Sonographie und farbkodierte Dopplersonographie zur Dignitätsbeurteilung von Ovarial- und Mammatumoren

Brigitta Reimann

2002

Frauenklinik und Poliklinik der Technischen Universität München,

Klinikum rechts der Isar (Direktorin: Univ.-Prof. Dr. M. B. Kiechle)

Sonographie und farbkodierte Dopplersonographie zur Dignitätsbeurteilung von Ovarial- und Mammatumoren

Brigitta Reglinde Reimann

Vollständiger Abdruck der von der Fakultät für Medizin der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Medizin

genehmigten Dissertation

Vorsitzender: Univ.-Prof. Dr. D. Neumeier Prüfer der Dissertation: 1. Univ.-Prof. Dr. K.-Th. M. Schneider 2. Priv.-Doz. Dr. C. Höβ

Die Dissertation wurde am 21.5.02 bei der Technischen Universität München eingereicht und durch die Fakultät für Medizin am 13.11.02 angenommen.

Reglinde

und

Klaus

in großer Liebe

INHALTSVERZEICHNIS

ABKÜRZUNGSVERZEICHNIS

1 EINLEITUNG		INLEITUNG	1
	1.1	Tumorassoziierte Neoangiogenese	1
	1.2	Diagnostische Ultraschallverfahren: physikalische und technische Grundlagen	2
	1.3	Einsatz von Sonographie und Dopplersonographie beim Ovarialkarzinom	4
	1.4	Einsatz von Sonographie und Dopplersonographie beim Mammakarzinom	6
	1.5	Einsatz der künstlichen Intelligenz in der Medizin: Neuronale Netze	7
2	F	RAGESTELLUNG	11
3 MATER		Iaterial und Methodik	13
	3.1	Qualitätskontrolle dopplersonographischer Blutflussmessungen	13
	3.	1.1 Entwicklung Eines In-Vitro-Strömungsmodells	13
	3.	1.2 Versuchsreihen zur Reproduzierbarkeit von Blutflussanalysen	15
	, -	3.1.2.1 Duplex-Dopplersonographie	15
		3.1.2.2 Farbkodierte Dopplersonographie	19
	3.2	Ovarialtumoren	22
	3.	2.1 Patientinnen	22
	3.	.2.2 Sonographische Untersuchung	22
	3.	.2.3 Dopplersonographische Untersuchung	23
	3.	2.4 Histopathologie	25
	3.	2.5 statistische Verfahren	27
	3.3	Mammatumoren	28
	3.	.3.1 Patientinnen und Untersuchungsdaten	29
	3.	3.2 Histopathologie	31
	3.	.3.3 Datenanalyse mit der multiplen logistischen Regression	33
	3.	.3.4 Auswertung mit Neuronalen Netzen	36

III

4	E	RGEBNISSE	39	
	4.1	Reproduzierbarkeit von dopplersonographischen Blutflussanalyse	en39	
	4.	1.1 Gepulste Dopplersonographie		
	4.	1.2 Farbkodierte Dopplersonographie	47	
	4.2	Ovarialtumoren		
	4.	2.1 Sonographische Dignitätsbeurteilung	55	
	4.	2.2 Dopplersonographische Dignitätsbeurteilung		
	4.	2.3 Kombinierte Dignitätsbeurteilung	69	
	4.3	Mammatumoren		
	4. D	3.1 Kombinierte sonographische und dopplersonographische gnitätsbeurteilung (multiple logistische Regression)	73	
	4. D	3.2 Kombinierte sonographische und dopplersonographische gnitätsbeurteilung (Neuronale Netze)	75	
	4.	3.3 Vergleich der neuronalen Netze mit statistischen Verfahren	78	
5	D	SKUSSION	81	
	5.1	Reproduzierbarkeit dopplersonographischer Untersuchungen	81	
	5.2 Bedeutung der Kombination von Sonographie und Dopplersonographie für die Diagnostik von Ovarialtumoren			
	5.3 Einsatz neuronaler Netze und der multiplen logistischen Regression zur Auswertung sonographischer und dopplersonographischer Untersuchungen99			
6	Z	JSAMMENFASSUNG	105	
7	Literatur Abbildungsverzeichnis Tabellenverzeichnis			
8				
9				
10	10 Anhang		125	

ABKÜRZUNGSVERZEICHNIS

μm	Mikrometer
A.	Arteria
CDE	color Doppler energy (energiekodierte Farbdopplersonographie)
CDI	color Doppler imaging (farbkodierte Dopplersonographie)
cm	Zentimeter
dB	Dezibel
dl	Deziliter
FIGO	Fédération Internationale de Gynécologie et d'Obstétrique
g	Gramm
h	Stunde
Hz	Hertz
LK	Lymphknoten
LMP	low malignant potential (Tumoren niedrig maligner Potenz)
m	Meter
Max.	Maximum
Med.	Median
MHz	Megahertz
Min.	Minimum
ml	Milliliter
mm	Millimeter
mW	Milliwatt
n	Anzahl
р	statistische Irrtumswahrscheinlichkeit
PI	Pulsatilitäts-Index
RI	Resistance-Index
ROC	receiver operating characteristics
S	Sekunde
TU	Tumor
V _{max}	systolische Maximalgeschwindigkeit
V _{mean}	über einen Herzzyklus gemittelte Maximalgeschwindigkeit
V_{min}	enddiastolische Maximalgeschwindigkeit

1 EINLEITUNG

Gynäkologische Tumoren sind die Ursache für einen erheblichen Teil der Morbidität und Mortalität der weiblichen Bevölkerung: Etwa jede zehnte Frau erkrankt im Laufe ihres Lebens an Brustkrebs. Obwohl das Ovarialkarzinom deutlich seltener auftritt, ist es dennoch bei Frauen die Krebs-Todesursache Nummer fünf nach den Karzinomen der Brust, des Dickdarms, der Lunge und der Bauchspeicheldrüse [9]. Trotz neuer Therapie-Strategien hat sich die Prognose des Ovarialkarzinoms in den letzten Jahrzehnten nicht entscheidend verbessert. Grund dafür ist vor allem, dass die Möglichkeiten der Frühdiagnostik noch immer begrenzt sind. Die Sonographie und die gepulste Farbdoppler-Sonographie sind nicht-invasive Verfahren, die sowohl bei Tumoren der Brust als auch der Eierstöcke einen wertvollen Beitrag zur Frühdiagnostik leisten können.

1.1 Tumorassoziierte Neoangiogenese

Eine wesentliche Voraussetzung für das Wachstum von Tumoren stellt die Neoangiogenese dar. Nicht nur maligne Tumoren, auch entzündliche Prozesse und benigne Proliferationen können zur Gefäßneubildung führen. Die in bösartigen Tumoren neu entstandenen Gefäße unterscheiden sich aber von normalen Blutgefäßen in vielen Punkten [33] (S. 281): Die Tunica Muskularis fehlt meist, so dass sich die Widerstände in karzinom-induzierten Gefäßen anders als in normalen Blutgefäßen verhalten. Außerdem sind unregelmäßige Verzweigungen, Kalibersprünge im Gefäßverlauf, Schlängelungen der Gefäße und vermehrt arteriovenöse Shunts typisch, die Veränderungen im Blutfluss nach sich ziehen.

Grund dieser extremen Gefäßinduktion sind von den metabolisch aktiven, sich schnell teilenden Karzinomzellen produzierte angiogenetische Faktoren wie etwa EGF (epidermal growth factor) und Angiogenin [32]. Es spielen jedoch auch durch die Gefäßendothelien produzierte Faktoren wie der VEGF (vascular endothelial growth factor), der bei lokaler Hypoxie vermehrt ausgeschüttet wird, eine Rolle [32]. Die Neovaskularisation ermöglicht dem Tumor das Größenwachstum und durch den Anschluss an das Gefäßsystem die Möglichkeit zur Metastasierung. Die erhöhte Vaskularisation und die Merkmale der neu eingesprossten, atypischen Blutgefäße mit dem veränderten Flussverhalten können mit der farbkodierten und mit der gepulsten Dopplersonographie dargestellt werden.

1.2 Diagnostische Ultraschallverfahren: physikalische und technische Grundlagen

Schall mit Frequenzen oberhalb der Hörgrenze, also oberhalb von 16 kHz, bezeichnet man als Ultraschall. In der medizinische Diagnostik wird im allgemeinen in einem Frequenzbereich zwischen 1 und 15 MHz gearbeitet, in der Gynäkologie setzt man bevorzugt Frequenzen von 3,5 bis 7,5 MHz ein [77].

Grundlage ist die von den Geschwistern Curie 1880 entdeckte Tatsache, dass die Deformierung von Kristallen Spannung erzeugen kann (piezoelektrischer Effekt). Umgekehrt kann das Anlegen einer elektrischen Ladung eine Deformierung des Kristalls bewirken (reziproker piezoelektrischer Effekt). Wird an Quarz- oder Bariumtitanat mit hoher Frequenz eine Wechselspannung gelegt, dehnt sich der Kristall aus und zieht sich wieder zusammen, was mechanische Schwingungen, also Schallwellen, erzeugt. Bei der Ultraschalltechnik werden sowohl der piezoelektrische als auch der reziproke piezoelektrische Effekt ausgenützt. Eine Piezokeramikscheibe mit Plattenelektroden dient so als Sender und als Empfänger [26] (S. 11-12).

