden von Immulite und Elecsys äquimolare Eigenschaften.

Es besteht weiterhin eine Kontroverse über die klinische Relevanz der Äquimolarität. Jung und Mitarbeiter (114) konnten 1996 anhand von ROC-Analysen keinen diagnostischen Vorteil von äquimolaren Methoden gegenüber dem zu diesem Zeitpunkt noch nicht-äquimolaren AxSYM erkennen. Dies lag möglicherweise daran, dass nur eine relativ geringe Anzahl von Patienten untersucht wurde und damit die so genannte statistische Power in der Studie nur gering war. Der Gebrauch methodenabhängiger Referenzwerte hatte zu gleicher Diskrimination von PCa- und BPH-Patienten geführt. Semjonow et al. (31) demonstrierten 2001 an „Molar-Response“-Graphen die erfolgreiche Modifizierung von AxSYM zu einer äquimolaren Methode durch den Ersatz der polyklonal-monoklonalen Antikörper durch monoklonal-monoklonale Antikörper. In dieser Studie konnten ROC-Analysen eine leichte Verbesserung der diagnostischen Validität der modifizierten Methoden beschreiben. In einer Folgestudie dieser Arbeitsgruppe konnte gezeigt werden, dass die Anzahl der unnötigen Biopsien durch die Modifizierung zu einer äquimolaren Methode verringert werden kann (29). Eine

Einschränkung hinsichtlich der Sensitivität musste dabei nicht in Kauf genommen werden. Dass die äquimolare Erfassung von fPSA und ACT-PSA weiterhin eine wichtige Voraussetzung für einen PSA-Test ist, konnte im Jahre 2004 in einer Computersimulationsstudie demonstriert werden (85). Dazu wurde die Anzahl der falsch-positiven bzw. falsch-negativen Biopsieindikationen zwischen äquimolaren und nicht-äquimolaren Methoden verglichen und auf die Gesamtbevölkerung von Großbritannien bezogen. Demnach würde der Einsatz von nicht-äquimolaren Methoden bei unveränderter Sensitivität eine 0,5 bis 13 % höhere Rate an falsch-positiver Biopsien verursachen. Unter der Annahme, dass nicht-äquimolare Methoden zu einer positiven Messabweichung führen, kamen die Autoren zu dem Ergebnis, dass der Einsatz dieser Methoden beim Screening einen signifikanten Anstieg der falsch-positiven sowie signifikanten Abfall der falsch-negativen Resultate verursachen würde. Eine negative Messabweichung hätte den umgekehrten Effekt.

In diesem Zusammenhang relevant ist eine von Roddam et al. (115) im Jahre 2006 veröffentlichte PSA-Wiederfindungsstudie. Aus WHO-Standards für tPSA (90:10 ACT-PSA:fPSA) und fPSA wurden 15 Proben mit tPSA-Konzentrationen von 2, 5 und 10 µg/l mit jeweils unterschiedlichen fPSA-Anteilen (10 %, 32,5 %, 55 %, 77,5 % und 100 %) zusammengestellt und zur Analyse an alle am UK NEQAS (National External Quality Assessment Services) teilnehmenden Laboratorien versandt. Hierbei wurden neben 10 anderen PSA-Testsystemen auch alle in dieser Arbeit untersuchten Methoden verwendet. Aus den gemessenen Konzentrationen und tatsächlichen Werten wurde für jede Probe eine Wiederfindungsrate errechnet. Zwischen den einzelnen Methoden bestanden klare Differenzen in der PSA-Bestimmung. Bei einer definierten tPSA-Konzentration von 5 µg/l variierten die Steigungen der Regressionsgeraden bei zunehmendem fPSA-Anteil zwischen -0,03 (AxSYM; Standardfehler 0,01), 0,09 (Centaur; 0,00), 0,13 (Access; 0,01), 0,18 (Elecsys; 0,00) und 0,28 (Immulite; 0,01) pro fPSA-Zunahme um 1 %. Der Access, der als Grundlage für die vergleichenden Tests meiner Untersuchungen diente, hatte demnach nicht-äquimolare Charakteristiken mit einer Steigung signifikant unterschiedlich zu 0. Diese Ergebnisse lagen zu Beginn meiner Untersuchungen noch nicht vor, und ich bin – entsprechend früherer Studienlage (90) – davon ausgegangen, dass Access äquimolar misst. Der Access-Test verwendet die gleichen Antikörper wie der Tandem-R-Test, hat aber offensichtlich nicht die gleichen äquimolaren Eigenschaften. Auch wenn diese Studie ansonsten zu ähnlichen Resultaten kommt – dass AxSYM

und Centaur die größte Äquimolarität aufweisen und Immulite und Elecsys eine nicht-äquimolare Messung mit einer bevorzugten Messung von fPSA haben – müssen die Ergebnisse meiner Arbeit relativiert werden. Bei Proben mit einer tPSA-Konzentration von 5 µg/l und %fPSA-Werten von 10, 30 und 50 %, stellvertretend für PCa-Patienten, gesunde Männer und BPH-Patienten, war die Messabweichung von den vorgegebenen PSA-Konzentrationen bei den Tests von Immulite und Access für alle drei Gruppen weit über 10 %, während die anderen Methoden geringere Abweichungen von maximal bis zu 10 % aufwiesen. Bei Proben mit niedrigem fPSA-Anteil hatte Immulite die größte positive Messabweichung (23 %) und Elecsys die negativste Messabweichung (-3 %). Weiterhin konnten Roddam und Mitarbeiter (115) demonstrieren, dass selbst kleine Unterschiede im Arbeitsablauf sowie ein gewisser Grad an Messabweichung und Nicht-Äquimolarität der Methoden zu unterschiedlichen Ergebnissen zwischen den Laboratorien führen können. Sie schlussfolgerten, dass es essentiell wichtig ist, äquimolare Assays anzuwenden und zur Vermeidung klinischer Fehler Kalibrationen mit internationalen Standards vorzunehmen. Diese Schlussfolgerungen beruhen jedoch ausschließlich an von präparierten PSA-Proben mit unphysiologischen fPSA-Anteilen bis zu 100 % ermittelten Molar-Response-Bestimmungen, wobei Matrixeffekte (Einfluss von Serumproteinen) nicht ausreichend berücksichtigt werden konnten.

