

Voraussetzung. Denn in der Praxis ist es ja gerade die Frage, dass z.B. der tPSA-Wert im Bereich von 4-10 µg/l keine Differenzierung zwischen PCa- und BPH-Patienten ermöglicht. Der statistische Nachweis eines diagnostischen Vorteils des neuen Parameters gelingt damit unter diesen Umständen häufig nicht. Aus diesem Grund wurde das Matching-Verfahren von unserer Arbeitsgruppe erarbeitet. Hierbei wird durch Zuordnung von BPH-Patienten zu PCa-Patienten mit weitgehend identischen tPSA-Werten eine AUC von etwa 0,5 für das tPSA angestrebt [77]. Dadurch kann jeder neue Parameter unter identischen Bedingungen mit dem tPSA verglichen werden [77].

Bei der Bewertung des %fPSA im Vergleich zum tPSA hinsichtlich der Differenzierung zwischen PCa und BPH werden dazu neben der AUC auch die Spezifitäten z.B. bei 80%, 90% oder 95% Sensitivität verglichen. Hierfür sind mittlerweile verschiedene Computerprogramme sowohl frei im Internet als auch kommerziell erhältlich. In einem Vergleich haben wir 8 Programme für die ROC-Analyse hinsichtlich verschiedener Kriterien (Richtigkeit, Bedienbarkeit, Signifikanzangaben, Grafikausgabe usw.) getestet (**Artikel Nr. E5**). Neben dem Programm GraphROC, das von uns bereits seit 1996 eingesetzt wird, hat sich hierbei vor allem das Programm MedCalc als sehr praktikabel erwiesen [78]. Es wird inzwischen in unserer Arbeitsgruppe bevorzugt genutzt [78].

Eine weitere Möglichkeit, zwei Parameter vor allem innerhalb eines begrenzten tPSA-Bereiches zu bewerten, ist die sogenannte Diskordanz-Analyse [79]. Damit wird ein exakter Vergleich z.B. des tPSA mit dem cPSA möglich, da nicht nur der Grenzwert des tPSA, sondern auch der des cPSA in die Bewertung einfließt [79]. Mit Hilfe dieses Verfahrens [79] haben wir inzwischen durch Auswertung weiterer Studiendaten belegt, dass cPSA, bedingt durch seinen signifikanten Vorteil gegenüber dem tPSA innerhalb eines begrenzten tPSA-Bereiches von 3-5 µg/l, als Initialparameter beim PCa-Screening gegenüber dem tPSA Vorteile bietet [80].

In diesem Kapitel habe ich die diagnostische Bewertung des PSA und seiner Subformen als Einzelbefunde erläutert. Im folgenden Kapitel werde ich das Arbeiten mit einem sogenannten artifiziellen neuronalen Netzwerk darstellen, mit dem gewissermaßen eine Integration der Einzelbefunde zu einem Gesamtbefund erfolgt. Damit kann eine deutlich verbesserte PCa-Diagnostik erzielt werden. Erst kürzlich haben wir dazu auf Bitte des Herausgebers der neuen Zeitschrift *Nature Clinical Practice Urology* in einem „Viewpoint“ den augenblicklichen Stand und die zukünftige Entwicklung für artifizielle neuronale Netzwerke in der Urologie aufgezeigt [81].

4. Artifizielle neuronale Netzwerke und neue Serummarker für das PCa

4.1 Aufbau artifizieller neuronaler Netzwerke

Die ersten Versuche, mit nicht-linearen Methoden Populationen aufgrund verschiedener Charakteristika zu trennen, wurden bereits in den 50er Jahren unternommen. Aufgrund des hohen technischen Aufwandes erfuhren diese Methoden lange Zeit keine breite klinische Anwendung. Erst die Entwicklung der Computertechnik führte zu einem sprunghaften Anstieg des Einsatzes nicht-

linearer Klassifikationsmethoden, insbesondere der künstlichen neuronalen Netzwerke (ANN). Besonders in den Fachgebieten der Radiologie (Mammografie, Bewertung von Lungenherden), der Kardiologie (Infarkt- und EKG-Diagnostik), der Onkologie und Pharmakologie wurden Studien zur Anwendung von ANN erfolgreich durchgeführt.

Strukturell ist ein ANN mit einem Netzwerk menschlicher Neuronen zu vergleichen, mit dessen Hilfe von der Funktion her Lernprozesse nachgebildet werden. Dabei fließt die Information in der Regel nur in eine Richtung (Abbildung 2), nämlich von den Eingangsneuronen zu den Zwischenneuronen und letztendlich zu dem(n) Ausgangsneuron(en). Die Zwischenneuronen können die Information wichten und ihre Weiterleitung veranlassen. Im Falle einer Mustererkennung (Klassifikationsaufgabe) führt dann die Verknüpfung unterschiedlich gewichteter Eingangssignale zu einem Wahrscheinlichkeitswert für das Vorliegen bestimmter Muster. Die Leistung dieser Netzwerke beruht dabei im Wesentlichen mehr auf der Komplexität der Verschaltungen als auf der Komplexität der Neuronen selbst. Somit sind ANN in ihrer Struktur und Funktion den neuronalen Netzen im menschlichen Gehirn ähnlich.