Im einfachsten Ultraschallprinzip, nach der Amplitude A-Mode genannt, lenkt das von Grenzflächen reflektierte Echo eines Ultraschallstrahles einen Elektronenstrahl aus, der mit gleichmäßiger Geschwindigkeit über den Bildschirm einer Kathodenstrahlröhre geführt wird. Dieser Modus wird unter anderem bei der Echoenzephalographie und bei der Augenund Orbitadiagnostik eingesetzt [77] (S. 6).

Beim B-Mode (engl. brightness = Helligkeit) werden die Echos einer Grenzfläche nicht als Amplitude, sondern als Bildpunkte dargestellt. Die Helligkeit der erzeugten Bildpunkte ist von der reflektierten Echostärke abhängig. Das B-Bild (B-Scan) baut durch die räumliche Verschiebung einer B-Mode-Zeile in einer bestimmten Ebene ein zweidimensionales Bild auf [77] (S. 6-7). Es ist das Verfahren der ubiquitär eingesetzten zweidimensionalen, morphologischen Ultraschalldiagnostik, also z.B. der Abdominal- und Schilddrüsensonographie und der gynäkologischen Sonographie. Die Dopplersonographie basiert auf dem von Christian Johann Doppler 1842 erstmals beschriebenen Effekt, dass sich die Frequenz des reflektierten Lichtes sich bewegender Sterne in Bezug zu einem feststehenden Beobachter verändert. Analog verhalten sich auch die Schallfrequenzen bewegter Objekte. Die Geschwindigkeit von Schallquellen kann durch ein Quantifizieren der Frequenzänderungen bestimmt werden [57] (S. 5). Bei der Dopplersonographie wird der ausgesandte Ultraschall an den korpuskulären Blutbestandteilen reflektiert. Das Spektrum der Flussgeschwindigkeiten in einem Gefäß kann so akustisch abgeleitet werden.

Durch die Kodierung in verschiedenen Farben ist zusätzlich eine visuelle, qualitative Darstellung der verschiedenen Geschwindigkeiten und Flussrichtungen möglich. In diesem, farbkodierte Dopplersonographie (Farbdoppler, Color Doppler Imaging, CDI), genannten Verfahren wird die Blutströmung auf den Schallkopf zu üblicherweise rot und die sich vom Schallkopf entfernende Strömung blau kodiert, Turbulenzen werden gelb dargestellt [26] (S.39). In der energiekodierten Dopplersonographie (Color Doppler Energy, CDE) ist nicht nur die Blutflussgeschwindigkeit, sondern zusätzlich die vom Blutfluss erzeugte Energie für die Farbkodierung verantwortlich. Dadurch soll sie langsamere Blutflussgeschwindigkeiten sensitiver aufzeichnen und unabhängiger vom Winkel sein [4]. Allerdings können unterschiedliche Flussrichtungen nicht mit verschiedenen Farben dargestellt werden.

Eine quantitative Erfassung der Flussgeschwindigkeiten und der Widerstände ist mit der dopplersonographischen Ableitung von Blutströmungsprofilen möglich. Diese wird als Duplex (gepulster Doppler) bezeichnet und besteht aus einer Kombination von B-Bild und Doppler-Spektrum. So kann man das "sample volume" (der Bereich, aus dem die Signale abgeleitet werden) unter Sicht am gewünschten Gefäßabschnitt positionieren und danach bei "eingefrorenem" zweidimensionalem Bild Dopplersignale ableiten [57] (S. 10). Die Dopplersonographie ist die Grundlage für die Durchgängigkeitsmessung von Arterien, für die Widerstandsmessung arterieller Gefäße in der Tumordiagnostik und für die fetale und kardiale Gefäßdiagnostik. [77] (S.7).

Die Diagnostik gynäkologischer Tumoren beruht vor allem auf den drei Ultraschallverfahren B-Bild (B-Scan), Duplex-Mode und farbkodierte Dopplersonographie.

1.3 Einsatz von Sonographie und Dopplersonographie beim Ovarialkarzinom

Die Ultraschalluntersuchung des Ovars gehört zur Standarddiagnostik bei der Dignitätsbeurteilung unklarer ovarieller Prozesse. In den letzten Jahren erfuhr sie eine stetige Weiterentwicklung.

Die ersten Versuche, Ovarialtumoren sonographisch zu beurteilen, gehen auf die frühen 70er Jahre zurück [79] (S. 602). Die Bildqualität der Geräte war zu dieser Zeit beschränkt. Bei der transabdominellen Untersuchung erschwerten die weitere Entfernung des Schallkopfes von den Ovarien und Darmgasüberlagerungen die Beurteilbarkeit zusätzlich, so dass sich die Beschreibung vor allem auf drei Merkmale stützte: Solide, zystisch oder zystisch-solide.

Mit der Weiterentwicklung der Technik konnten neue morphologischen Aspekte hinzugezogen und Kriterienkataloge für die Malignität von Tumoren erarbeitet werden [76], was eine differenziertere Diagnostik ermöglichte. Neben Größe, Lokalisation und Konsistenz des Tumors wurde nun auch die Regelmäßigkeit des Randes mit bewertet [28]. Anfang der 80er-Jahre begannen auch erste Versuche, den transabdominellen Ultraschall als Screening-Methode einzusetzen [38]. Die Ergebnisse groß angelegter Screening-Studien fielen aber nicht besonders ermutigend aus, die Operationsraten waren für die wenigen gefundenen Karzinome viel zu hoch [14], [16] (S. 309), [20], [38].

Einige Vorteile und Verbesserungen auch für das Screening brachte die Umstellung auf transvaginalen Ultraschall mit sich, die in den späten 80er-Jahren erfolgte: Der Schallkopf kommt näher an den Beckenorganen zu liegen, so dass mit höherer Ultraschallfrequenz von 5 bis 7 MHz und damit besserer Bildauflösung gearbeitet werden kann [7] (S. 168). Besonders bei adipösen Patientinnen ist durch den transvaginalen Schall eine erheblich bessere Darstellung der Ovarien möglich. Zudem ist die Schallqualität nicht mehr von einer gefüllten Blase abhängig [104]. Durch die bessere Bildqualität wurde eine noch genauere morphologische Klassifikation möglich. Verschiedene Studien deuteten auf eine Verbesserung beim Screening asymptomatischer Frauen, vor allem in der Postmenopause, hin [11], [45], [81]. Durch morphologische Score-Systeme schien die Spezifität des Screenings noch verbesserbar [23], [43], [90]. Sassone et al. etwa kombinierten die

sonomorphologischen Kriterien Wandstruktur, Wanddicke, Septierung und Echogenität und entwickelten damit ein Punktesystem mit Werten von 4 bis 15, die Grenze zur Malignität wurde ab 9 Punkten gesetzt [90]. Damit konnten immerhin eine Sensitivität von 100 % und eine Spezifität von 83 % erreicht werden. Die Schwierigkeit, maligne von benignen Prozessen zu unterscheiden und damit den Frauen mit gutartigen Tumoren die Operation zu ersparen, ohne bösartige übersehen zu haben, wird damit zwar etwas verringert, eine diagnostische Unsicherheit bleibt jedoch weiterhin bestehen.

Ende der 80er Jahre wurden Vaginalsonden mit integriertem gepulstem und farbkodiertem Doppler entwickelt. Die ersten klinischen Erfahrungen gaben Anlass zu enthusiastischen Prognosen [10]. Frühe Studien an kleinen Patientengruppen nannten Sensitivitäten von 100 % und Spezifitäten von 97 % [63]. In späteren Studien konnten diese Werte nicht reproduziert werden, zum Teil wohl aufgrund verschiedener Geräte und uneinheitlicher Untersuchungsbedingungen [52]. Viele Untersucher verwendeten den niedrigsten aufgezeichneten Doppler-Wert für die statistische Auswertung [13], [18], [41], [105], andere gebrauchten den Mittelwert[17], [27]; weniger gebräuchlich war der höchste Wert [88] oder der Median. Zudem wurden unterschiedliche Doppler-Parameter benutzt, vor allem der Pulsatilitäts-Index und der Resistance-Index, gelegentlich auch die systolische Maximalgeschwindigkeit. Die verschiedenen Grenzwerte besonders beim PI und beim RI variierten von Studie zu Studie, die Angaben schwankten zwischen 0,4 und 0,72 für den RI [17], [43], [62] und zwischen 0,62 und 1,25 für den PI [56], [105]. In den meisten Studien wurde nur einer der beiden Parameter beurteilt. Die Gefäßlokalisation wurde oft nicht genau definiert, eine strikte Einteilung in zentral (im Tumor), am Rand des Tumors oder außerhalb des Tumors mit Beziehung zu ihm gelegen, fand sich selten [61]. Die Vielfalt der Parameter zeugt von einer intensiven Suche, erschwert aber vergleichende Analysen und somit eine einheitliche Ergebnisfindung.

Nach wenigen Jahren der Euphorie tauchten bald kritische Stimmen unter den Publikationen auf [25], insbesondere das Screening betreffend. In nur in wenigen Veröffentlichungen wurde versucht, einen aus Sonomorphologie- und Dopplersonographie-Parametern kombinierten Screening-Score zu entwickeln [62] oder mit statistischer Analyse wie der multiplen logistischen Regression zu arbeiten [99]. Versuche, die Daten mit Methoden der künstlichen Intelligenz wie den neuronalen Netzen auszuwerten, gibt es in der Gynäkologie erst seit kürzerer Zeit [98] (S. 21). Gerade bei der Entdeckung von Ovarialkarzinomen in einem potentiell noch heilbaren Stadium (LMP, FIGO I) besteht auch heute bei der Dopplersonographie noch die größte Schwierigkeit [53] (S. 500).