### **Bewertung der PSA-Bestimmungen mit verschiedenen Testsystemen aus klinischer Sicht**

Die klinische Wertung der mit den verschiedenen Testsystemen gemessenen tPSA- und fPSA-Werte sowie der errechneten %fPSA-Werte zeigte unterschiedliche Ergebnisse und muss unter verschiedenen Gesichtspunkten betrachtet werden. ROC-Analysen zeigten nur geringe, klinisch nicht-signifikante Unterschiede zwischen den tPSA-Methoden (Tab. 7, Abb. 10). Diese Ergebnisse relativieren die oben angesprochene Notwendigkeit, unbedingt äquimolare Tests einsetzen zu müssen. Wahrscheinlich schränken geringfügige Abweichungen von der Äquimolarität die klinische Aussagekraft eines Tests nicht wesentlich ein. Wird jedoch aufgrund der anfangs dargestellten unterschiedlichen absoluten PSA-Werte der verschiedenen Testsysteme ein Grenzwert für klinische Einordnungen zugrunde gelegt, ergeben sich gravierende Schlussfolgerungen. Bei einem konventionellen tPSA-Grenzwert von 4 µg/l, der allgemein zur Entscheidung für eine Prostatabiopsie verwendet wird, resultieren je nach Assay eine merklich unterschiedliche Anzahl an PCa- und BPH-

Patienten mit falsch-positiver bzw. falsch-negativer Einordnung (Tab. 9a). Da die Unterschiede zwischen den PSA-Werten bei den verschiedenen Testen linear sind, würden diese fehlerhaften Einstufungen sich auch bei einer Wahl von niedrigeren Grenzwerten zur Biopsieentscheidung ergeben (116,117). Da nach der tPSA-Messung die erste Entscheidung über den Fortgang der Diagnostik getroffen wird, ist die Suche nach dem optimalen tPSA-Grenzwert besonders wichtig. In der ärztlichen Routine wird nach wie vor der konventionelle Grenzwert von 4 µg/l bevorzugt. Die Verfasser der European Randomized Study for Prostate Cancer senkten den Grenzwert auf 3 µg/l, weil etwa 20 % der klinisch wichtigen Tumoren im niedrigen PSA-Bereich durch einen Grenzwert von 4 µg/l nicht erfasst werden (13). In den USA besteht sogar die Tendenz, den Grenzwert auf 2,5 µg/l herabzusetzen, um eine Sensitivität von mindestens 90 % zu erreichen (118). Tatsächlich würde in dieser Arbeit erst bei einer Absenkung des Grenzwertes auf 3,1 µg/l eine 90%ige Sensitivität für alle Methoden erreicht werden (Tab. 8a). Weiterhin konnten Thompson et al. (12) zeigen, dass die Detektionsrate vor allem von 2-4 µg/l dem Grauzonenbereich von 4-10 µg/l entspricht. Es ist bedauerlich, dass innerhalb dieser Diskussionen die weiterhin unzureichende Variabilität der PSA-Konzentrationen unterschiedlicher Methoden fehlt. Dabei konnte ich in dieser Arbeit zeigen, zu welchen Fehlinterpretationen die Verwendung einheitlicher Grenzwerte führt. Die in Tabelle 8a aufgeführte Spanne der ermittelten tPSA-Grenzwerte von 3,1 bis 4,0 µg/l, um eine 90%ige Sensitivität zu erzielen, unterstreichen dieses Problem.

In der von Link et al. (103) veröffentlichten Screeningstudie wurden ebenfalls die klinischen Auswirkungen der höheren tPSA-Bestimmung am Access gegenüber dem Centaur untersucht. Bei einem fixen Grenzwert von 4 µg/l wären mit dem Access-Test etwa 20 % mehr Patienten pathologisch eingestuft worden als mit dem Centaur. Da in dieser Studie aber keine Biopsien entnommen wurden, konnten die diagnostischen Validitäten jedoch nicht miteinander verglichen werden. Fillela et al. (104) verglichen mit der ROC-Analyse die diagnostische Leistung der PSA-Bestimmungen am Centaur und Access an Serumproben von 128 PCa- und 72 BPH-Patienten. Obwohl am Access deutlich höhere tPSA-Konzentrationen ermittelt wurden, hatten beide Methoden fast identische Flächen unter der ROC-Kurve. Dies entspricht genau den von mir oben beschriebenen Ergebnissen. Die Autoren beobachteten außerdem, dass der Ersatz des %fPSA durch f/cPSA die diagnostische Validität steigern konnte, wobei eine Anhebung der Spezifität bei 90%iger Sensitivität von 35 % auf 44 % im Bereich der tPSA-Werte

von 4-10 µg/l ermittelt wurde. Der Umstand, dass dieser Quotient aus Methoden verschiedener Hersteller gewonnen wurde, liefert aber eine weitere potentielle Fehlerquelle, die zu falschen klinischen Entscheidungen führen könnte.

Die ähnlichen diagnostischen Validitäten der tPSA-Tests (Tab. 8a) beweisen, dass diese Methoden bezüglich der diagnostischen Leistung vergleichbar sind, solange methodenabhängige Grenzwerte gewählt werden. Wählt man hingegen einen fixen Grenzwert von 2,5 µg/l oder 4 µg/l, ist die Anzahl der richtig eingestuftten PCa- und BPH-Patienten zwischen den einzelnen Methoden signifikant unterschiedlich (Tab. 9a). TPSA-Methoden, die tendenziell niedrigere tPSA-Konzentrationen bestimmen (AxSYM und Centaur), würden bei der Wahl einheitlicher Grenzwerte Indikation für weniger Prostatabiopsien sein. Dies würde zwar auf der einen Seite einigen BPH-Patienten eine unnötige Biopsie ersparen, aber andererseits dazu führen, dass Patienten mit einem PCa übersehen würden. Die mit dem am Immulite gemessenen höheren tPSA-Konzentrationen haben hingegen entgegengesetzte Konsequenzen.

Auch wenn zum Teil signifikant unterschiedlich, kann auch bei den Flächen unter der ROC-Kurve von ähnlichen diagnostischen Leistungen der %fPSA-Bestimmungen der verschiedenen Methoden ausgegangen werden (Abb. 11). Die Daten in Tabelle 8b legen die Schlussfolgerung nahe, dass die diagnostische Validität der verschiedenen %fPSA-Bestimmungen nahezu gleich sein würde, wenn methodenabhängige Grenzwerte bei 90%iger oder 95%iger Sensitivität benutzt würden. Bei festgelegten Grenzwerten für %fPSA (Tab. 9b) ist hingegen die Anzahl der falsch eingestuftten Patienten zwischen den meisten Methoden signifikant verschieden. Aus den Daten ermittelte Spezifitäten und Sensitivitäten bei festgelegten Diskriminationspunkten ergeben viel größere Unterschiede als methodenabhängige Grenzwerte. Durch die signifikant unterschiedlichen Mediane, durch verschiedene Grenzwerte bei vorgegebener 90%iger Sensitivität und durch unterschiedliche Ergebnisse bei der ROC-Analyse sind folglich die %fPSA-Werte als klar methodenabhängig belegt. Dieses Ergebnis ist von besonderem klinischem Interesse, weil %fPSA-Werte üblicherweise benutzt werden, um im so genannten tPSA-Graubereich von 2-10 µg/l zwischen BPH- und PCa-Patienten zu differenzieren.

Im Jahre 2000 konnte ein Vergleich der diagnostischen Leistung sechs verschiedener Methoden zur Bestimmung von fPSA- und tPSA, einschließlich des AxSYM-, Immulite- und Elecsys-Tests, ähnliche Beobachtungen machen (102). Die Anzahl der anhand fixer tPSA- und %fPSA-Grenzwerte korrekt eingestuftener Patienten variierte zwischen den Methoden. So hatte im tPSA-Bereich von 0-10 µg/l bei einem Diskriminationspunkt von 15 % die %fPSA-Bestimmung von Immulite mit 100 % die höchste Sensitivität, gefolgt von AxSYM (69 %) und Elecsys (64 %), aber gleichzeitig mit 50 % eine deutlich niedrigere Spezifität als Elecsys (59 %) bzw. AxSYM (69 %). Anhand der Verteilung der Patienten errechneten die Autoren testspezifische Grenzwerte für %fPSA zur optimalen Differenzierung zwischen PCa und BPH im Bereich von 0-10 µg/l: 20 % für Elecsys, 15 % für AxSYM und 10 % für Immulite. Sie nannten eine Grauzone für %fPSA Werte von 10-20 %, die durch methodenabhängige Grenzwerte präzisiert werden könnte. Da das Patientenkollektiv dieser Studie sehr heterogen war und nur etwa bei einem Drittel der 282 Patienten ein tPSA-Wert unter 20 µg/l vorlag sowie Variationskoeffizienten bei interseriellen Präzisionskontrollen von bis zu 20 % auftraten, müssen die Ergebnisse dieser Studie mit Vorsicht betrachtet werden.