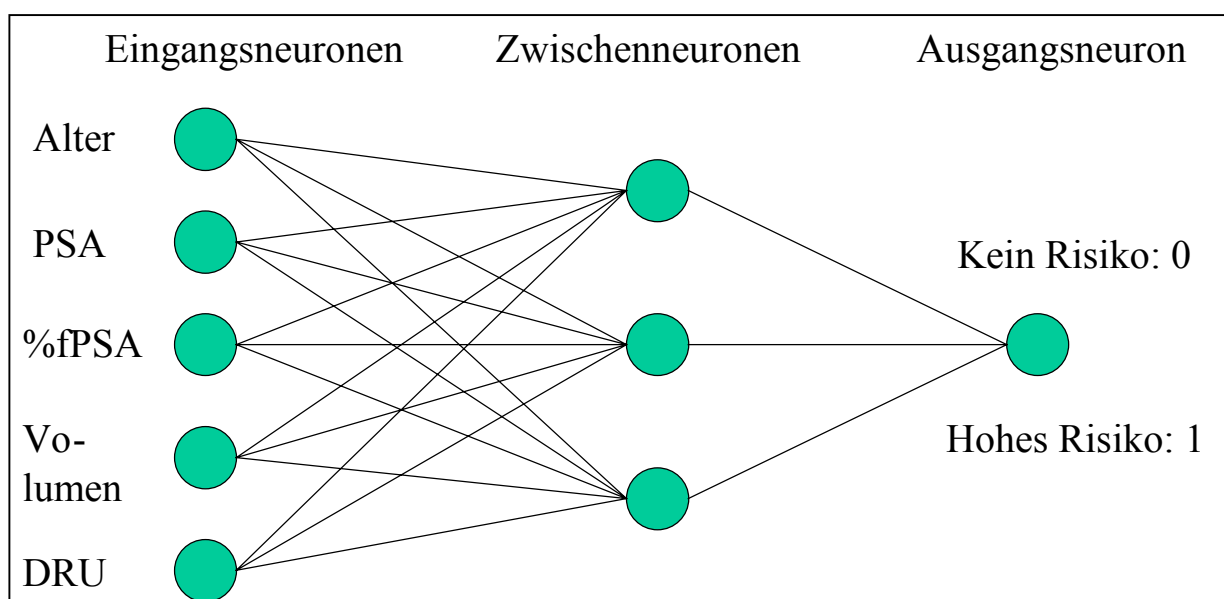


Abbildung 2: Aufbau des genutzten künstlichen neuronalen Netzwerks (ANN)

Während eines Trainingsprozesses werden die Gewichte im Hinblick auf eine optimale Mustererkennung verändert. Dabei lassen sich die ANN anhand der verwendeten Trainingsstrategien in zwei große Klassen einteilen, die der kontrollierten und unkontrollierten Lernmethoden. Bei den unkontrollierten Lernmethoden erhält das ANN nur die Eingangsinformationen und sucht dann selbstständig nach verborgenen Beziehungen der Eingangsvariablen. Im Gegensatz dazu erfolgt das kontrollierte Lernen der ANN mittels Eingabe der Eingangs- als auch der Ausgangsinformationen. Während der Trainingsphasen werden ständig die aktuellen Ausgangswerte des ANN (Istwerte) und Zielgrößen (Sollwerte) miteinander verglichen und abhängig davon wird die Gewichtung der

Schaltstellen verändert. Eines der am häufigsten verwendeten ANN ist das Multilayer Perceptron Netzwerk. Ein in diesem Aufbau von uns verwendetes Netzwerk besteht aus 5 Eingangsneuronen, die mit 3 Neuronen der Zwischenschicht (sogenannte hidden layer) und letztendlich mit dem Ausgangsneuron verbunden sind (Abbildung 2). Zu Beginn der Berechnung werden den Verschaltungen zwischen den Neuronen zufällige Gewichtungen zugeteilt. Anschließend wird nach Anlegen eines Datensatzes die Größe der Abweichung von Ist- und Sollwert verwendet, um die Gewichte der Neuronen zu modifizieren. In der Folge wird dem ANN eine große Zahl an Fällen präsentiert und bei jedem Trainingsdurchlauf eine Korrektur vorgenommen. Dieser Korrekturmechanismus wird „Backpropagation“ genannt, da die Fehler vom Ausgang in Richtung Eingang zurückgegeben werden und zur Korrektur der Gewichte an den Schaltstellen verwendet werden.

Der Stichprobenumfang für die Erstellung eines ANN ist von der Anzahl der Eingangsvariablen abhängig. Dabei muss die Anzahl der Datensätze mit abnehmender Anzahl an Eingangsvariablen zunehmen, um ein gut validiertes ANN zu erhalten. Nach der Trainingsphase ist es notwendig, das erstellte Netz mit einer bis dahin nicht dem ANN präsentierten Datenmenge zu testen, um Aussagen über die Validität des ANN machen zu können. Bei großen Datenmengen werden von den Datensätzen der Gesamtstichprobe ca. 70% zum Trainieren und 30% zum Testen verwendet. Bei kleineren Datenmengen wird die Methode der k-fachen Kreuzvalidierung verwendet, bei der die Gesamtanzahl n der Datensätze in k Untermengen geteilt wird. In k Trainingsschritten wird jeweils eine der k Untermengen zum Testen der mit den übrigen $k-1$ Untermengen trainierten ANN verwendet. Die Validität des aus den n Datensätzen gebildeten endgültigen ANN wird durch Zusammenfassen der k Klassifikatoren bestimmt. Der Extremfall der Kreuzvalidierung ist das „leave-one-out“-Verfahren, das bei kleineren Datenmengen eingesetzt wird. Hier wird n mal ein ANN mit $n-1$ Datensätzen trainiert und anschließend wird der zurückbehaltene Datensatz getestet. Die letztere Methode wird in unseren aktuellen Studien ausnahmslos verwendet. Damit ist bei Subgruppen mit relativ geringer Probandenanzahl eine ausreichende Datenmenge zum Training vorhanden und bei allen ANN-Berechnungen wird die gleiche Methode unabhängig vom Stichprobenumfang verwendet.