1.4 Einsatz von Sonographie und Dopplersonographie beim Mammakarzinom

Ähnlich wie bei den Ovarialtumoren geht die Sonographie der Mamma auf die frühen 70er Jahre zurück. Seit Mitte der 70-er werden intensive Untersuchungen zur Identifizierung maligner Tumoren und zur Abgrenzung dieser von benignen Tumoren unternommen. Trotz der Überlegenheit der Mammographie konnte sich die Sonographie, im Gegensatz zu Untersuchungsmethoden wie anderen der Thermographie, als wichtiges Diagnoseinstrument behaupten. Neben ihrem Einsatz zur Differenzierung und Lokalisierung gewann sie in den letzten Jahren vor allem Bedeutung bei der Feinnadelbiopsie. ultraschallgesteuerten Auch als wertvolle Ergänzung zur Mammographie bei radiologisch dichteren Brüsten von Frauen in der Prämenopause ist sie nicht mehr wegzudenken [51] (S. 5), [54] (S. 4).

Einen deutlichen Zugewinn an diagnostischer Treffsicherheit brachte auch hier die Einführung der Dopplersonographie, die 1977 erstmals beschrieben wurde [108]. Vor allem der Einsatz der gepulsten, farbkodierten Dopplersonographie der Mamma gegen Ende der 80er Jahre gab den Forschungen zur Dignitätsdiagnostik einen erneuten Schub [19], [47]. Die erreichten Sensitivitäten und Spezifitäten blieben jedoch hinter den Erwartungen zurück. Dies mag, wie bei der Sonographie des Ovars, unter anderem an je nach Studiengruppen differierenden Kriterienkatalogen und an Uneinheitlichkeiten zum Beispiel bei der Wahl der Grenzwerte für die Flussparameter liegen. Zunächst waren mit dem Gedanken, dass maligne Gefäße durch die fehlende Muskularis einen geringeren Flusswiderstand aufweisen, niedrige Werte für die Widerstandsindizes propagiert worden [1], [96]. Bei weiterführenden Untersuchungen stellte sich jedoch heraus, dass bei der Mamma, ganz im Gegensatz zum Ovar, hohe Indizes für Malignität sprechen [8], [72], [89]. Ein Einsatz zum routinemäßigen Screening als Alternative zur Mammographie, der viele Frauen wegen der Strahlenbelastung ablehnend gegenüberstehen, ist aus diesen Gründen, insbesondere aber wegen der noch immer unterlegenen Treffsicherheit gerade bei nicht palpablen Tumoren [54], noch nicht absehbar.

1.5 Einsatz der künstlichen Intelligenz in der Medizin: Neuronale Netze

Künstliche neuronale Netze ("artificial neural networks", ANN) sind eine Methode der Informationsverarbeitung, die von Informatikern ursprünglich aus der Medizin bzw. der Biologie entliehen wurde. Ihr liegt die Idee zugrunde, das "Urteilsvermögen" eines elektronischen Rechners zu verbessern, indem der Computer die Funktionsweise des menschlichen Gehirns, also das Zusammenspielen von Neuronenverbänden, nachahmt. Dies soll eine Hilfe bei der Aufgabe bieten, Daten zu klassifizieren und auszuwerten; zum Beispiel in einer medizinischen Anwendung anhand von verschiedenartigen Untersuchungsergebnissen die richtige Diagnose zu stellen.

Die neuronalen Netze finden jedoch nur langsam ihren Weg zurück in die Medizin, ein Grund dafür ist unter anderem der "Black Box"-Charakter des Netzesystems [6]: Analog zum biologischen Vorbild wird die Funktionsweise eines einzelnen Neurons weitgehend verstanden, die Mechanismen in größeren Netzwerken dagegen sind nicht ohne weiteres analytisch nachvollziehbar [111] (S. 27).

Ein wesentliches Element der künstlichen neuronalen Netze ist ihre Lernfähigkeit, also das Vermögen, ohne explizierte Programmierung eine Aufgabe selbständig aus Trainingsbeispielen zu lernen [111] (S. 26). Die Zellen eines künstlichen Netzwerkes sind stark idealisierte Neuronen. In Anlehnung an das biologische Vorbild bestehen sie aus drei Komponenten: Einem Zellkörper, den Dendriten und einem Axon. Die Dendriten führen die Eingaben des Netzes in die Zelle, das Axon leitet die Ausgabe der Zelle weiter, verzweigt sich und tritt mit den Dendriten nachfolgender Neuronen über Synapsen in Kontakt. Die Stärke der Synapsen wird in der künstlichen Intelligenz durch einen numerischen Wert, das Verbindungsgewicht, dargestellt [111] (S. 72). Neuronale Netze lernen, indem sie z. B. ihre Verbindungsgewichte modifizieren, wenn ihnen wiederholt Datensätze zum Training eingegeben werden [111] (S. 73). Ein Maß für die Veränderung der Gewichte ist bei häufig verwendeten Lernverfahren wie der so genannten Backpropagation die Differenz zwischen gewünschtem und erhaltenem Ausgabewert des Netzes. Die Aufgabe besteht darin, die Gewichte so zu optimieren, dass für alle Eingaben mit möglichst hoher Genauigkeit der richtige Ausgabewert gefunden wird und damit der Ausgabefehler möglichst klein ist.

Backpropagation ist ein Gradientenabstiegsverfahren. Trägt man den Fehler eines neuronalen Netzes mit zwei Eingängen als Funktion der Gewichte des Netzwerkes graphisch auf, erhält man eine Fehlerfläche, wie z.B. in **Abb. 1** dargestellt [111] (S. 105). Ausgehend von einem vorgegebenen oder zufällig gewählten Startpunkt sucht sich das Netz in fortlaufenden Schritten immer wieder den steilsten Abstieg, bis es in der Umgebung ein Fehlerminimum erreicht hat. Dieser Punkt stellt allerdings nur ein lokales Minimum dar, das nicht zwingend identisch sein muss mit dem absoluten Fehlerminimum [111] (S. 112).



Abb. 1: Fehlerfläche (Error) als Funktion der Gewichte w1 und w2 bei Backpropagation (nach [111], S. 105).

Neuronale Netze können verschiedenartig aufgebaut sein. In der Regel bestehen sie aus einer Schicht von Eingabeneuronen und einer Schicht von Ausgabeneuronen; dazwischen können eine oder mehrere so genannte verdeckte Schichten (hidden layers) liegen (Abb. 2). Die Rechenergebnisse der verdeckten, mittleren Schichten sind für den Anwender normalerweise nicht einsehbar.

Die Informationen werden in die Eingabeneuronen eingespeist, es können nicht nur quantitative Kriterien (z.B. Tumorgröße) verwendet werden, sondern auch qualitative (z.B. Vaskularisation vorhanden oder nicht vorhanden). Letztere müssen aber, um für das Netz auswertbar zu sein, in numerische Werte übertragen werden. Meist werden dafür Binärzahlen verwendet, also z.B. keine Vaskularisation vorhanden = 0, Vaskularisation vorhanden = 1. Die Ausgabe der vom Netz ermittelten Ergebnisse erfolgt ebenfalls in Zahlen, die eine Art Wahrscheinlichkeit des angenommenen Ereignisses darstellen. Durch Rücktransformation der Ergebnisse kann wieder eine qualitative Aussage erzielt werden, z.B. "das Ereignis trifft zu, wenn der Ausgabewert größer als der Schwellenwert 0,5 ist".



Abb. 2: Beispiel eines neuronalen Netzes mit zwei verdeckten Schichten (nach [111], Seite 73).

Eines der einfachsten Netze besteht aus verschiedenen Eingängen e_1 bis e_n , einem speziellen Eingang e_0 und einem Ausgang a (Abb. 3). Die eingegangenen Informationen werden durch die Zahlen w gewichtet. Aus der Summe der gewichteten Eingangswerte berechnet sich der Ausgabewert a (Formel 1). Der Eingang $e_0=1$ wird benötigt für den Fall, dass alle Eingänge null betragen, der Ausgang aber einen anderen Wert haben soll. Idealerweise sollten die Eingänge alle positive Werte haben oder Null betragen.

$$a = w_0 \cdot 1 + w_1 \cdot e_1 + w_2 \cdot e_2 + \dots + w_n \cdot e_n$$

= $w_0 + \sum_{i=1}^n w_i \cdot e_i$ (1)

Die Gewichte w_1 bis w_n müssen nun so bestimmt werden, dass für alle Eingaben mit hoher Wahrscheinlichkeit der richtige Ausgabewert gefunden wird. Zunächst wird die Lernphase durchgeführt, in der die Ausgabewerte vorgegeben sind, dabei wird für "wahr" bzw. "ja" eine 1 gesetzt, für "falsch" bzw. "nein" eine 0. Man ändert die Gewichte in jedem neuen Lerndurchgang stets nur um kleine Beträge, damit sich das System auf einen stationären Zustand, d. h. auf ein lokales Minimum nach dem Lernsystem der Backpropagation, zubewegen kann [111] (S. 111).



Abb. 3: Schematisiertes Beispiel eines einfachen neuronalen Netzes

2 FRAGESTELLUNG

Die Arbeit besteht aus drei eng miteinander verbundenen Teilen, alle basierend auf der Sonographie und der Dopplersonographie: Einem experimentellen Strömungsmodell, der Dignitätsdiagnostik am Ovar und der Dignitätsdiagnostik an der Mamma. Zur Auswertung wurde die multiple logistische Regression mit den neuronalen Netzen, einer Methode der künstlichen Intelligenz, verglichen.