Im Rahmen der ERSCP-Studie (European Randomized Study of Screening for Prostate Cancer) wurde der Einfluss der Methodenabhängigkeit der PSA-Bestimmung auf die klinische Entscheidungsfindung untersucht (18). Im tPSA-Bereich von 1 bis 6 µg/l wurden Serumproben von 360 PCa- und 96 Nicht-PCa-Patienten an fünf Methoden zur Bestimmung von fPSA und tPSA, darunter auch Access und Elecsys, untersucht. Die Methoden wiesen signifikant verschiedene PSA-Ergebnisse auf und zeigten ein uneinheitliches Ansprechen auf den WHO-Standard. Vor allem zwischen den %fPSA-Bestimmungen traten große Differenzen auf: %fPSA-Werte von Elecsys und Access variierten um 32 % bzw. 36 % unter nicht-PCa bzw. PCa-Patienten. Auf das gesamte Patientenkollektiv von 17 334 Männern übertragen, würden bei einem Wechsel des Testsystems, das 20 % höhere PSA-Konzentrationen bestimmt, bis zu 1008 Patienten zusätzlich biopsiert werden. Dabei muss angemerkt werden, dass hierzu von den Autoren keine näheren Berechnungen durchgeführt wurden. Insgesamt zeigten die verschiedenen Methoden Unterschiede in der Anzahl der detektierten Tumorpatienten, nicht aber pathologisch hinsichtlich Tumorstadium oder -grad (dargestellt in einer weiteren Veröffentlichung (97)).

Die vorgestellten Ergebnisse meiner Arbeit zeigen, dass trotz erheblicher Anstrengung der letzten 15 Jahre weiterhin eine inadäquate Vergleichbarkeit von tPSA, fPSA und %fPSA-Werten unterschiedlicher Testsysteme vorliegt und das Ziel der Standardisierung zum jetzigen Zeitpunkt noch nicht erreicht werden konnte. Im Folgenden möchte ich mich deshalb mit den möglichen Ursachen dafür auseinandersetzen und einen Blick in die Zukunft werfen.

### **Standardisierung und Harmonisierung der PSA-Testsysteme**

Standardisierung im eigentlichen Sinn beinhaltet den Abbau von Messabweichungen der einzelnen Bestimmungsmethoden. Sie soll die klinische Interpretation der Messergebnisse erleichtern und den Vergleich unterschiedlicher Studien ermöglichen. Die Standardisierung von Immunoassays betrifft die Bestimmung verschiedenster Moleküle. Im Gegensatz zu den meisten Steroidhormonassays gibt es für die PSA-Bestimmung keine Referenzmethode, keinen klaren Grenzwert und ein sehr heterogenes Analysat, das sich vom Standard unterscheidet.

Die Heterogenität der Antigene in biologischen Medien stellt ein besonderes Problem bei der Standardisierung von Immunoassays dar (119). Genau genommen ist die Standardisierung nur dann möglich, wenn Standard und Analysat identisch sind. Das PSA verfügt jedoch über eine ausgesprochen große Heterogenität mit verschiedenen Vorläufern, Untergruppen, Fragmenten und unterschiedlichen Aktivitätsformen (Kapitel 1.2.2). Neben dem fPSA, dem  $\alpha_2$ -Makroglobulin, dem ACT-PSA und den anderen PSA-Komplexen kann in Seren das Vorläuferprotein proPSA mit zusätzlichen Aminosäuren bestimmt werden. Darüber hinaus weist das aus 237 Aminosäuren bestehende Glykoprotein eine Mikroheterogenität auf, bestehend aus Glykosylierungen sowie Einkerbungen („nicked“-PSA) und Strangbrüchen in der Primärstruktur. Allein schon bei der Erfassung der beiden großen Untergruppen, dem ACT-PSA und dem fPSA, gibt es unterschiedlich starke Spezifitäten zwischen den Antikörpern der verschiedenen Methoden. Wie diese Arbeit und die Studie von Kort et al. (113) zeigten, wurde das Problem der äquimolaren Erfassung der beiden Untergruppen noch nicht behoben. Auch die Erfassung von nicked-PSA ist methodenabhängig. Während manche Antikörper gespaltene und ganze Aminosäurenkette gleichermaßen erfassen, scheinen andere die Anzahl der gespaltenen unterzubewerten (120). Deshalb ist es wichtig, die Reaktivität der Antikörper gegenüber Standards und dem Analysat genau zu kennen.

Die Spezifität einer Methode hängt hauptsächlich von der Epitop-Spezifität und -affinität des verwendeten Antikörpers ab. Die unterschiedlichen Epitop-Bindungsplätze der Antikörper verschiedener Methoden führen zu signifikant unterschiedlichen PSA-Bestimmungen (20,25,80,121).

Die Entwicklung einer Referenzmethode ist schwierig, da neben den Antiköperspezifität auch andere Faktoren, wie etwa das Testdesign, die Messung von PSA beeinflussen. Um akzeptable Ergebnisse zu liefern, sollte eine Referenzmethode unanhängig von Matrixeffekten sein, auf verschiedene Testformate übertragen werden können, nach internationalen Standards kalibriert werden und monoklonale Antikörpern mit bekannter Epitoperkennung verwenden (122).

Eine weitere Quelle für unterschiedliche PSA-Bestimmungen können Matrix-Effekte in den Standardpräparationen sein. Die Variabilität der PSA-Methoden könnte teilweise auf die unphysiologischen Grundsubstanz (Matrix) des WHO-Referenzmaterials zurückzuführen sein. Die WHO-PSA-Standards für tPSA (90:10 ACT-PSA:fPSA) und fPSA sind in einer Grundsubstanz aus Rinderserum-Albumin erhältlich (30,123). Grundsubstanzen, die dem menschlichen Serum ähnlich sind, verbessern die Präzision von PSA-Bestimmungen signifikant (26). In Europa wurde deshalb unter Aufsicht des Referenzbüros der Europäischen Gemeinschaft für Standards, Messung und Tests ein matrixfreies gereinigtes fPSA namens CRM 613 (zertifiziertes PSA-Referenzmaterial) hergestellt. Es ist seit 1999 erhältlich und kann zur Kalibrierung verwendet werden. Bisher ist dies jedoch nur sehr selten erfolgt (27).