Der Vorteil der Netzwerke liegt darin, dass auch bei vielen Eingangsvariablen mit nicht-linearen Relationen ein adäquater Klassifikator gebildet werden kann, mit dessen Hilfe eine optimale Mustererkennung möglich ist. Dies ist mit herkömmlichen Methoden, wie der linearen Regression oder linearen Diskriminanzanalyse, oftmals nicht der Fall. Eine neuere Untersuchung [82], in der bei einem Vergleich von Klassifikationsleistungen die ANN in vielen Fällen günstiger abschneiden als die oftmals verwendete logistische Regression (LR), bestätigt unsere Literaturbewertung. Im Folgenden werden die ersten Anwendungen von multivariaten Modellen und insbesondere die Anwendung von ANN bei der Diagnostik des PCa beschrieben.

4.2 Die Anwendung von ANN zur Diagnostik des PCa

Im Jahr 1994 demonstrierten Snow et al. [83] den Nutzen eines ANN für die Prädiktion eines späteren Prostatakarzinomprogresses anhand des PSA und anderer präoperativer Parameter. 1998 wurde erstmals über die Einbeziehung des %fPSA in ein LR-Modell berichtet [84]. Mit den drei Eingangsvariablen PSA, %fPSA und Patientenalter konnte eine 11%ige Verbesserung der Spezifität gegenüber dem %fPSA erzielt werden [84]. Finne et al. [85] zeigten im PSA-Bereich 4-10 µg/l unter Einbeziehung der Variablen PSA, %fPSA, Prostatavolumen und DRU bei vorgegebener Sensitivität von 95% eine Verbesserung der Spezifität von 19% (alleinige Nutzung des %fPSA) auf 33% mit Hilfe eines ANN. Eine andere Arbeitsgruppe wies mit der Kombination aus drei Netzwerken bei vorgegebener Sensitivität von 92% eine Verbesserung der Spezifität um 49% nach [86]. Djavan et al. [87] erarbeiteten ein ANN für den Bereich 2,5-10 µg/l, wobei für PSA-Werte bis 4 µg/l eine 95%ige Spezifität und bei Werten von 4-10 µg/l eine 95%ige Sensitivität zugrunde gelegt wurden. An diesen Grenzwerten wurde durch das jeweilige ANN eine Sensitivitäts- bzw. Spezifitätsverbesserung von über 25% erzielt [87]. Seit 2003 wurden weitere ANN- und LR-Studien z.T. unter Einschluss anderer Kriterien, wie Familienanamnese, ethnischer Zugehörigkeit, des Internationalen Prostata Symptome Score oder verschiedener TRUS-Variablen veröffentlicht [88-90]. Alle Arbeiten zeigten eine verbesserte Spezifität gegenüber den Einzelparametern. Die eigenen Arbeiten zur Erstellung und Anwendung von ANN sind im folgenden Kapitel dargestellt.

4.3 Eigene Arbeiten zur Nutzung von ANN bei der Diagnostik des PCa

4.3.1 Entwicklung des Programms „ProstataClass“

Aufbauend auf erste eigene Anwendungsstudien des %fPSA zur Verbesserung der Biopsieindikation unter Berücksichtigung des Prostatavolumens und des tPSA-Bereiches [53, 59, 61] erfolgte in Zusammenarbeit mit dem Institut für Medizinische Biometrie der Charité Campus Mitte (Dr. Cammann) eine Untersuchung an 928 Männern [91]. Die Variablen tPSA, %fPSA, Lebensalter, Prostatavolumen und der Status der DRU wurden einbezogen (**Artikel Nr. E6**). Im Rahmen der Auswertung dieser fünf Eingangsvariablen wurde ein ANN entwickelt, das die Wahrscheinlichkeit für das Vorliegen eines PCa berechnet [91]. Auf dieser Grundlage wurden für verschiedene tPSA-Bereiche Algorithmen abgeleitet, die als gezielte Entscheidungshilfen für die Durchführung der Prostatastanzbiopsie herangezogen werden können [91]. Die Entwicklung dieses ANN durch Nutzung der von 1996 bis 2001 kontinuierlich erhobenen Daten ist die konsequente Weiterführung unseres Anliegens, bessere Entscheidungshilfen zur Biopsieindikation zu erarbeiten. Die Bedeutung dieser Publikation [91] wird auch dadurch sichtbar, dass der Artikel als „Spotlight“ in der Ausgabe des *International Journal of Cancer* besonders hervorgehoben wurde (siehe Beiblatt zu Artikel Nr. E6).