Am experimentellen Strömungsmodell sollten Fehlermöglichkeiten bei der dopplersonographischen Tumordiagnostik aufgezeigt werden, die vom Untersucher für die Erzielung möglichst korrekter Ergebnisse unbedingt beachtet werden müssen. Nicht nur die Art des Gewebes und die Erfahrung des Untersuchenden, sondern auch die Qualität und die Einstellung des Gerätes beeinflussen das Untersuchungsergebnis nachhaltig. Ein falsch positives oder negatives Resultat etwa kann allein aus einem schlecht getroffenen Winkel resultieren. Der Einfluss der Geräteeinstellung auf die farbkodierte und gepulste dopplersonographische Untersuchung sollte durch Variieren von Geschwindigkeitsskala, Farbintensität, Farbdopplerverstärkung, Winkelkorrektur und Dopplerfenster gezeigt werden. Außerdem wurden physikalische Faktoren wie die Messgenauigkeit in unterschiedlicher Tiefe und die Doppler- und Farbduplexableitung bei Simulation untereinander liegender Gefäße sowie biologische Faktoren wie die Hämatokritabhängigkeit des Dopplersignals analysiert.

Für die **Dignitätsdiagnostik bei Ovarial- und Mammatumoren** sollten die wichtigsten in der Literatur beschriebenen Kriterien der sonographischen und dopplersonographischen Untersuchung aufgearbeitet und hinsichtlich ihrer Wertigkeit überprüft werden (Ovar: [12], [17], [23], [28], [49], [53] (S.497), [62], [81], [86] (S. 347), [90], Mamma: [5], [24], [40], [58] (S. 383), [60] (S. 89), [70] (S. 20), [73], [74], [83], [96], [97]). Dadurch sollten die für die Dignitätsdiagnostik von Ovarial- und Mammatumoren relevanten sonographischen und dopplersonographischen Parameter ermittelt werden.

Die Auswertung mit Methoden der Statistik und der künstlichen Intelligenz sollte ermitteln, ob durch den additiven Einsatz der Dopplersonographie die sonographische Vorhersagegenauigkeit bei Ovarial- und Mammatumoren verbessert werden kann. Außerdem beinhaltete sie die Suche nach neuen Wegen, um die Parameter sinnvoll zu kombinieren und die diagnostische Treffsicherheit zu erhöhen, ohne auf ein starres Punkte-Score-System angewiesen zu sein. Die Ergebnisse der statistischen Methode der multiplen logistischen Regression sollte mit der künstlichen Intelligenz, den neuronalen Netzen, hinsichtlich ihrer Validität verglichen werden.

3 MATERIAL UND METHODIK

Für alle Untersuchungen wurde das hoch auflösende Realtime-Gerät Acuson 128 XP (Firma Acuson, Mountain View, CA, USA) verwendet. Die Untersuchungen des Ovars wurden mit einem 5,0 MHz transvaginalen Schallkopf mit integriertem 5 MHz Pulse Wave Doppler Transducer und einem Sektorwinkel von 140 Grad (Acuson EV 5) durchgeführt. Für die Sonographie und Dopplersonographie der Mamma sowie die In-Vitro-Versuche wurde ein 7,0 MHz Linearschallkopf mit einem 5 MHz Pulse Wave Doppler Transducer benutzt.

3.1 Qualitätskontrolle dopplersonographischer Blutflussmessungen

An einem In-Vitro-Strömungsmodell sollten durch das Variieren einzelner Parameter bei sonst konstanter Einstellung des Ultraschallgerätes Einflussgrößen und Fehlermöglichkeiten bei der dopplersonographischen Messung untersucht werden.

3.1.1 ENTWICKLUNG EINES IN-VITRO-STRÖMUNGSMODELLS

Um einen intravasalen Blutfluss zu simulieren, wurde ein Perfusor (Injectomat CP-IS, Firma Fresenius) mit einem Pumpvolumen von 0,5 bis 199,5 ml/h und konstantem Fluss verwendet. Als "Blutgefäß" diente ein Silikonschlauch (Infusiomatenschlauch) von 1 mm Innendurchmesser und 2 mm Außendurchmesser, durch den mit isotoner Kochsalzlösung verdünnte Erythrozytenkonzentrate gepumpt wurden. Der Hämatokrit betrug, mit Ausnahme eines Versuches zur Überprüfung der Hämatokritabhängigkeit, 40 % bis 50 %. Die Konzentrate stammten aus dem hämatologischen Labor des Klinikums rechts der Isar; sie hatten das Verfallsdatum knapp überschritten oder waren nach dem Auftauen zur Verwendung für Patienten nicht gebraucht worden. Die verdünnten Erythrozytenkonzentrate wurden an einem Infusionsständer befestigt, der Silikonschlauch wurde mit einem Infusionsbesteck daran angeschlossen. Nach dem Einspannen in die Pumpe wurde der Schlauch an den beiden Längsseiten einer Plastikwanne so befestigt, dass er in der mit Ultraschallgel gefüllten Wanne einen schrägen Verlauf nahm. Er endete in einem Messzylinder, in dem das Erythrozytenkonzentrat aufgefangen wurde.

Der Ultraschallkopf wurde mit einer Halterung oberhalb der Plastikwanne so befestigt, dass die Kontaktfläche leicht in das Gel eintauchte und der Schlauch vom Ultraschallbild in seinem gesamten Verlauf erfasst wurde. Bei einigen Versuchen wurde er zum Simulieren von In-Vivo-Bedingungen von Hand gehalten.

Zu jedem zu untersuchenden Punkt wurden vier bis zwölf Messreihen an verschiedenen Tagen durchgeführt. Die Stromzeitvolumina variierten zwischen 5 und 199,5 ml/h. Der gemessene Schlauchabschnitt lag in zwei bis vier Zentimetern Tiefe. Vor und nach jedem Versuch wurde der Pumpfehler durch einen probeweisen Durchlauf bei 100 ml/h ermittelt. Das im Messzylinder nach genau 15 Minuten aufgefangene Volumen wurde abgelesen. Die durchschnittliche prozentuale Abweichung der tatsächlichen Strömungsgeschwindigkeit von ihrem Soll-Wert wurde berechnet.

Zunächst wurde das Ultraschallbild eingestellt, danach der Cursor über den gesamten Schlauchdurchmesser gelegt und die Dopplerflusskurve abgeleitet. Die Pulsrepetitionsfrequenz wurde dabei auf 400-800 Hz gestellt, der Wandbewegungsfilter mit 50-100 Hz möglichst niedrig gehalten. Die Sendeleistung wurde kleiner 800 mW/cm² gehalten.

In einigen Versuchen wurde die durch das konstante Pumpvolumen - wie bei venösen Gefäßen - nicht-pulsatile Flusskurve aufgezeichnet und abgemessen. Nach dem Gesetz von Hagen-Poiseuille (Formel 2) wurde aus Pumpvolumen \dot{V} und Schlauchradius *R* die Flussgeschwindigkeit errechnet (Formel 3) [37] (S. 103).

$$\dot{V} = \frac{\pi (p_1 - p_2)}{8 \cdot \eta \cdot l} \cdot R^4 \tag{2}$$

In das Gesetz von Hagen-Poiseuille zur Berechnung laminarer Strömungen gehen neben dem Schlauchradius R und der Länge l die Eigenschaften der Flüssigkeit mit der Viskosität η in die Rechnung mit ein. p_1 stellt dabei den höheren Druck an der Anfangsstelle des angenommenen zylindrischen Rohres dar, p_2 den niedrigeren an der Endstelle. Über das ganze Rohr gemittelt beträgt die Strömungsgeschwindigkeit \bar{v} [37] (S. 103):

$$\overline{v} = \frac{V}{\pi \cdot R^2} = \frac{1}{2} \cdot v_0 \tag{3}$$

Dabei entspricht die Geschwindigkeit v_0 dem Maximum in der Flussmitte eines Rohres.

Die verwendeten Pumpvolumina von 5 bis 199,5 ml/h entsprechen damit rechnerisch durchschnittlichen Strömungsgeschwindigkeiten im Rohr von 0,0018 bis 0,07 m/s bzw. maximalen Strömungen in der Flussmitte von 0,0035 bis 0,14 m/s.

In einigen Experimenten wurde der Blutfluss rein qualitativ mit der farbkodierten Dopplersonographie aufgezeichnet.

3.1.2 VERSUCHSREIHEN ZUR REPRODUZIERBARKEIT VON BLUTFLUSSANALYSEN

In verschiedenen Versuchsreihen wurde gezeigt, wie reproduzierbar Blutflussanalysen im Modell sind. Dabei wurden sowohl Parameter der duplex- als auch der farbkodierten dopplersonographischen Messung auf mögliche Einflüsse und Fehlerquellen untersucht.

3.1.2.1 Duplex-Dopplersonographie

Die im Doppler-Mode gemessenen Blutflusskurven können durch die Geräteeinstellung oder durch biologische Faktoren verändert werden. Dies kann die Ergebnisse der Geschwindigkeits- und Widerstandsmessung verfälschen.