Der für die meisten kommerziell erhältlichen PSA-Tests verwendete Kalibrator besteht aus einer 90:10 Mischung aus ACT-PSA und fPSA. Es ist der von der WHO als erstes Referenzmaterial für PSA angenommene Standard (30). Durch Kalibrierung der Methoden nach diesem Standard konnten unterschiedliche PSA-Ergebnisse verschiedener Messmethoden signifikant verringert werden (26,99). Eine Rekalibrierung führt zu einer besseren Vergleichbarkeit der PSA-Konzentrationen unterschiedlicher Methoden (27). Die größere Variabilität der fPSA-Konzentrationen lässt sich zum Teil darauf zurückführen, dass Hersteller für diese Bestimmung weiterhin eigene Kalibratoren herstellen, obwohl auch hierfür ein Standard besteht (30). Dass aber auch die Zusammensetzung des zweiten „Arbeits“-Kalibrators die Leistung der Tests beeinflusst

sen kann, konnte in einem Vergleich zwischen dem Europäischen Zertifizierten Referenzmedium (CRM 613) und fünf PSA-Reagenzien aus fünf kommerziellen Quellen mittels Massenspektrometrie gezeigt werden (124). Die Präparationen unterschieden sich merklich in ihrer Heterogenität und Menge. Die Autoren schlussfolgerten, dass der Gebrauch von PSA-Standards aus unterschiedlichen Bezugsquellen zur Kalibrierung zu signifikant unterschiedlichen PSA-Bestimmungen führen kann. Die in dieser Studie untersuchten Reagenzien enthielten PSA, das aus Seminalplasma extrahiert wurde. Da das fPSA aus Seminalplasma aber andere Charakteristiken aufweist als das im Serum, entsteht bei der Kalibrierung mit diesen PSA-Reagenzien eine weitere potentielle Fehlerquelle (125,126).

Durch die Einführung von Qualitätsstandards und die Verfügbarkeit externer Kontrollprogramme können Hersteller unter Beweis stellen, dass ihre Methoden den allgemeinen Anforderungen entsprechen. Es ist wichtig, dass die PSA-Kontrollen aus möglichst physiologischen Lösungen bestehen, da synthetische Proben dazu neigen, methodenabhängigen Unterschiede zu verdecken (126,127). Ein Vergleich der Proben konnte zeigen, dass die Variabilität der Methoden im Bereich des WHO-Standards minimiert ist (94).

Neben den zuvor genannten Gründen hat der Aufbau des Tests einen wesentlichen Einfluss auf die Testresultate. Hier nimmt insbesondere die Inkubationszeit Einfluss auf die Spezifität des Testes (20). Auch die Wahl des Testprinzips hat entscheidenden Einfluss auf die Richtigkeit der Messung. Früher beobachtete „high-dose-hook“-Effekte, die zu falsch niedrigen Ergebnissen führten, werden heutzutage nicht mehr beobachtet (23). Polyklonale Antikörper wurden bemängelt, weil sie typischerweise größere fPSA-Konzentrationen messen, da sie auch eine gewisse Affinität zu Epitopen des ACT-PSA-Komplexes besitzen (54). Serum von Patienten, die mit monoklonalen Antikörpern behandelt wurden, kann mit den Anti-PSA-Antikörpern kreuzreagieren und zu falsch-positiven Ergebnissen führen (128). Grundsätzlich gibt es etliche Faktoren, die zur methodenabhängigen PSA-Bestimmungen beitragen.

Die Standardisierung ist höchst aufwendig. Allein für den WHO-PSA-Standard sind von der Idee bis zur Fertigstellung fünf bis sechs Jahre vergangen (30). Zwar kann auch ohne eine Standardisierung an der besseren Harmonisierung von PSA-Ergebnissen

gearbeitet werden, aber schon allein die korrekte Kalibrierung in Abwesenheit einer Referenzmethode ist problematisch. Die Entwicklung von Korrekturfaktoren zur Umrechnung von tPSA-Werte zwischen unterschiedlichen Methoden scheint momentan nicht möglich zu sein, weil weiterhin eine unzureichende Äquimolarität vorhanden ist und Vergleiche selbst zwischen den Erkrankungsgruppen unterschiedliche Ergebnisse liefern können (siehe Kapitel 4.3.1). Zum gegenwärtigen Stand sollten Referenzwerte für jeden Test separat bestimmt werden. Diese müssten nach einheitlichen Kriterien, zum Beispiel nach den Empfehlungen des IFCC (International Federation of clinical chemistry) (26), etabliert werden – was aber bisher selten der Fall ist. Dies kann recht aufwendig sein und erfordert eine möglichst gute Zusammenarbeit verschiedener Laboren, Hersteller und Kliniken. Idealerweise sollten Proben von einer gut charakterisierten Referenzpopulation gesammelt und den Herstellern zugänglich gemacht werden. Ein starrer Grenzwert von 4 gilt schon deshalb als problematisch, weil allein die biologische Variation von tPSA-Konzentrationen bei über 50jährigen Männern im Bereich von 0,1-20 µg/l bei 20 % liegt (61). Die Ersetzung eines Grenzwertes durch ein Risikoprofil, wie dies zurzeit bei an %fPSA orientierten artifiziellen neuronalen Netzwerken vorgenommen wird, kann die schwache Spezifität des Tumormarkers substantiell verbessern. Mittlerweile wurden hierfür auch schon methodenabhängige artifizielle neuronale Netzwerke entwickelt (16). Vor dem Hintergrund, dass die Hersteller immer wieder ihre Methoden verändern, sollte auch die Übertragbarkeit der Ergebnisse alter Studien, die etwa zur Bestimmung eines Grenzwertes durchgeführt wurden, kritisch betrachtet werden.

**Fazit:**

Diese Arbeit untersuchte die Vergleichbarkeit von PSA-Methoden unterschiedlicher Hersteller. Auch wenn im Allgemeinen von Seiten der Hersteller behauptet wird, dass ihre PSA-Methoden nach dem WHO-PSA-Referenzmaterial kalibriert werden und äquimolar messen, konnte diese Studie beweisen, dass die Variabilität zwischen den Methoden noch nicht behoben werden konnte. Es besteht weiterhin eine unzureichende Vergleichbarkeit der PSA-Ergebnisse. Dabei entsprechen die Abweichungen der Konzentrationen der einzelnen Methoden weitestgehend den Ergebnissen anderer Studien (100,103,104,113,115). Das untersuchte Patientenkollektiv ist nicht repräsentativ für die Prävalenz des PCa und kann nicht auf die Gesamtbevölkerung bezogen werden. Dennoch zeigt diese Studie, dass durch die analytischen Einschränkungen weiterhin

klinische Entscheidungen, etwa die der Prostatabiopsieempfehlung, beeinflussen werden. Andere Studien kamen diesbezüglich zu ähnlichen Ergebnissen (18,85,102,103). Um die Situation realistischer darzustellen, verzichtete ich auf die Rekalibrierung der Methoden mit einem bekannten Kalibrator. Das Ziel methodenunabhängiger fPSA- und tPSA-Konzentrationen konnte bisher nicht erreicht werden. Wegen der Heterogenität des PSA, einschließlich struktureller Diversitäten zwischen malignen oder benignen Ursprungs, unterschiedlichen analytischen Bedingungen durch den Gebrauch verschiedenster Anti-PSA-Antikörperkombinationen mit unterschiedlichen Epitop-Affinitäten und -Spezifitäten sowie unterschiedlichen technischen Prinzipien wird dieses Ziel auch in naher Zukunft schwer zu erreichen sein (80). Um Fehlinterpretation von PSA-Ergebnissen, die mit verschiedenen Methoden bestimmt wurden, zu minimieren, sollten Ärzte auf bestehende Unterschiede zwischen den Ergebnissen unterschiedlicher Methoden und die biologische Variation des PSA aufmerksam gemacht werden (61). Zusätzlich ist es erforderlich, methodenabhängige Referenzintervalle zu erarbeiten.