Auf der Basis dieser Kombination von klinischen Daten und klinisch-chemischen Parametern wurde das Computerprogramm „ProstataClass“ entwickelt, welches seit 2002 auf der Webseite der

Klinik für Urologie der Charité (www.charite.de/ch/uro) als Download frei verfügbar ist [91]. Es wird seitdem in der klinischen Routine unserer Klinik verwendet (Abbildung 3). Seit September 2003 kann das Risikobewertungsprogramm auch direkt online wahlweise in Deutsch oder Englisch genutzt werden. Außerdem sind seit 2005 auch die Erklärungen und Zusatzinformationen in der englischen Sprache erhältlich. Hiermit wird eine optimale Anwendung im klinischen Alltag sowie die internationale Umsetzung der ausgesprochenen Biopsieempfehlungen ermöglicht.

Risikoabschätzung für Prostatakrebs
Artifizielles neuronales Netz
(ProstataClass HCS2004/01)

Klinik für Urologie Charité - Universitätsmedizin Berlin
ProstataClass HCS 2004/01

Risikoabschätzung für Prostatakrebs Patient: XY
Datum: 6/2/2004

	beobachtete Werte	(mw, s) des Trainingsdatensatzes
Alter [Jahre]	60.0	(64.9, 7.11)
Gesamt-PSA [ng/ml]	6.0	(6.8, 1.68)
%-fPSA	16.7	(13.3, 8.27)
Prostatavolumen [ml]	25.0	(37.5, 19.22)
DRU	0	(0.4, 0.50)

Sens. 95 %
Schwelle: 37.0

ANN-Ausgang 31.5
"geringes Risiko"

PSA-Werte gemessen mit den IMMULITE Systemen der Diagnostic Products Corporation (DPC).

2-4 Spezifität: 90 % 95 %
4-10 Sensitivität: 90 % 95 %
10-20 Sensitivität: 90 % 95 %

Bild drucken Neuer Patient Exit

Abbildung 3: Layout des Programms „ProstataClass“

(modifiziert nach einer Abbildung von Stephan et al. in Clin Chem 2002;48:1279-87)

Zahlreiche Vorträge und Posterpräsentationen auf regionalen und internationalen Konferenzen belegen die Bemühungen, diese Entscheidungshilfe für eine bessere Patientenauswahl zur Prostatastanzbiopsie auch anderen Einrichtungen zugänglich zu machen. Dieses ist der Vorteil des für den gesamten tPSA-Bereich 2-20 µg/l gültigen und für jeden frei zugänglichen Computerprogramms „ProstataClass“ [91] im Vergleich zu anderen veröffentlichten Arbeiten [85-87]. Kürzlich erfolgte deshalb meine Berufung in die „International Task Force on Early Diagnosis of Cancer“ der European School of Oncology. In diesem Gremium ist die Ausarbeitung von Richtlinien zur Früherkennung des PCa zusammen mit Jonas Hugosson (Schweden) mein Verantwortungsbereich [92]. Eine aktuelle

Kurzübersicht in der Zeitschrift *Nature Clinical Practice Urology* [81] sowie aktuelle Buchbeiträge [93, 94] zeigen das hohe wissenschaftliche Interesse am Einsatz von ANN bei der PCa-Diagnostik.

In Fortführung des ersten ANN-Projekts habe ich eine multizentrische Studie mit insgesamt fast 1200 Patienten initiiert und erfolgreich durchgeführt [95]. Diese Studie zeigt, dass durch die Nutzung dieses ANN etwa 20-30% der Biopsien vermeidbar sind (**Artikel Nr. E7**). Zusätzlich kann bei den Patienten, die bisher aufgrund nur eines Kriteriums ($\text{PSA} < 4 \mu\text{g/l}$) als unverdächtig hinsichtlich eines PCa eingestuft werden, das Karzinomrisiko besser abgeschätzt werden. Wurden diese Patienten biopsiert, war die Karzinomentdeckungsrate mit 46% deutlich über der allgemein zu erwartenden 20-25%igen Entdeckungsrate [95]. Das ANN-Programm „ProstataClass“ wurde zusätzlich mit weiteren prospektiven Patientendaten validiert und die Ergebnisse in der führenden deutschen urologischen Zeitschrift *Der Urologe* publiziert [96]. Vor der endgültigen Entscheidung für eine Biopsie sollte allerdings bei Daten nahe des ANN-Grenzwertes immer eine Wiederholung der tPSA- und fPSA-Messung erfolgen [97].

Die Anwendung des Programms „ProstataClass“ erfolgt in unserer Klinik von Beginn an mit dem seit 1996 in die klinische Routine eingeführten Immulite-Testsystem der Fa. DPC. Dieses Testsystem war als eines der ersten für das fPSA kommerziell erhältlich [40]. Bedingt durch die Vielzahl von verschiedenen tPSA- und fPSA-Testsystemen, deren Ergebnisse z.T. erheblich variieren können [98], zeigten eigene Voruntersuchungen relevante Abweichungen in den ANN-Biopsieempfehlungen. Allein bedingt durch ein anderes tPSA- und fPSA-Testsystem werden bis zu 10% der Patienten mit PCa falsch klassifiziert.