Geschwindigkeitsmessung bei verschiedenen Pumpvolumina

Das Pumpvolumen wurde sukzessive variiert, um die Genauigkeit der dopplersonographischen Messung zu überprüfen. Ausgehend von 199,5 ml/h wurde das Pumpvolumen kontinuierlich in Schritten von zunächst 25 ml/h, unterhalb von 50 ml/h in Schritten von 5 ml/h bis zu einem minimalen Pumpvolumen von 5 ml/h verringert. Zu jedem Pumpvolumen wurde in acht Versuchsreihen dopplersonographisch die bei konstantem Fluss nicht-pulsatile Flusskurve abgeleitet. Der Winkel des in der Halterung fixierten Ultraschallkopfes zum Schlauch betrug jeweils 40 bis 60 Grad.

Die dopplersonographisch ermittelten Maximalgeschwindigkeiten wurden zur errechneten, realen Geschwindigkeit korreliert. Wie bereits oben erläutert, beträgt die dabei gemessene Maximalgeschwindigkeit des Dopplerspektrums (Flussgeschwindigkeit in der Rohrmitte) das Doppelte der Gesamtgeschwindigkeit im Rohr (Formel 3). Zu jeder dopplersonographisch errechneten Geschwindigkeit wurde der Mittelwert aus den acht Versuchen errechnet. Er wurde gemeinsam mit den Einzelwerten und dem 10 %-Konfidenzintervall in einem Diagramm in Bezug zum realen Pumpvolumen gesetzt.

Winkelabhängigkeit der Geschwindigkeitsmessung

Der Winkel θ zwischen Doppler-Ultraschallstrahl und Blutgefäß stellt eine nicht zu unterschätzende Fehlerquelle bei dopplersonographischen Untersuchungen dar. Die exakte Einstellung des Winkels θ ist deshalb so wichtig, weil der Winkel als Cosinus-Funktion in den Zähler der Doppler-Gleichung (Formel 4) eingeht, mit der sich die Strömungsgeschwindigkeit v des Blutes errechnen lässt (Formel 5) [94] (S. 5):

$$\Delta F = F_r - F_0 = \frac{2 \cdot F_0 \cdot v \cdot \cos\theta}{c} \tag{4}$$

$$v = \frac{\Delta F \cdot c}{2 \cdot F_0 \cdot \cos\theta} \tag{5}$$

Die Dopplerfrequenz ΔF entspricht dabei der von den korpuskulären Blutbestandteilen reflektierten Frequenz F_r abzüglich der vom Gerät ausgestrahlten Sendefrequenz F_0 . Die Schallausbreitungsgeschwindigkeit c wird im menschlichen Weichteilgewebe mit ca. 1540 m/s konstant angenommen [94] (S. 6).

Bei Winkelabweichungen wächst mit zunehmendem Winkel der Fehler. Bei 90 Grad ist keine Geschwindigkeitsmessung möglich ($\cos\theta = 0$). Bei einem Winkel von 60 Grad führt eine Fehlmessung von 5 Grad zu einer Unterschätzung von 13 % bzw. einer Überschätzung von 18 % [94] (S. 23). Winkel oberhalb von 60 Grad sollen aufgrund der zu hohen Fehlschätzung bei der quantitativen Flussmessung verworfen werden. In vivo ist das genaue Einstellen des Winkels jedoch, bedingt durch atypische Gefäßverläufe, Gefäßkrümmungen oder Gefäßaufzweigungen, häufig sehr schwer.

In zehn Versuchsreihen sollte diese Winkelabhängigkeit der Dopplersonographie aufgezeigt werden. In den Versuchen sollte insbesondere ermittelt werden, ab welchen Winkeln gültige Messungen nicht mehr anzunehmen sind. Dazu wurde in mehreren Schritten die Steilheit des Silikonschlauches in der Plastikwanne und somit die Winkeleinstellung zum Schallkopf verändert. Der Schallkopf wurde dabei von Hand gehalten, um die Bedingungen in vivo mit leichten Fehleinschätzungen des Winkels zu simulieren. Das Pumpvolumen betrug immer 100 ml/h, was einer maximalen Fließgeschwindigkeit von 0,07 m/s entspricht. Die Flussgeschwindigkeit wurde nach Einstellung des richtigen Winkels mit dem Doppler abgeleitet. Die so ermittelten Fließgeschwindigkeiten wurden in einem Diagramm in Relation gesetzt zur tatsächlichen Geschwindigkeit von 0,07 m/s.

Außerdem wurde zum Darstellen der im Gerät integrierten Winkelfunktion ein "Leerversuch" durchgeführt. Dazu wurde ohne Ableiten der Dopplerfunktion am Gerät, die gemäß der Doppler-Gleichung (Formel 4) eine Cosinus-Funktion ist, bei einem Winkel von 50 Grad eine Geschwindigkeit von 0,07 m/s eingestellt. Anschließend wurde der Winkel sukzessive variiert und die dabei vom Gerät angegebene Geschwindigkeit aufgezeichnet. Zur Auswertung wurde sie in dem Diagramm mit den gemessenen Geschwindigkeiten graphisch dargestellt.

Variieren der Empfangsverstärkung

Eine zu hoch oder zu niedrig gewählte Doppler-Empfangsverstärkung kann möglicherweise die Geschwindigkeit zu hoch oder zu niedrig erscheinen lassen. Um das Ausmaß dieses Fehlers zu ermitteln, wurden acht Versuchsreihen durchgeführt. Der Winkel des in der Halterung fixierten Schallkopfes zum Schlauch betrug konstant 55 Grad. Bei einem Pumpvolumen von 199,5 ml/h (maximale Geschwindigkeit in Flussmitte 0,14 m/s) wurden Dopplerflusskurven unter systematischer Veränderung der Doppler-Empfangsverstärkung ("Doppler gain") abgeleitet. Die gemessenen Geschwindigkeiten wurden mit der Sollgeschwindigkeit verglichen.

Dopplerfenster

Das Strömungsprofil einer laminaren Strömung in peripheren Arterien verläuft infolge von Trägheitskräften am Gefäßrand parabelförmig: In der Gefäßmitte sind die Geschwindigkeiten höher, am Gefäßrand langsamer (**Abb. 4**).



Abb. 4: Geschwindigkeitsprofil einer laminaren Strömung. Die Länge der Pfeile entspricht der jeweiligen Geschwindigkeit (modifiziert nach [94] (S. 5) und [37] (S. 102).

Wird das Dopplerfenster, also die an den Geräten einstellbare Länge der Schallkeule, in der das Dopplersignal abgeleitet wird, nur in der Gefäßmitte platziert, resultiert eine fälschlicherweise zu hohe Geschwindigkeit. Erfasst es das Gefäß nur am Rand, ist die Geschwindigkeit zu niedrig. Ein Dopplerfenster ("cursor", "sample volume"), das nicht das gesamte Gefäß umfasst, verfälscht auf diese Weise die Dopplerflusskurve [94] (S. 26-29).

Um diesen Effekt zu überprüfen, wurden acht Versuchsreihen an verschiedenen Tagen durchgeführt. Der Cursor wurde beim Ableiten der Flusskurve zunächst nur in die Flussmitte gelegt, danach umfasste er das ganze Gefäß. Das Pumpvolumen betrug 199,5 ml/h (maximale Geschwindigkeit 0,14 m/s). Der Winkel zwischen dem fixierten Ultraschallkopf und dem Schlauch betrug 50 bis 60 Grad, die Empfangsverstärkung lag bei 50 %. Die dopplersonographisch gemessenen Geschwindigkeiten wurden mit der Sollgeschwindigkeit (berechnet nach den Formeln 4 und 5) verglichen.

Genauigkeit der Geschwindigkeitsmessung in unterschiedlicher Tiefe

Um den Einfluss unterschiedlicher Messtiefen auf die dopplersonographische Geschwindigkeitsmessung zu überprüfen, wurde in acht Versuchsreihen an drei verschiedenen Tagen die Dopplerflusskurve zunächst in einem Schlauchabschnitt nahe der Oberfläche (1 cm Entfernung vom Schallkopf) und anschließend in einem tiefer gelegenen Schlauchabschnitt (4 cm Entfernung vom Schallkopf) abgeleitet. Die gemessene maximale Geschwindigkeit bei einem konstanten Pumpvolumen von 199,5 ml/h (maximale Flussgeschwindigkeit 0,14 m/s) wurde dokumentiert. Der Winkel des fixierten Schallkopfes zum Schlauch betrug konstant 50 Grad, die Doppler-Empfangsverstärkung war auf 50 % eingestellt. Die gemessenen Werte wurden mit dem Soll-Wert verglichen.

Simulation untereinander liegender Gefäße bei der dopplersonographischen Geschwindigkeitsmessung

In sechs Versuchsreihen sollten die Interferenzen bei der Analyse eines Gefäßes durch ein zweites Gefäß im Schallbereich erfasst werden. Dazu wurden zwei Silikonschläuche gleichen Durchmessers (1 mm Innendurchmesser, 2 mm Außendurchmesser) in einem Abstand von 1 cm in gleichem Winkel (60 Grad zum Ultraschallkopf) so untereinander angebracht, dass beide von dem fixierten Schallkopf in 1 bzw. 2 cm Tiefe optimal erfasst werden konnten. Beide wurden von getrennten Infusiomaten mit Erythrozytenkonzentrat gleichen Hämatokrits (40-45 %) durchspült. Das Pumpvolumen beider Infusiomaten betrug 199,5 ml/h (maximale Fließgeschwindigkeit 0,14 m/s). Die Empfangsverstärkung des Gerätes wurde auf 50 % gestellt. Anschließend wurde die Dopplerkurve erst des oberen, dann des unteren Gefäßes abgeleitet. Alle Messwerte wurden mit dem tatsächlichen Gefäßdurchmesser verglichen.