## 6. Zusammenfassung

Der Vergleich von fünf Methoden zur Bestimmung von tPSA und fPSA bzw. cPSA (ADVIA Centaur) an Serumproben von 798 Patienten mit benigner Prostatahyperplasie bzw. Prostatakarzinom zeigte eine inadäquate Variabilität zwischen den tPSA-, fPSA- und %fPSA-Werten der einzelnen Methoden. Diese hochsignifikanten und klinisch relevanten Unterschiede wurden auch bei Methoden beobachtet, die nach dem WHO-Standard kalibriert wurden. Solange noch keine bessere Harmonisierung der PSA-Ergebnisse unterschiedlicher Methoden erreicht wurde, kann die Diskussion über tPSA- und %fPSA-Grenzwerte für eine Biopsieempfehlung zu keiner endgültigen Empfehlung führen. Weitere Untersuchungen bezüglich des Referenzmaterials oder einer Referenzmethode sind notwendig, um diese schwierige Situation in der Praxis zu verbessern. Zurzeit muss festgestellt werden, dass der Gebrauch einheitlicher Grenzwerte für unterschiedliche Testsysteme zur Bestimmung von tPSA und fPSA zu ernststen Problemen führt.

Zusammenfassend können folgende Schlußfolgerungen gezogen werden:

- Jeder der hier verwendeten PSA-Teste verfügt über eine gute analytische Zuverlässigkeit.
- Die tPSA-, fPSA- und %fPSA-Ergebnisse variieren stark zwischen den Methoden.
- Die äquimolare Erfassung von freiem und gebundenem PSA ist bei den meisten tPSA-Testsystemen weiterhin unzureichend. Manche Hersteller verzichten weiterhin auf eine Kalibrierung ihrer Methoden nach dem WHO-Standard.
- Hinsichtlich der klinischen Aussagekraft gibt es aufgrund der ROC-Analysen keine wesentlichen Unterschiede zwischen den Methoden. Werden jedoch feste Diskriminationspunkte, z.B. der konventionelle Grenzwert 4 µg/l oder 2,5 µg/l für Biopsieentscheidungen, herangezogen, resultieren aufgrund der absoluten Differenzen der PSA-Werte auch erhebliche Unterschiede in der Richtigkeit der Zuordnung.
- Die Unerlässlichkeit der Standardisierung/ Harmonisierung der PSA-Bestimmungen wurde in dieser Studie belegt. Es bedarf erheblicher Anstrengungen, um dieses in der gegenwärtigen Situation aufgrund der Komplexität der Standardisierung bestehende Problem im Sinne des Patienten zu lösen.

## 7. Literaturverzeichnis

1. Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer statistics, 2007. *CA Cancer J Clin* 2007;57:43-66.
2. Lutz JM, Francisci S, Mugno E, Usel M, Pompe-Kirn V, Coebergh JW, et al. Cancer prevalence in Central Europe: the EUROPREVAL Study. *Ann Oncol* 2003;14:313-22.
3. Rodriguez C, Freedland SJ, Deka A, Jacobs EJ, McCullough ML, Patel AV, et al. 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. Scherr D, Swindle PW, Scardino PT. National Comprehensive Cancer Network guidelines for the management of prostate cancer. *Urology* 2003;61:14-24.
5. Thompson IM, Ernst JJ, Gangai MP, Spence CR. Adenocarcinoma of the prostate: results of routine urological screening. *J Urol* 1984;132:690-2.
6. 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.
7. Schroder FH, Roobol-Bouts M, Vis AN, van der KT, Kranse R. Prostate-specific antigen-based early detection of prostate cancer--validation of screening without rectal examination. *Urology* 2001;57:83-90.
8. Gohagan JK, Prorok PC, Kramer BS, Cornett JE. Prostate cancer screening in the prostate, lung, colorectal and ovarian cancer screening trial of the National Cancer Institute. *J Urol* 1994;152:1905-9.
9. 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.
10. Bunting PS. A guide to the interpretation of serum prostate specific antigen levels. *Clin Biochem* 1995;28:221-41.
11. 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.
12. Thompson IM, Pauler DK, Goodman PJ, Tangen CM, Lucia MS, Parnes HL, et al. Prevalence of prostate cancer among men with a prostate-specific antigen level  $\leq$  4.0 ng per milliliter. *N Engl J Med* 2004;350:2239-46.
13. Schroder FH, van dC-K, I, de Koning HJ, Vis AN, Hoedemaeker RF, Kranse R. Prostate cancer detection at low prostate specific antigen. *J Urol* 2000;163:806-12.
14. 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.
15. Brawer MK, Meyer GE, Letran JL, Bankson DD, Morris DL, Yeung KK, et al. Measurement of complexed PSA improves specificity for early detection of prostate cancer. *Urology* 1998;52:372-8.

16. 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.
17. Semjonow A, De AG, Oberpenning F, Schmid HP, Brandt B, Hertle L. The clinical impact of different assays for prostate specific antigen. *BJU Int* 2000;86:590-7.
18. Blijenberg BG, Yurdakul G, van Zelst BD, Bangma CH, Wildhagen MF, Schroder FH. Discordant performance of assays for free and total prostate-specific antigen in relation to the early detection of prostate cancer. *BJU Int* 2001;88:545-50.
19. Stamey TA. Second Stanford Conference on International Standardization of Prostate-Specific Antigen Immunoassays: September 1 and 2, 1994. *Urology* 1995;45:173-84.
20. Graves HC. Standardization of immunoassays for prostate-specific antigen. A problem of prostate-specific antigen complexation or a problem of assay design? *Cancer* 1993;72:3141-4.
21. Nakamura RM. Current and future directions regarding quality assurance and standardization of prostate specific antigen immunoassays. *Cancer* 1994;74:1655-9.
22. Stenman UH, Leinonen J, Zhang WM. Standardization of PSA determinations. *Scand J Clin Lab Invest Suppl* 1995;221:45-51.
23. Semjonow A, Brandt B, Oberpenning F, Roth S, Hertle L. Discordance of assay methods creates pitfalls for the interpretation of prostate-specific antigen values. *Prostate Suppl* 1996;7:3-16.
24. Graves HC, Wehner N, Stamey TA. Comparison of a polyclonal and monoclonal immunoassay for PSA: need for an international antigen standard. *J Urol* 1990;144:1516-22.
25. Zhou AM, Tewari PC, Bluestein BI, Caldwell GW, Larsen FL. Multiple forms of prostate-specific antigen in serum: differences in immunorecognition by monoclonal and polyclonal assays. *Clin Chem* 1993;39:2483-91.
26. Stamey TA, Chen Z, Prestigiacomo AF. Reference material for PSA: the IFCC standardization study. International Federation of Clinical Chemistry. *Clin Biochem* 1998;31:475-81.
27. Blijenberg BG, Storm BN, van Zelst BD, Kruger AE, Schroder FH. New developments in the standardization of total prostate-specific antigen. *Clin Biochem* 1999;32:627-34.
28. Chan DW, Sokoll LJ. WHO first international standards for prostate-specific antigen: the beginning of the end for assay discrepancies? *Clin Chem* 2000;46:1291-2.
29. Oberpenning F, Weining C, Brandt B, De AG, Heinecke A, Hamm M, et al. A new modification of the Chiron ACS assay for total prostate-specific antigen achieves equimolar response characteristics and improves the detection of prostate cancer. *Clin Chem Lab Med* 2003;41:90-4.
30. Rafferty B, Rigsby P, Rose M, Stamey T, Gaines DR. Reference reagents for prostate-specific antigen (PSA): establishment of the first international standards for free PSA and PSA (90:10). *Clin Chem* 2000;46:1310-7.
31. Semjonow A, Oberpenning F, Weining C, Schon M, Brandt B, De AG, et al. Do modifications of nonequimolar assays for total prostate-specific antigen improve detection of prostate cancer? *Clin Chem* 2001;47:1472-5.