Deshalb wurde vor der geplanten Erweiterung des Programms „ProstataClass“ für andere häufig genutzte Testsysteme eine Studie mit fünf verschiedenen tPSA- und fPSA-Testsystemen und insgesamt 4480 Patienten aus fünf Zentren durchgeführt. Das Originalprogramm „ProstataClass“ als auch PSA-Test spezifische ANN wurden im Vergleich zum %fPSA bewertet [99]. In dieser Studie war das ANN des Originalprogramms „ProstataClass“ in 11 von 15 Vergleichen mittels ROC-Analyse (AUC) in den verschiedenen tPSA-Bereichen dem %fPSA unabhängig vom Testsystem überlegen [99]. Allerdings zeigte die Berechnung der PSA-Test spezifischen ANN einen signifikanten Vorteil gegenüber dem Originalprogramm „ProstataClass“ in 12 von 15 Vergleichen anhand der AUC-Werte und in 14 von 15 Grenzwertvergleichen bei 95% Sensitivität. Diese Ergebnisse sprechen dafür, dass eine optimale Bewertung der Befunde nur mit einem testspezifischen ANN für das jeweilige PSA-Testsystem zu erzielen ist [99]. Da eine endgültige Bewertung der ANN mit verschiedenen PSA-Testsystemen nur durch einen Vergleich in derselben Population erfolgen kann, wurden parallele Messungen mit den fünf zur Zeit am häufigsten genutzten Testssystemen an ca. 800 Patienten durchgeführt (**Artikel Nr. E8**). Diese Daten zeigen Unterschiede in den Medianen des %fPSA bis zu 10%, die u.a. durch unterschiedliche Kalibrationen und nicht-äquimolare Bestimmung von PSA-Formen zu erklären sind [100]. Eine Bewertung innerhalb der verschiedenen ANN soll in Kürze die Anwendung des Programms „ProstataClass“ für andere PSA-Testsysteme ermöglichen.

Neben der Orientierung auf die klinisch-praktische Anwendung des ANN-Programms für weitere PSA-Testsysteme ist auch die mögliche Integration neuer Marker und ggf. die Umstellung des ANN nur auf laborchemische Parameter zur besseren Objektivierbarkeit der Befunde ein weiterer Schwerpunkt meiner Forschungstätigkeit. Dieser Schwerpunkt wird im folgenden Kapitel dargelegt.

4.3.2 ProPSA als neuer Serummarker im ANN

Ein wesentlicher Punkt für den nutzbringenden Einsatz neuer Tumormarker innerhalb eines ANN ist der Nachweis, dass der Parameter auch bei alleiniger Anwendung einen Zugewinn in der PCa-Diagnostik erbringt. Der Übersichtlichkeit halber soll dies am Beispiel des proPSA demonstriert werden, obwohl die fPSA-Subformen bereits im Kapitel 2 (ab Seite 8) beschrieben wurden.

Mehreren Arbeitsgruppen gelang der Beweis, dass proPSA-Formen sowohl die Unterscheidung zwischen PCa und BPH als auch besonders die Entdeckung der aggressiven Formen des PCa verbessern kann [101-105]. Eine von uns initiierte multizentrische Studie zur Bewertung der -5 und -7-Formen des proPSA mit einem Testkit der Fa. Roche an 1727 Patienten im tPSA-Bereich von 0-10 µg/l ergab keinen Vorteil des proPSA gegenüber dem %fPSA [106]. Eine andere Studie stellte zwar einen signifikanten Unterschied mit denselben proPSA-Formen zwischen PCa- und BPH-Patienten fest, es ergab sich aber ebenfalls keine zusätzliche Verbesserung zum %fPSA [107].

In einer gemeinsamen Studie mit einer Schweizer Arbeitsgruppe wurden von 898 Patienten mit tPSA-Werten von 1-10 µg/l die Serumkonzentrationen von proPSA, fPSA, tPSA sowie die klinischen Befunde Prostata-Volumen und Status der DRU bewertet [108]. Dabei wies das -5 und -7 proPSA nur im tPSA-Bereich 4-10 µg/l einen signifikanten Unterschied zwischen PCa- und BPH-Patienten auf [108]. Innerhalb einer Subgruppe erreichte ein ANN nur mit den Laborparametern tPSA, %fPSA und proPSA denselben AUC-Wert (0,84) wie das klassische „ProstataClass“-ANN mit den zusätzlichen klinischen Parametern Prostata-Volumen und Status der DRU, aber ohne proPSA [108]. Diese Resultate des proPSA-basierten ANN spiegeln sich allerdings nicht am klinisch wichtigen Grenzwert bei 95% Sensitivität wider [108]. Somit bleibt das proPSA nur in Ausnahmefällen eine zusätzliche Entscheidungshilfe zur verbesserten Unterscheidung zwischen PCa und BPH. Diese Schlussfolgerung untermauern auch die Resultate von Bangma et al. [107].

Erste Untersuchungen des -2 und -4 proPSA mit einem von der Fa. Beckman Coulter entwickelten Test zeigten, dass möglicherweise diese Formen eine höhere diagnostische Aussagekraft als die -5 und -7 proPSA-Formen haben [101, 102, 104]. Eine endgültige Bewertung steht aber noch aus.