Einfluss des Hämatokritwertes

Der Hämatokrit kann bei Patientinnen mit malignen Neoplasien deutlich erniedrigt sein. Durch Versuche sollte festgestellt werden, ob diese Tumoranämie einen Einfluss auf die duplexsonographischen Messungen hat, und unterhalb welches Hämatokrits kein Farbduplex- oder kein Dopplersignal mehr ableitbar sind. Der Ultraschallkopf war dabei mit einem Winkel von 48 Grad zum Schlauch in der Halterung fixiert Das Pumpvolumen betrug konstant 100 ml/h (maximale Fließgeschwindigkeit in der Flussmitte 0,07 m/s). Das Erythrozytenkonzentrat wurde erst unverdünnt (Hämatokrit von 82,7 % bzw. 97,6 %) durch den Silikonschlauch gepumpt und danach progredient mit isotoner Kochsalzlösung bis hinab zu einem nicht mehr messbaren Hämatokrit verdünnt. Mit der Duplex-Dopplersonographie wurde die maximale Geschwindigkeit abgeleitet. Bei den stärkeren Verdünnungsstufen wurde das Gefäß zudem in der farbkodierten Dopplersonographie dargestellt. Der Hämatokrit jeder Verdünnungsstufe wurde vom hämatologischen Labor des Klinikums rechts der Isar bestimmt. Aufgrund des eindeutigen Ergebnisses bei zwei Messreihen wurde auf eine statistische Auswertung und die Durchführung weiterer Versuchsreihen verzichtet.

3.1.2.2 Farbkodierte Dopplersonographie

Das Auffinden von Gefäßen im Farbduplex-Mode ist eng an die Wahl einer sensiblen Geräteeinstellung gebunden. Der Untersucher hat somit über die Geräteeinstellung möglicherweise einen maßgeblichen Einfluss auf die Aussage "Vaskularisation vorhanden / nicht vorhanden".

Simulation untereinander liegender Gefäße bei der farbkodierten Dopplersonographie

In sechs Versuchsreihen sollte untersucht werden, ob die Darstellung eines Gefäßes durch ein zweites Gefäß im Schallbereich möglicherweise verschlechtert werden kann. Der Versuch wurde aufgebaut wie bei der Simulation untereinander liegender Gefäße bei der dopplersonographischen Geschwindigkeitsmessung: Zwei Silikonschläuche gleichen Durchmessers (1 mm Innendurchmesser, 2 mm Außendurchmesser) wurden in einem Abstand von 1 cm in gleichem Winkel (60 Grad zum Ultraschallkopf) so untereinander angebracht, dass beide von dem fixierten Schallkopf in 1 bzw. 2 cm Tiefe optimal erfasst werden konnten. Beide wurden von getrennten Infusiomaten mit Erythrozytenkonzentrat gleichen Hämatokrits (40-45 %) durchspült. Das Pumpvolumen beider Infusiomaten betrug 199,5 ml/h (maximale Fließgeschwindigkeit 0,14 m/s). Die maximale Geschwindigkeitsskala wurde auf 0,023 m/s eingestellt, die Position des Farbintensitäts-Drehreglers stand auf 50 %. Der Durchmesser des sichtbaren Spektrums beider Gefäße im parallelen Abschnitt wurde abgemessen. Alle Messwerte wurden mit der tatsächlichen Strömungsgeschwindigkeit verglichen.

Variieren der maximalen Geschwindigkeitsskala (CDI-Modus)

Die maximale Geschwindigkeitsskala zeigt im CDI-Modus die maximale durchschnittliche Geschwindigkeit an, die durch die farbkodierte Dopplersonographie dargestellt wird. Durch eine zu hoch gewählte maximale Geschwindigkeitsskala kann in der Tumordiagnostik ein langsamer Fluss kleiner Gefäße möglicherweise übersehen werden. In sechs Versuchsreihen an vier Tagen wurde deshalb gezeigt, unterhalb welches Pumpvolumens bei verschiedenen maximalen Geschwindigkeitsskalen kein Fluss mehr darstellbar ist und ab welchem Volumen das Gefäß durchgehend zu sehen ist. Dabei wurde mit niedrigen Pumpvolumina begonnen und die Pumpgeschwindigkeit langsam gesteigert. Der Ultraschallkopf war in der Halterung fixiert, der Winkel des Silikonschlauches zum Schallkopf betrug 50 Grad. Die Farbdopplerverstärkung stand konstant auf mittlerer Position. Zur Auswertung wurden die Mittelwerte der Pumpvolumina, ab denen gerade noch ein Fluss sichtbar war und ab denen ein durchgehender Fluss sichtbar war, graphisch dargestellt.

Winkelabhängige Grenze der Gefäßdarstellung bei unterschiedlicher maximaler Geschwindigkeitsskala (CDI-Modus)

In einem weiteren Versuch zur maximalen Geschwindigkeitsskala sollte die winkelabhängige Grenze der Gefäßdarstellung gezeigt werden. Analog dem zuvor beschriebenen Versuch wurde hier bei schlechten - also flachen - Winkeln getestet, wann ein sichtbarer Fluss abreißt. Die Winkelgrade wurden dabei sukzessive erhöht. Das Pumpvolumen war konstant auf 100 ml/h eingestellt.

Variieren des Dynamikbereiches der Energie (CDE-Modus)

Zum Vergleich von CDI und CDE wurden mit gleicher Versuchsanordnung wie bei den Versuchen zur maximalen Geschwindigkeitsskala sechs Versuchsreihen im CDE-Modus durchgeführt. Da das Signal hier vor allem von der Bewegungsenergie der korpuskulären Blutbestandteile erzeugt wird, existiert keine maximale Geschwindigkeitsskala. Eine sensitivere Geräteeinstellung wird unter anderem durch das Erhöhen des Dynamikbereiches der Energie ("dynamic range of energy"). erreicht. Die Mittelwerte aus sechs Versuchsreihen wurden, wie bei der Geschwindigkeitsskala im CDI, graphisch dargestellt.

Variieren der Farbintensität (CDI, CDE)

Durch die eine hoch eingestellte Farbintensität ("color level") können möglicherweise sehr dünne Gefäße noch dargestellt werden. Andererseits wächst durch ein Erhöhen der Farbintensität die Gefahr von Artefakten. Im Extremfall kann das ganze Bild als Farbflimmern erscheinen und dadurch das Erkennen von Gefäßen fast unmöglich machen. Die Versuche zur Farbintensität sollten zeigen, wie der Gefäßdurchmesser durch ein sensitiveres Einstellen verändert werden kann. Außerdem sollte ein Vergleich zwischen CDI und CDE gezogen werden.

Im CDI- und im CDE-Mode wurde bei fixiertem Ultraschallkopf die Farbintensität (engl. level) abgestuft, zur quantitativen Dokumentation wurde die Breite des in der Farbdarstellung bei 199,5 ml Pumpvolumen (entsprechend 0,14 m/s maximale Geschwindigkeit in der Schlauchmitte) sichtbaren Blutstromes gemessen. Im CDE-Modus wurden vier Versuchsreihen bei einem Dynamikbereich der Energie von 35 dB durchgeführt, im CDI-Modus je vier Versuchsreihen mit einer maximalen Geschwindigkeitsskala von 0,023 m/s und von 0,09 m/s. Die gemessene Breite des Gefäßes wurde jeweils für CDI und für CDE in Relation zum tatsächlichen Innendurchmesser gesetzt.

3.2 Ovarialtumoren

Einschlusskriterien in die Studie waren der klinisch, sonographisch, radiologisch oder anhand einer Tumormarker-Erhöhung gestellte Verdacht auf das Bestehen eines Ovarialtumors bzw. -karzinoms.

3.2.1 PATIENTINNEN

63 Patientinnen mit unklaren Ovarialbefunden wurden an der Frauenklinik der TU München präoperativ in einer Sitzung sonographisch und mit farbkodierter, gepulster Dopplersonographie untersucht. Das Durchschnittsalter betrug 58 Jahre, der Median lag bei 57 Jahren. Die jüngste Patientin war 24, die älteste 90 Jahre alt.

3.2.2 SONOGRAPHISCHE UNTERSUCHUNG

Die Untersuchungen wurden einen bis fünf Tage vor der Operation von drei Untersuchern nach einem festgelegten Protokoll vorgenommen und dauerten im Schnitt etwa 25 Minuten (20-45 Minuten). Die Untersuchung erfolgte in Rückenlage mit erhöhtem Becken, es wurde ein 5,0 MHz transvaginaler Schallkopf mit einem Sektorwinkel von 140 Grad (Acuson 128 XP, Fa. Acuson, CA, USA) verwendet.

Beide Ovarien wurden, falls möglich, dargestellt und hinsichtlich Größe und Struktur beurteilt. Anschließend erfolgte die Lokalisierung des Tumors und seine Charakterisierung bezüglich der folgenden, in der Literatur beschriebenen, etablierten Kriterien ([12], [17], [23], [28], [49], [53] (S. 497), [62], [81], [86] (S. 347), [90]).

Pathologische Prozesse wurden zur Volumenbestimmung in allen drei Dimensionen abgemessen. Das *Volumen* wurde nach der Formel zur Berechnung eines Ellipsoids bestimmt (Formel 6):

$$V[cm^{3}] = \frac{\pi \bullet a \bullet b \bullet c}{6}$$
(6)

Alle Raumforderungen wurden zunächst nach den Kriterien *solide*, *zystisch-solide* und *rein zystisch* klassifiziert.