32. 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.
33. Ablin RJ, Soanes WA, Bronson P, Witebsky E. Precipitating antigens of the normal human prostate. *J Reprod Fertil* 1970;22:573-4.
34. 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. VII. *Nippon Hoigaku Zasshi* 1971;25:322-4.
35. Li TS, Beling CG. Isolation and characterization of two specific antigens of human seminal plasma. *Fertil Steril* 1973;24:134-44.
36. Wang MC, Valenzuela LA, Murphy GP, Chu TM. Purification of a human prostate specific antigen. *Invest Urol* 1979;17:159-63.
37. 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.
38. Papsidero LD, Wang MC, Kuriyama M, Horoszewicz JS, Leong SS, Croghan GA, et al. Human prostate epithelial cell-specific antigen. *Prog Clin Biol Res* 1981;75A:435-43.
39. 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.
40. Eide IA, Angelsen A. Prostate-specific antigen. *Tidsskr Nor Laegeforen* 2000;120:2528-31.
41. Chu TM, Kuriyama M, Johnson E, Papsidero LD, Killian CS, Murphy GP, et al. Circulating antibody to prostate antigen in patients with prostate cancer. *Transplant Proc* 1984;16:481-5.
42. Malm J, Lilja H. Biochemistry of prostate specific antigen, PSA. *Scand J Clin Lab Invest Suppl* 1995;221:15-22.
43. Riegman PH, Vlietstra RJ, Suurmeijer L, Cleutjens CB, Trapman J. Characterization of the human kallikrein locus. *Genomics* 1992;14:6-11.
44. 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.
45. Yousef GM, Diamandis EP. The new human tissue kallikrein gene family: structure, function, and association to disease. *Endocr Rev* 2001;22:184-204.
46. 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.
47. Lilja H. A kallikrein-like serine protease in prostatic fluid cleaves the predominant seminal vesicle protein. *J Clin Invest* 1985;76:1899-903.
48. Diamandis EP. Prostate-specific antigen: a cancer fighter and a valuable messenger? *Clin Chem* 2000;46:896-900.
49. Christensson A, Laurell CB, Lilja H. Enzymatic activity of prostate-specific antigen and its reactions with extracellular serine proteinase inhibitors. *Eur J Biochem* 1990;194:755-63.
50. Lilja H, Christensson A, Dahlen U, Matikainen MT, Nilsson O, Pettersson K, et al. Prostate-specific antigen in serum occurs predominantly in complex with alpha 1-antichymotrypsin. *Clin Chem* 1991;37:1618-25.

51. Zhang WM, Finne P, Leinonen J, Vesalainen S, Nordling S, Rannikko S, et al. Characterization and immunological determination of the complex between prostate-specific antigen and alpha2-macroglobulin. *Clin Chem* 1998;44:2471-9.
52. Bjork T, Ljungberg B, Piironen T, Abrahamsson PA, Pettersson K, Cockett AT, et al. Rapid exponential elimination of free prostate-specific antigen contrasts the slow, capacity-limited elimination of PSA complexed to alpha 1-antichymotrypsin from serum. *Urology* 1998;51:57-62.
53. 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.
54. 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.
55. Christensson A, Bjork T, Nilsson O, Dahlen U, Matikainen MT, Cockett AT, et al. Serum prostate specific antigen complexed to alpha 1-antichymotrypsin as an indicator of prostate cancer. *J Urol* 1993;150:100-5.
56. Allard WJ, Zhou Z, Yeung KK. Novel immunoassay for the measurement of complexed prostate-specific antigen in serum. *Clin Chem* 1998;44:1216-23.
57. Catalona WJ, Richie JP, Ahmann FR, Hudson MA, Scardino PT, Flanigan RC, et al. 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.
58. Price CP, Allard J, Davies G, Dawnay A, Duffy MJ, France M, et al. Pre- and post-analytical factors that may influence use of serum prostate specific antigen and its isoforms in a screening programme for prostate cancer. *Ann Clin Biochem* 2001;38:188-216.
59. Sasagawa I, Adachi M, Sawamura T, Iijima Y, Hayami S, Kubota Y, et al. Serum levels of total and free testosterone in men undergoing hemodialysis. *Arch Androl* 1998;40:153-8.
60. Wang LG, Liu XM, Kreis W, Budman DR. Down-regulation of prostate-specific antigen expression by finasteride through inhibition of complex formation between androgen receptor and steroid receptor-binding consensus in the promoter of the PSA gene in LNCaP cells. *Cancer Res* 1997;57:714-9.
61. Soletormos G, Semjonow A, Sibley PE, Lamerz R, Petersen PH, Albrecht W, et al. Biological variation of total prostate-specific antigen: a survey of published estimates and consequences for clinical practice. *Clin Chem* 2005;51:1342-51.
62. Raaijmakers R, Kirkels WJ, Roobol MJ, Wildhagen MF, Schrder FH. Complication rates and risk factors of 5802 transrectal ultrasound-guided sextant biopsies of the prostate within a population-based screening program. *Urology* 2002;60:826-30.
63. Carter HB. Prostate cancers in men with low PSA levels--must we find them? *N Engl J Med* 2004;350:2292-4.
64. Seaman E, Whang M, Olsson CA, Katz A, Cooner WH, Benson MC. PSA density (PSAD). Role in patient evaluation and management. *Urol Clin North Am* 1993;20:653-63.
65. Brawer MK, Beatie J, Wener MH, Vessella RL, Preston SD, Lange PH. Screening for prostatic carcinoma with prostate specific antigen: results of the second year. *J Urol* 1993;150:106-9.