4.3.3 Humanes Kallikrein 11 (hK11), Makrophagen Inhibitor Cytokin-1 (MIC-1) und Makrophagen Migrationsinhibitor Faktor (MIF) als neue Serummarker des PCa

Das humane Kallikrein 11 (hK11) gehört ebenso zur Gruppe der Kallikreine wie das PSA, das auch als humanes Kallikrein 3 (hK3) bezeichnet wird. Kallikreine sind Serinproteasen, welche große Makromoleküle in kleinere Peptide spalten. Bis vor ca. 6 Jahren ging man davon aus, dass die Gruppe der Kallikrein-Gene nur 3 Mitglieder hat. Dabei kodieren das pankreatisch-renale Kallikrein (KLK1), das humane glanduläre Kallikrein 2 (KLK2) und das KLK3 für die entsprechenden Proteine hK1, hK2 und PSA (hK3). Mittlerweile konnten 12 weitere Kallikrein-Gene, namentlich KLK4 bis KLK15 und die entsprechenden Proteine hk4 bis hk15 entdeckt werden. Ein wesentlicher Anteil dieser neu entdeckten Kallikreine geht auf Arbeiten meines Kooperationspartners Prof. Diamandis (Toronto, Kanada) zurück [109]. Eine Übersicht über die auf dem Gen 19q 13.4 lokalisierten Kallikrein-Loci zeigt die Abbildung 4.

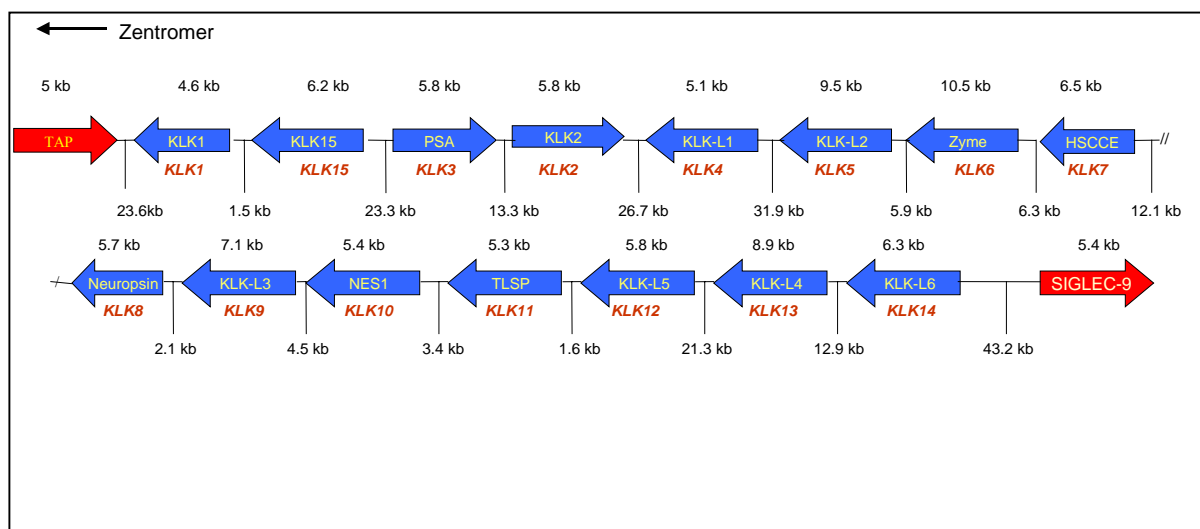


Abbildung 4: Alle Kallikrein-Gene auf Chromosom 19, Region 19q13.4, Gesamtlänge ca. 300 kb (modifiziert nach einer Abbildung von Stephan et al. in Urology 2002;59:2-8)

Im Jahr 2002 gelang erstmals die Isolation des hK11 [110]. Die höchsten Konzentrationen wurden mittels Immunoassay im Prostatagewebe und im Seminalplasma nachgewiesen [110]. Das hK11-Protein wurde dabei als potentieller Marker für das PCa postuliert, da eine signifikante Erhöhung im Serum bei 60% aller PCa-Patienten vorhanden war [110]. Ergebnisse einer anschließenden Studie an 150 Patienten der Charité mit BPH und PCa zeigten, dass hK11 (AUC: 0,81) ein ähnliches Potential wie das %fPSA (AUC: 0,82) hat, um zwischen PCa und BPH zu differenzieren [111]. Die Kombination des %fPSA und tPSA mit dem Parameter hK11 ermöglicht eine weitere Einsparung an unnötigen Biopsien [111]. Etwa bei der Hälfte aller BPH-Patienten mit einem suspekten %fPSA-Wert <20% kann durch die Einbeziehung der hK11/tPSA-Ratio die Prostatabiopsie vermieden werden [111].

Als weiterer neuer potentieller Marker für das PCa wird das Makrophagen Inhibitor Cytokin-1 (MIC-1) angesehen [112]. MIC-1 ist Mitglied in der Familie der Wachstumsfaktoren. Das Molekül wurde 1997 zuerst in aktivierten Makrophagen entdeckt, konnte aber auch in vielen anderen Geweben wie z.B. Niere, Pankreas, Plazenta und Prostata nachgewiesen werden [113]. Ein Serumtest für MIC-1 wurde von der kooperierenden australischen Arbeitsgruppe unter Leitung von Prof. Breit (Sydney) entwickelt [114, 115]. Im Rahmen der Kooperation zeigten sich in 1000 Serumproben von PCa- und BPH-Patienten der Klinik für Urologie der Charité beim MIC-1 signifikante Unterschiede zwischen beiden Patientengruppen [112]. In Zukunft könnte mit Hilfe des MIC-1 in Kombination mit %fPSA eine wesentliche Anzahl von Biopsien zusätzlich eingespart werden [112]. Außerdem scheint ebenfalls eine bessere Unterscheidung von aggressiven (Gleason Score >7) und weniger aggressiven Tumoren unter Verwendung des MIC-1 möglich zu sein [112].