Solide Prozesse wurden auf ihre *Binnenstruktur*, also die Beschaffenheit der Echomuster im Tumor, untersucht. Dabei wurden sie als überwiegend homogen oder als überwiegend inhomogen einstuft. Die *Randkontur* eines soliden Befundes oder Anteils erscheint bei benignen Tumoren im Ultraschall eher regelmäßig, bei malignen Prozessen eher unregelmäßig. Die *Begrenzung* vom umgebenden Gewebe kann sich scharf abzeichnen oder als Folge der Invasivität maligner Tumoren unscharf imponieren. Das Vorhandensein eines *echodichten Randsaumes*, das ein typisches Schallphänomen bei Fibroadenomen darstellt, wurde notiert.

Bei zystischen Gebilden oder Anteilen des Ovars wurden die *Struktur der Zystenwand* (glatt oder irregulär) und ihre *Dicke*, die Anzahl der *Septierungen* mit der genauen *Septendicke* und das Vorhandensein von *papillären Strukturen* sowie deren Größe beurteilt.

Zudem wurden bei allen Tumoren eine etwaige tumoröse *Infiltration* in Nachbarorgane und das Vorhandensein von *Aszites* registriert. Konnten Kriterien aufgrund einer schlechten Darstellbarkeit nicht beurteilt werden oder bestand Unsicherheit im Hinblick auf ein Kriterium, wurde es als "nicht untersuchbar" vermerkt und bei der Auswertung nicht berücksichtigt.

3.2.3 DOPPLERSONOGRAPHISCHE UNTERSUCHUNG

Anschließend an die sonographische Untersuchung erfolgte die dopplersonographische Beurteilung. Für sie wurde ein 5 MHz Pulse-Wave Doppler Transducer (Acuson 128 XP, Fa. Acuson, CA, USA) verwendet. Um eine bestmögliche Dopplerempfindlichkeit auch für langsame Flussgeschwindigkeiten zu erreichen, wurde der Wandbewegungsfilter mit 50 Hz möglichst niedrig gehalten Die Pulsrepetitionsfrequenz wurde auf 400-800 Hz gestellt. Die Farbdopplerverstärkung wurde bis knapp an die Grenze des Hintergrundrauschens gedreht und die maximale Geschwindigkeitsskala auf die niedrigste Stufe gestellt, so dass auch kleine Gefäße noch gut zur Darstellung kamen.

Zunächst erfolgte die **qualitative** Beurteilung des Durchblutungsstatus des Befundes (*Vaskularisation* vorhanden oder nicht vorhanden). Ein Gefäß wurde durch deutliche und konfluierende Farbpixel klassifiziert. Weiterhin wurde die annähernde *Gefäßzahl* des Tumors ermittelt und die *Gefäßverteilung* in Bezug zum Tumor bestimmt: Innerhalb eines

soliden Befundes oder in papillären Strukturen liegende Gefäße wurden als zentral bezeichnet. Gefäße am Tumorrand bzw. in einem Umfeld von bis zu 2 cm um ihn herum mit nachgewiesener Beziehung zum Tumor, in der Zystenwand oder in Septen wurden als peripher eingestuft.

Die **quantitative** dopplersonographische Bewertung wurde an Gefäßen vorgenommen, die innerhalb eines soliden Befundes oder in papillären Strukturen (zentral), an dessen Rand bzw. in einem Umfeld von bis zu 2 cm um ihn herum mit nachgewiesener Beziehung zum Tumor, in der Zystenwand oder in Septen (peripher) dargestellt werden konnten. Schlängelungen und atypische Verläufe wurden dabei gesondert vermerkt.

Zum Ableiten der Blutflusskurven wurde das Dopplerfenster gefäßdeckend platziert, der Insonationswinkel wurde korrigiert. Gefäße mit Winkelwerten über 60 Grad zum Schallkopf wurden verworfen. Eine kleine Gruppe (18 Patientinnen) wurde untersucht, ohne den Winkel zu berücksichtigen. Sie wurden aufgrund der geringen Fallzahl in die Auswertung mit einbezogen. Das abgeleitete Frequenzspektrum (Blutflusskurve) zeigt die unterschiedlichen Geschwindigkeiten der korpuskulären Blutbestandteile im Gefäßquerschnitt. Aus ihm wurden die folgenden Flussgeschwindigkeiten ausgemessen:

- Maximale systolische Flussgeschwindigkeit $(V_{max}[m / s])$
- Maximale enddiastolische Flussgeschwindigkeit $(V_{min}[m / s])$
- Gemittelte maximale Flussgeschwindigkeit $(V_{mean}[m / s])$

Anhand dieser Flussgeschwindigkeiten wurden die folgenden Indizes für den Gefäßwiderstand ermittelt:

• Resistance-Index (
$$RI = \frac{V_{max} - V_{min}}{V_{max}}$$
)

• Pulsatilitäts-Index (
$$PI = \frac{V_{max} - V_{min}}{V_{mean}}$$
)

• AB-Ratio $(AB = \frac{V_{max}}{V_{min}})$

Das Vorhandensein oder Fehlen eines "notch", einer frühdiastolischen Einkerbung im Verlauf der Blutflusskurve, wurde für jedes Gefäß vermerkt. Alle Parameter wurden getrennt für zentrale, periphere und für die Gesamtheit aller nachgewiesenen, in Beziehung zum Tumor stehenden Gefäße erfasst. Beim Vorhandensein mehrerer Gefäße wurden sowohl der Mittelwert als auch der Median, der Maximalwert und der Minimalwert analysiert.

Die Ergebnisse der sonographischen und dopplersonographischen Untersuchung wurden in ein Protokollblatt eingetragen und mit Angaben aus der Patientenakte zu Menopausenstatus, Zyklusphase, Hormoneinnahme und Umgebungsuntersuchungen vervollständigt. Die Daten wurden postoperativ mit der histologisch gesicherten Diagnose korreliert und der statistischen Auswertung zugeführt (siehe Kap. "statistische Verfahren").

3.2.4 HISTOPATHOLOGIE

Bei allen 63 Patientinnen wurde der Ovarialtumor operativ entfernt. Die Operationspräparate wurden im Institut für allgemeine Pathologie und pathologische Anatomie der TU München (Direktor: Prof. Dr. H. Höfler) feingeweblich untersucht. 41 der Ovarialtumoren wurden im histologischen Präparat als benigne, 22 als maligne klassifiziert. Dies entspricht einem Anteil maligner Tumoren von 35 %. Die Tumorklassifikation der malignen Tumoren erfolgte nach der FIGO-Stadieneinteilung (**Tab. 1**).

Die histologische Analyse der malignen Prozesse ergab elf seröse Adenokarzinome, vier Tumoren niedrig maligner Potenz (LMP), zwei undifferenzierte Adenokarzinome, ein muzinöses Adenokarzinom, ein klarzelliges Karzinom, ein Endometrioid-Karzinom, ein Dysgerminom und ein Non-Hodgkin-Lymphom (**Tab. 2**). Fünf der Karzinome und das Dysgerminom wurden im Stadium T₁ (beschränkt auf die Ovarien) ohne befallene Lymphknoten oder Metastasen diagnostiziert. Mit den vier Tumoren niedrig maligner Potenz konnten damit fast die Hälfte der Patientinnen in einem frühen Stadium mit kurativer Zielsetzung (FIGO I) operiert werden (**Tab. 3**). Neun der bösartigen Tumoren wurden im Stadium FIGO III aufgefunden, zwei Neoplasien hatten bereits Fernmetastasen gesetzt (FIGO IV).

FIGO I	Karzinom beschränkt auf die Ovarien
FIGO II	Karzinom befällt ein oder beide Ovarien mit Ausdehnung ins Becken

FIGO III	Peritonealmetastasen außerhalb des kleinen Beckens und /		
	oder regionärer Lymphknotenbefall (N1)		
FIGO IV	Fernmetastasen (ausgenommen Peritonealmetastasen)		

Tab. 1: Stadieneinteilung von Ovarialkarzinomen nach der Fédération Internationale de Gynécologie et d'Obstétrique (FIGO), vereinfacht nach [93].

Benigne Prozesse	n = 41	Tumorstadium (FIGO)	n = 21
	Pat.zahl	<u> </u>	Pat.zahl
Einfache Ovarialzyste	12	FIGO I	10
Seröses Zystadenom	12	FIGO II	0
Muzinöses Zystadenom	4	FIGO III	9
Endometriosezyste	4	FIGO IV	2
Teratom	3		
Hydrosalpinx	2	Non-Hodgkin-Lymphom nicht berü	cksichtigt
Hämatosalpinx	1		
Mesotheliom	1	Lymphknoten (N)	n = 17
Hämatom	1		Pat.zahl
Abszedierende Entzündung	1	N _x	6
		N ₀	6
Maligne Prozesse	n = 22	N ₁	5
	Pat.zahl		
Seröses Adenokarzinom	11	Dignität (G)	n = 17
Niedrig maligne Potenz (LMP)	4		Pat.zahl
Undifferenziertes Adenokarz.	2	G ₁	1
Muzinöses Adenokarzinom	1	G ₂	5
Klarzelliges Karzinom	1	G ₃	11
Endometrioid-Karzinom	1		•
Dysgerminom	1	Non-Hodgkin-Lymphom und LMP bei	
Non-Hodgkin-Lymphom	1	Lymphknoten und Dignität nicht eingeschlosse	

Tab. 2: Histologie der Ovarialtumoren (n=63)

Tab. 3: Klassifikation der malignen Ovarialtumoren (n=21)

Zwölf der benignen Tumoren waren einfache Ovarialzysten, weitere zwölf seröse Zystadenome. Außerdem wurden vier muzinöse Zystadenome, vier Endometriosezysten, drei Teratome, zwei Hydrosalpinges, eine Hämatosalpinx, ein Mesotheliom, ein Hämatom und eine abszedierende Entzündung gefunden (**Tab. 2**).