66. Schroder FH, Kranse R, Barbet N, Hop WC, Kandra A, Lassus M. Prostate-specific antigen: A surrogate endpoint for screening new agents against prostate cancer? *Prostate* 2000;42:107-15.
67. Oesterling JE, Jacobsen SJ, Chute CG, Guess HA, Girman CJ, Panser LA, et al. Serum prostate-specific antigen in a community-based population of healthy men. Establishment of age-specific reference ranges. *JAMA* 1993;270:860-4.
68. Catalona WJ, Partin AW, Slawin KM, Brawer MK, Flanigan RC, Patel A, et al. Use of the percentage of free prostate-specific antigen to enhance differentiation of prostate cancer from benign prostatic disease: a prospective multicenter clinical trial. *JAMA* 1998;279:1542-7.
69. Piironen T, Pettersson K, Suonpaa M, Stenman UH, Oesterling JE, Lovgren T, et al. In vitro stability of free prostate-specific antigen (PSA) and prostate-specific antigen (PSA) complexed to alpha 1-antichymotrypsin in blood samples. *Urology* 1996;48:81-7.
70. Partin AW, Brawer MK, Bartsch G, Horninger W, Taneja SS, Lepor H, et al. Complexed prostate specific antigen improves specificity for prostate cancer detection: results of a prospective multicenter clinical trial. *J Urol* 2003;170:1787-91.
71. Veltri RW, Chaudhari M, Miller MC, Poole EC, O'Dowd GJ, Partin AW. Comparison of logistic regression and neural net modeling for prediction of prostate cancer pathologic stage. *Clin Chem* 2002;48:1828-34.
72. Stephan C, Stroebel G, Heinau M, Lenz A, Roemer A, Lein M, et al. 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.
73. Stephan C, Jung K, Soosaipillai A, Yousef GM, Cammann H, Meyer H, et al. Clinical utility of human glandular kallikrein 2 within a neural network for prostate cancer detection. *BJU Int* 2005;96:521-7.
74. Stephan C, Meyer HA, Kwiatkowski M, Recker F, Cammann H, Loening SA, et al. A (-5, -7) Pro PSA Based Artificial Neural Network to Detect Prostate Cancer. *Eur Urol* 2006;50:1014-20.
75. Stephan C, Xu C, Brown DA, Breit SN, Michael A, Nakamura T, et al. Three new serum markers for prostate cancer detection within a percent free PSA-based artificial neural network. *Prostate* 2006;66:651-9.
76. Kuriyama M, Wang MC, Papsidero LD, Killian CS, Shimano T, Valenzuela L, et al. Quantitation of prostate-specific antigen in serum by a sensitive enzyme immunoassay. *Cancer Res* 1980;40:4658-62.
77. Ishibashi M. Standardization of prostate-specific antigen (PSA) assays: can interchangeability of PSA measurements be improved? *Clin Chem* 2006;52:1-2.
78. Hortin GL, Bahnson RR, Daft M, Chan KM, Catalona WJ, Ladenson JH. Differences in values obtained with 2 assays of prostate specific antigen. *J Urol* 1988;139:762-5.
79. Wu JT. Assay for prostate specific antigen (PSA): problems and possible solutions. *J Clin Lab Anal* 1994;8:51-62.
80. Stenman UH, Paus E, Allard WJ, Andersson I, Andres C, Barnett TR, et al. Summary report of the TD-3 workshop: characterization of 83 antibodies against prostate-specific antigen. *Tumour Biol* 1999;20 Suppl 1:1-12.
81. Graves HC. Issues on standardization of immunoassays for prostate-specific antigen: a review. *Clin Invest Med* 1993;16:415-24.

82. Semjonow A, Oberpenning F, Brandt B, Zechel C, Brandau W, Hertle L. Impact of free prostate-specific antigen on discordant measurement results of assays for total prostate-specific antigen. *Urology* 1996;48:10-5.
83. Jacobsen SJ, Lilja H, Klee GG, Wright GL, Jr., Pettersson K, Oesterling JE. Comparability of the Tandem-R and IMx assays for the measurement of serum prostate-specific antigen. *Urology* 1994;44:512-8.
84. Onur R, Ilhan N, Orhan I, Ilhan N. Increased discrimination between benign prostatic hyperplasia and prostate cancer with equimolar total prostate specific antigen measurement. *World J Urol* 2003;21:43-7.
85. Roddam AW, Price CP, Allen NE, Ward AM. Assessing the clinical impact of prostate-specific antigen assay variability and nonequimolarity: a simulation study based on the population of the United Kingdom. *Clin Chem* 2004;50:1012-6.
86. Corey E, Wegner SK, Stray JE, Corey MJ, Arfman EW, Lange PH, et al. Characterization of 10 new monoclonal antibodies against prostate-specific antigen by analysis of affinity, specificity and function in sandwich assays. *Int J Cancer* 1997;71:1019-28.
87. Stephenson RA, Greskovich FJ, Fritsche HA, James BC. Ratio of polyclonal-monoclonal prostate-specific antigen levels. Discrimination of nodal status in prostate tumors that produce low marker levels. *Urol Clin North Am* 1991;18:467-71.
88. Stenman UH, Leinonen J, Zhang WM. Problems in the determination of prostate specific antigen. *Eur J Clin Chem Clin Biochem* 1996;34:735-40.
89. Strobel SA, Sokoloff RL, Wolfert RL, Rittenhouse HG. Multiple forms of prostate-specific antigen in serum measured differently in equimolar- and skewed-response assays. *Clin Chem* 1995;41:125-7.
90. Laffin RJ, Chan DW, Tanasijevic MJ, Fischer GA, Markus W, Miller J, et al. Hybritech total and free prostate-specific antigen assays developed for the Beckman Coulter access automated chemiluminescent immunoassay system: a multicenter evaluation of analytical performance. *Clin Chem* 2001;47:129-32.
91. Murphy GP. The Second Stanford Conference on International Standardization of Prostate Specific Antigen Assays. *Cancer* 1995;75:122-8.
92. Segurado OG, Volmer W, Dowell B. PSA standardization: a review of NCCLS, Stanford and Abbott efforts. *Anticancer Res* 1997;17:2919-20.
93. Catalona WJ, Smith DS, Wolfert RL, Wang TJ, Rittenhouse HG, Ratliff TL, et al. Evaluation of percentage of free serum prostate-specific antigen to improve specificity of prostate cancer screening. *JAMA* 1995;274:1214-20.
94. Fox MP, Reilly AA, Schneider E. Effect of the ratio of free to total prostate-specific antigen on interassay variability in proficiency test samples. *Clin Chem* 1999;45:1181-9.
95. Cheli CD, Marcus M, Levine J, Zhou Z, Anderson PH, Bankson DD, et al. Variation in the quantitation of prostate-specific antigen in reference material: differences in commercial immunoassays. *Clin Chem* 1998;44:1551-3.
96. Blase AB, Sokoloff RL, Smith KM. Five PSA methods compared by assaying samples with defined PSA ratios. *Clin Chem* 1997;43:843-5.
97. Yurdakul G, Bangma CH, Blijenberg BG, van Zelst BD, Wildhagen MF, van der Kwast TH, et al. Different PSA assays lead to detection of prostate cancers with identical histological features. *Eur Urol* 2002;42:154-8.