Der Makrophagen Migrationsinhibitor Faktor (MIF) wurde ebenfalls als Marker für das PCa beschrieben. Als Produkt der T-Lymphozyten, welches Makrophagen aktiviert, wurde das Zytokin bereits vor ca. 40 Jahren erstmals nachgewiesen. Dabei wurde MIF mit verschiedenen Funktionen, wie der Regulierung von entzündlichen Prozessen und der Immunantwort, der Induktion der Zellproliferation, der Angiogenese oder der Hemmung von Tumorsuppressorgenen in Verbindung gebracht [116]. Eine Hochregulierung der mRNA des MIF in Prostataepithelzellen des PCa im Vergleich zu normalen Zellen wurde 1998 aufgezeigt [117]. Dieselbe Arbeitsgruppen entwickelte einen Serum-Test für MIF und wies eine höhere Konzentration in PCa-Patienten im Vergleich zu BPH- und gesunden Kontrollpatienten nach [118]. Ermutigt durch diese Resultate, testeten wir an einer größeren Gruppe von PCa- und BPH-Patienten die diagnostische Aussagefähigkeit von MIF. Wir fanden allerdings im Gegensatz zu den genannten Ergebnissen niedrigere Konzentrationen bei PCa-Patienten [119]. Eine weitere Analyse an einer anderen eigenen Patientengruppe bestätigte die geringeren MIF-Konzentrationen beim PCa und zeigte keine Korrelation zum tPSA [120].

4.3.4 MIC-1, MIF und hK11 als neue Serummarker des PCa im ANN

In weiteren Untersuchungen war von Interesse, wie der Einschluss der Ergebnisse dieser drei neuen Serummarker hK11, MIC-1 und MIF [110, 112, 118-120] die diagnostische Aussagekraft eines tPSA- und %fPSA-basierten ANN beeinflusst. Die Hypothese einer Spezifitätsverbesserung wurde in einer eigenen aktuellen Studie (**Artikel Nr. E9**) an 371 Patienten im relevanten tPSA-Bereich 0,5-20 µg/l überprüft [121]. Die Analyse ergab für alle drei neuen Marker als Einzelparameter (AUC: 0,56-0,60) keine Verbesserung gegenüber dem tPSA (AUC: 0,77) oder %fPSA (AUC: 0,81). Ein ANN mit den Eingangsvariablen hK11, MIC-1, MIF, tPSA und %fPSA zeigte aber bezüglich der AUC (0,86) eine signifikante Verbesserung gegenüber dem %fPSA [121]. An den klinisch wichtigen Sensitivitäts-Grenzwerten von 90 und 95% ergab allerdings nur ein ANN mit der zusätzlichen Eingangsvariable Prostatavolumen einen signifikanten Vorteil gegenüber dem %fPSA [121]. Dabei war eine bemerkenswerte Steigerung der Spezifität von 30% auf 80% bei 90% Sensitivität bzw. von

24% auf 40% bei 95% Sensitivität zu verzeichnen [121]. Trotz geringerer diskriminatorischer Trennschärfe als tPSA tragen alle drei neuen Tumormarker, die z.T. kaum mit dem tPSA korrelierten, innerhalb eines ANN zur Verbesserung der PCa-Diagnostik bei. Wahrscheinlich kann nur mit Hilfe neuer Marker für das PCa die Diskrepanz zwischen einer Überdiagnostik einerseits und einer Tumorrezidivquote von ca. 30% andererseits gelöst werden.

4.4 Humanes glanduläres Kallikrein 2 (hK2) als Tumormarker für das PCa

Ein weiterer Marker aus der Gruppe der Kallikreine ist das humane glanduläre Kallikrein 2, welches in eigenen Studien im Gewebe [122] und als Serumparameter zur PCa-Detektion innerhalb eines ANN [123] sowie zur Beurteilung der Tumorstadien [124] herangezogen wurde.

Das KLK2 (Protein hK2) als direkter Genomnachbar des PSA-Gens auf Chromosom 19q13.4 hat eine 80%ige bzw. 78%ige Übereinstimmung zum PSA auf der DNA- und Protein-Ebene. Trotz deutlich geringerer Serumkonzentrationen des hK2 im Vergleich zum PSA zeigten die ersten Bestimmungen in Seren von PCa- und BPH-Patienten eine signifikant höhere Spezifität gegenüber der alleinigen Ermittlung des PSA bzw. des %fPSA [125-127]. In mehreren Studien wurde durch die Quotienten hK2/fPSA bzw. hK2/%fPSA die Trennschärfe zwischen Patienten mit BPH und PCa verbessert [125-127]. In einem Übersichtsartikel der Zeitschrift *Urology* zur PSA-Thematik [49] diskutierte ich neben der bedeutenden Rolle des %fPSA auch ausführlich die vielversprechenden Ergebnisse des hK2 v.a. zur Detektion aggressiver Karzinome im tPSA-Bereich $<4 \mu\text{g/l}$ (**Artikel Nr. E10**).