98. Oberpenning F, Weining C, Brandt B, De AG, Heinecke A, Hamm M, et al. Combining free and total prostate specific antigen assays from different manufacturers: the pitfalls. *Eur Urol* 2002;42:577-82.
99. Ward AM, Catto JW, Hamdy FC. Prostate specific antigen: biology, biochemistry and available commercial assays. *Ann Clin Biochem* 2001;38:633-51.
100. Butch AW, Crary D, Yee M. Analytical performance of the Roche total and free PSA assays on the Elecsys 2100 immunoanalyzer. *Clin Biochem* 2002;35:143-5.
101. Kano S, Ishibashi M, Itoh Y. The present status of standardization in serum total PSA measurement--a result of "Survey 2000" conducted by PSA ad hoc Committee of Japanese Urological Association. *Rinsho Byori* 2001;49:967-73.
102. Patel D, White PA, Milford WA. A comparison of six commercial assays for total and free prostate specific antigen (PSA): the predictive value of the ratio of free to total PSA. *BJU Int* 2000;85:686-9.
103. Link RE, Shariat SF, Nguyen CV, Farr A, Weinberg AD, Morton RA, et al. Variation in prostate specific antigen results from 2 different assay platforms: clinical impact on 2304 patients undergoing prostate cancer screening. *J Urol* 2004;171:2234-8.
104. Filella X, Truan D, Alcover J, Quinto L, Molina R, Luque P, et al. Comparison of several combinations of free, complexed, and total PSA in the diagnosis of prostate cancer in patients with urologic symptoms. *Urology* 2004;63:1100-3.
105. Gordi T, Khamis H. Simple solution to a common statistical problem: interpreting multiple tests. *Clin Ther* 2004;26:780-6.
106. Passing H, Bablok. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. *J Clin Chem Clin Biochem* 1983;21:709-20.
107. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8:135-60.
108. Krouwer JS, Monti KL. A simple, graphical method to evaluate laboratory assays. *Eur J Clin Chem Clin Biochem* 1995;33:525-7.
109. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem* 1993;39:561-77.
110. Dewitte K, Fierens C, Stockl D, Thienpont LM. Application of the Bland-Altman plot for interpretation of method-comparison studies: a critical investigation of its practice. *Clin Chem* 2002;48:799-801.
111. Bruns DE, Huth EJ, Magid E, Young DS. Toward a checklist for reporting of studies of diagnostic accuracy of medical tests. *Clin Chem* 2000;46:893-5.
112. Jung K, Stephan C, Lein M, Henke W, Schnorr D, Brux B, et al. Analytical performance and clinical validity of two free prostate-specific antigen assays compared. *Clin Chem* 1996;42:1026-33.
113. 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.
114. Jung K, Lein M, Schnorr D, Brux B, Henke W, Loening S. Comparison between equimolar- and skewed-response assays of prostate specific antigen: is there an influence on the clinical significance when measuring total serum prostate specific antigen? *Ann Clin Biochem* 1996;33 ( Pt 3):209-14.

115. Roddam AW, Rimmer J, Nickerson C, Ward AM. Prostate-specific antigen: bias and molarity of commercial assays for PSA in use in England. *Ann Clin Biochem* 2006;43:35-48.
116. Tornblom M, Norming U, Adolfsson J, Becker C, Abrahamsson PA, Lilja H, et al. Diagnostic value of percent free prostate-specific antigen: retrospective analysis of a population-based screening study with emphasis on men with PSA levels less than 3.0 ng/mL. *Urology* 1999;53:945-50.
117. Catalona WJ, Ramos CG, Carvalhal GF, Yan Y. Lowering PSA cutoffs to enhance detection of curable prostate cancer. *Urology* 2000;55:791-5.
118. Catalona WJ, Partin AW, Finlay JA, Chan DW, Rittenhouse HG, Wolfert RL, et al. 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.
119. Ekins R. Immunoassay standardization. *Scand J Clin Lab Invest Suppl* 1991;205:33-46.
120. Leinonen J, Zhang WM, Paus E, Stenman UH. Reactivity of 77 antibodies to prostate-specific antigen with isoenzymes and complexes of prostate-specific antigen. *Tumour Biol* 1999;20 Suppl 1:28-34.
121. Wener MH, Daum PR, Brawer MK. Variation in measurement of prostate-specific antigen: importance of method and lot variability. *Clin Chem* 1995;41:1730-7.
122. Stenman UH. Immunoassay standardization: is it possible, who is responsible, who is capable? *Clin Chem* 2001;47:815-20.
123. Chen Z, Prestigiacomo A, Stamey TA. Purification and characterization of prostate-specific antigen (PSA) complexed to alpha 1-antichymotrypsin: potential reference material for international standardization of PSA immunoassays. *Clin Chem* 1995;41:1273-82.
124. Satterfield MB, Welch MJ. Comparison by LC-MS and MALDI-MS of prostate-specific antigen from five commercial sources with certified reference material 613. *Clin Biochem* 2005;38:166-74.
125. Tewari PC, Williams JS. Analytical characteristics of seminal fluid PSA differ from those of serum PSA. *Clin Chem* 1998;44:191-3.
126. Garg UC, Howanitz JH, Nakamura RM, Plous RH, Eckfeldt JH. Production, analysis, and characterization of reference materials for prostate-specific antigen. *Arch Pathol Lab Med* 1995;119:1104-8.
127. Schreiber WE, Endres DB, McDowell GA, Palomaki GE, Elin RJ, Klee GG, et al. Comparison of fresh frozen serum to proficiency testing material in College of American Pathologists surveys: alpha-fetoprotein, carcinoembryonic antigen, human chorionic gonadotropin, and prostate-specific antigen. *Arch Pathol Lab Med* 2005;129:331-7.
128. Morgan BR, Tarter TH. Serum heterophile antibodies interfere with prostate specific antigen test and result in over treatment in a patient with prostate cancer. *J Urol* 2001;166:2311-2.

**Erklärung**

Ich, Moritz Klaas, erkläre, dass ich die vorgelegte Dissertationsschrift mit dem Thema: „Analytische und klinische Validierung der fünf meistbenutzten Testkombinationen zur Bestimmung des totalen und freien prostata-spezifischen Antigens: Vergleichsmessungen an 798 Seren von Patienten mit Prostatakarzinom und benigner Prostatahyperplasie“ selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, ohne die (unzulässige) Hilfe Dritter verfasst und auch in Teilen keine Kopien anderer Arbeiten dargestellt habe.

Berlin, 08. Mai 2007

Moritz Klaas

## **Lebenslauf**

„Mein Lebenslauf wird aus Datenschutzgründen in der elektronischen Version meiner Arbeit nicht mit veröffentlicht.“

## **Danksagung**

Mein besonderer Dank gilt Herrn Prof. Dr. med. Klaus Jung für die hervorragende Betreuung dieser Arbeit. Ich möchte mich bei ihm für die stetige Unterstützung bei meiner wissenschaftlichen Ausbildung, die vielen anregenden Diskussionen und wertvollen Ratschläge bedanken.

Ich danke Herrn Dr. med. Carsten Stephan für die Betreuung der Arbeit und Hilfe bei der Erstellung des Artikels.

Besonders danken möchte ich Frau Janett Reiche (Medizinisch-technische Assistentin), die mich ganz wesentlich bei der technischen Durchführung unterstützte.

Ich danke allen Mitarbeiterinnen und Mitarbeitern der Forschungsabteilung der Klinik für Urologie der Charité, die mir mit wertvollen Hinweisen und freundlicher Unterstützung zur Seite standen.

Mein Dank gilt auch dem Institut für Laboratoriumsmedizin und Pathobiochemie unter Leitung von Prof. Dr. E. Köttgen für die Bereitstellung der Testsysteme.

Mein herzlicher Dank gilt meinen Eltern für ihre ständige Unterstützung. Meiner Familie und allen Freunden sei herzlich gedankt für ihr Interesse und ihre ständige Anteilnahme an meiner Arbeit. Für das Lesen der Arbeit und die nützlichen Hinweise möchte ich Philipp Klaas und Tom Rojo Poller danken. Meiner Frau Sunčica danke ich für ihr liebevolles Verständnis.