In einem Kooperationsprojekt mit dem Kallikrein-Labor von Prof. Diamandis aus Toronto [123] wurde an 475 BPH- und PCa-Patienten aus der Charité der dort entwickelte Serumtest für das hK2 weiter validiert (**Artikel Nr. E11**). In dieser Studie zeigte sich der Nutzen des hK2 im ANN v.a. im niedrigen tPSA-Bereich von $1-4 \mu\text{g/l}$ und ergab im Vergleich zum %fPSA signifikant bessere Werte [123]. Im tPSA-Bereich von $4-10 \mu\text{g/l}$ war mit dem ANN allerdings kein Vorteil gegenüber dem %fPSA oder dem Quotienten hK2/%fPSA nachweisbar [123]. Da in allen anderen tPSA-Bereichen außer $1-4 \mu\text{g/l}$ kein Vorteil des ANN gegenüber dem %fPSA an den Grenzwerten der 90- und 95-%igen Sensitivität festzustellen war, ist der Einsatz eines hK2-basierten ANN aus klinischer Sicht abwartend zu bewerten [123].

In weiteren Pilotstudien trennte hK2 signifikant zwischen gut bis mäßig differenzierten Karzinomen (G1 und G2) und schlecht differenzierten G3-Tumoren [128] sowie zwischen lokal begrenzten (pT2) und fortgeschrittenen (pT3-pT4) Tumoren [129, 130]. Diese Ergebnisse führten zu der Hypothese, dass hK2 als mögliches präoperatives Vorhersagekriterium in Bezug auf das Tumorstadium und den histologischen Differenzierungsgrad beim PCa eingesetzt werden könnte [128-130]. In einer eigenen Studie an 222 radikal prostatektomierten PCa-Patienten waren die hK2-Werte sowohl zwischen G2 und G3 als auch zwischen Gleason Score <7 und ≥ 7 Tumoren nicht signifikant unterschiedlich [124]. Die Quotienten hK2/fPSA und hK2/%fPSA unterschieden sich allerdings

analog zum %fPSA zwischen den Tumorstadien pT2 und pT3 [124]. Obwohl in einer aktuellen Untersuchung das hK2 sogar als Indikator eines biochemischen PSA-Rezidivs empfohlen wurde [131], sehe ich keinen Anlass, hK2 zur Differenzierung der Tumorstadien bzw. der histologischen Differenzierungsgrade einzusetzen.

Frühere Untersuchungen zur Expression des KLK2-Gens postulierten eine Überexprimierung im PCa und auch in Metastasen des PCa im Vergleich zum BPH- oder normalen Prostatagewebe [132, 133]. Damit wurden die anfänglich ausschließlich positiv bewerteten hK2-Studien im Serum begründet. Gemeinsam mit dem kanadischen Kooperationspartner haben wir hingegen eine Herunterregulierung im PCa-Gewebe im Vergleich zum normalen Prostatagewebe nachgewiesen [122]. Das PSA war davon stärker betroffen als das hK2 [122]. Wir konnten damit beweisen, dass unterschiedliche Konzentrationen des hK2-Proteins im Gewebe offenbar nicht die Begründung für die veränderten hK2-Konzentration im Serum sind. Die Regulierungsmechanismen vom Übergang der im Prostatagewebe nachgewiesenen Kallikreine in das Serum bleiben damit weiterhin unklar.

5. Neue Tumormarker des PCa im Gewebe

Im folgenden Kapitel wird die Evaluierung weiterer Kallikreine und eines weiteren Gens im PCa-Gewebe als erster Schritt in der Biomarker-Evaluierung für ein Organ beschrieben [134]. Dies beinhaltet die Beurteilung der Expressionshöhen zwischen Karzinom- und Nichtkarzinomgewebe.

5.1 Die Kallikreine 14 und 15

Wie im Abschnitt 3.3.2.2 bereits kurz erläutert, gehören der Gruppe der Kallikreine mittlerweile 15 Gene und die entsprechenden Proteine an. Im Jahr 2001 wurden zwei neue Mitglieder identifiziert und charakterisiert [135, 136]. Die beiden Kallikrein-Gene KLK14 und KLK15 stellen dabei in der parallelen Betrachtung eine vielversprechende Kombination bezüglich der Expression im Prostatagewebe dar [135, 136]. So zeigt das KLK14 in 80% der Fälle eine starke Herunterregulierung im malignen im Vergleich zum normalen Prostatagewebe des selben Patienten [135].

An insgesamt 58 Gewebeproben (jeweils gepaarte maligne und benigne Proben) von 29 PCa-Patienten wurde ein Vergleich der Gen-Expression des KLK15 vorgenommen [136]. Im Gegensatz zum PSA, KLK2 und zum KLK14 wurde keine generelle Expressionsminderung, sondern in fast der Hälfte der Patienten (13 von 29) eine höhere Expression im Tumorgewebe im Vergleich zum nicht malignen Gewebe nachgewiesen [136]. Diese ersten Untersuchungen zum KLK14 und KLK15 erfolgten nur qualitativ und nicht quantitativ.

Die Nutzung der quantitativen Echtzeit-Polymerase-Kettenreaktion (Real-time PCR) für Gen-Expressions-Studien war während meines Forschungsaufenthaltes im Labor von Prof. Diamandis (Toronto) ein Schwerpunkt meiner Tätigkeit. Mit der Einführung des LightCyclers im Jahr 1997 konnte erstmals die fluoreszenzgestützte photometrische Echtzeit-Messung von PCR-