Das Alter der Patientengruppe mit histologisch gesichertem bösartigem Tumor unterschied sich nicht signifikant von der Gruppe der Patientinnen mit gutartigen Prozessen. Das Durchschnittsalter der Patientinnen mit bösartigen Tumoren lag bei 59 Jahren. Die jüngste Patientin war 29, die älteste 90 Jahre alt. Das Mittel der Patientinnen mit gutartigen Tumoren belief sich auf ein Alter von 57 Jahren mit einem Minimum von 24 und einem Maximum von 81 Jahren. Das Gesamtdurchschnittsalter betrug 58 Jahre.

Auch beim Menopausenstatus zeigte sich kein signifikanter Unterschied zwischen der malignen und der benignen Gruppe. Unter den untersuchten Frauen befanden sich 21 im prämenopausalen Alter und 42 in der Postmenopause (**Tab. 4**). Ein bösartiges Geschwür hatten sieben der Patientinnen in der Prämenopause (33,3 %), 15 der Frauen in der Postmenopause (36 %). Mit einem Wert von 1,0 im Fisher's Exact Test unterschieden sich die beiden Kategorien hinsichtlich der Häufigkeit bösartiger Tumoren nicht. Auch die Einnahme von Hormonen und unterschiedliche Zyklusstadien während der Untersuchung prämenopausaler Frauen ergab in diesem Kollektiv keine signifikante Abweichung im Hinblick auf die Diagnose.

	Gesamt	Benigne	Maligne
Prämenopausal	n = 21	n = 14 (66 %)	n = 7 (33 %)
Präovulatorisch	6	4	2
Ovulation	2	2	0
Postovulatorisch	10	5	5
Blutung	3	3	0
Postmenopausal	n = 42	n = 27 (64 %)	n = 15 (36 %)
Hormoneinnahme	8	6	2
Keine Hormoneinnahme	34	21	13

Tab. 4: Menopausenstatus der Patientinnen mit Ovarialtumor (n=63).

3.2.5 STATISTISCHE VERFAHREN

Aus sonographischen und dopplersonographischen Parametern sowie aus den anamnestischen und histologischen Daten wurde eine Datenbank in Microsoft[®] Excel 7.0 für Windows erstellt. Anschließend erfolgten die statistischen Auswertungen am Institut für medizinische Statistik und Epidemiologie der TU München (Direktor: Prof. Dr. A. Neiß) mit SAS (Statistical analysis system), Version 6.12 für Windows.

Im Rahmen der statistischen Auswertung wurde zunächst untersucht, welche der sonographischen und dopplersonographischen Parameter sich in maligner und benigner Gruppe signifikant unterschieden. Für die quantitativen Parameter wurde mit ROC-Kurven (receiver operating characteristics) der beste Grenzwert festgestellt. Alle Einzelparameter wurden in der univariaten Auswertung mit Hilfe des Chi²-Tests (nominal skalierte Paramenicht-parametrischen Rangsummentest nach ter) und dem Mann-Whitney-U (verhältnisskalierte Parameter) auf signifikante Unterschiede hin überprüft. Werte mit einer Irrtumswahrscheinlichkeit von p < 0.05 wurden als signifikant gewertet. Von allen Werten wurde die Sensitivität und die Spezifität getrennt berechnet sowie positive und negative Vorhersagewahrscheinlichkeit bestimmt. Die Sensitivität entspricht dem prozentualen Teil der malignen Tumoren, die nach dem jeweiligen Kriterium als maligne eingestuft werden, zu der Gesamtzahl der malignen Tumoren. Dementsprechend bezeichnet die Spezifität den mit dem Kriterium richtig erkannten Teil der gutartigen Neoplasien zur Gesamtzahl der benignen Tumoren. Durch die Verbesserung beider Werte kann die Validität, die Zahl der richtig eingestuften Prozesse bezogen auf die Gesamtzahl, erhöht werden.

Alle Parameter wurden, getrennt nach Sonographie und Dopplersonographie sowie in Kombination beider, in die multivariate Analyse mit einbezogen. Bei dieser wurden die Variablen mittels multipler logistischer Regression (forward stepwise method, SPSS for Windows, Version 8.0) auf ihre Unabhängigkeit voneinander überprüft und gewichtet. Um unter klinischen Bedingungen eine bestmögliche Treffsicherheit zu erzielen, wurde für Sonographie und Dopplersonographie getrennt und anschließend für beide Verfahren gemeinsam aus den Ergebnissen der multiplen logistischen Regression eine Entscheidungsmatrix erstellt. Die dadurch erhöhte Gesamt-Sensitivität und Gesamt-Spezifität der Methode wurde ermittelt.

3.3 Mammatumoren

In die Studie aufgenommen worden waren Patientinnen mit unklaren Befunden der Brust, die durch Palpation, Mammographie, Sonographie oder einen verdächtigen Lymphknoten-Befund auffällig geworden waren.
3.3.1 PATIENTINNEN UND UNTERSUCHUNGSDATEN

Die Datenanalyse für die Beurteilung von Mammatumoren erfolgte an insgesamt 219 Patientinnen mit tastbarem oder mammographisch dargestelltem Mammatumor, die an der Frauenklinik der TU München präoperativ sonographisch und mit farbkodierter, gepulster Dopplersonographie untersucht worden waren. Ein Teil der dadurch gewonnenen Daten wurde zum Erstellen des Entscheidungsbaumes für die multiple logistische Regression (n = 89, siehe [91]) und zum Trainieren der neuronalen Netze (Lerngruppe) verwendet. Diese Gruppe wird hier zur Unterscheidung als "retrospektive Patientengruppe" bezeichnet, obwohl die Untersuchungen mit prospektivem Ansatz erfolgt waren, also ohne Wissen der erst durch die Operation gewonnenen histologischen Diagnose. Zwei Patientinnen dieser Gruppe wiesen trotz des Verdachtes auf das Bestehen eines Mammatumors kein sonomorphologisches Korrelat auf, sie wurden deshalb aus der Auswertung ausgeschlossen.

Die übrigen 130 Patientinnen wurden nach Abschluss der multivariaten Analyse untersucht und als "prospektiv" klassifiziert. Die Beurteilung der Dignität erfolgte anhand des Entscheidungsbaumes und wurde mit dem in der Operation gewonnenen histologischen Resultat verglichen. Parallel dazu wurden ihre Daten in das Neuronale-Netze-Programm als Testgruppe aufgenommen.

Das Durchschnittsalter der retrospektiven Gruppe betrug 59 Jahre (Median 59 Jahre). Die jüngste Patientin war 24, die älteste 92 Jahre alt. In der prospektiven Gruppe belief sich das Durchschnittsalter auf 62 Jahre (Median 61, Minimum 25, Maximum 87 Jahre). Das Gesamtdurchschnittsalter und der Median lagen bei 60 Jahren.

Die sonographischen Untersuchungen erfolgten nach einem festgelegten Protokoll (0). Die Patientinnen wurden mit über den Kopf erhobenen Armen auf dem Rücken gelagert. Beide Mammae wurden mit B-Mode-Sonographie systematisch in allen vier Quadranten mäanderförmig untersucht. Pathologische Prozesse wurden zur Volumenbestimmung in allen drei Dimensionen abgemessen und nach den folgenden in der Literatur beschriebenen, etablierten Kriterien ([5], [24], [40], [58] (S. 383), [60] (S. 89), [70] (S. 20), [73], [74], [83], [96], [97]) untersucht.

Die *Volumenberechnung* erfolgte durch die Formel zur Bestimmung eines Ellipsoids (Formel 6, Kap. 3.2.2). Anschließend wurden die Befunde auf ihre *Echodensität* (hypodens, isodens, hyperdens), *Randkontur* (regelmäßig, unregelmäßig), *Begrenzung* (scharf, unscharf) und *Binnenstruktur* (homogen, inhomogen) untersucht. Die *Umgebungsreaktion* (Verdrängung oder Abbruch der Cooper'schen Ligamente) wurde bewertet ebenso wie das Vorhandensein von *hyperreflexiven Ausläufern*, und die folgenden *Schallphänomene*:

- laterale Schallauslöschung (einseitig oder beidseitig)
- dorsale (retrotumorale) Schallauslöschung
- dorsale (retrotumorale) Schallverstärkung

Die Orientierung des Tumors wurde erkundet, wobei die Haut die Referenzebene darstellte. Verlief der größte Tumordurchmesser annähernd parallel zur Haut ($0^{\circ} \pm 20^{\circ}$), wurde die Orientierung als tangential eingestuft. Verlief die Tumorachse senkrecht zur Haut ($90^{\circ} \pm 20^{\circ}$), wurde er als sagittal eingestuft. War die Ausdehnung des Tumors in allen Ebenen ähnlich, wurde er als Rundherd bezeichnet. Zudem wurde eine etwaige tumoröse *Infiltration* in Haut oder Muskel registriert. Konnten Kriterien aufgrund einer schlechten Darstellbarkeit nicht beurteilt werden oder bestand Unsicherheit im Hinblick auf ein Kriterium, wurde es als "nicht untersuchbar" vermerkt und bei der Auswertung nicht berücksichtigt.

Bei der dopplersonographischen Untersuchung wurde rein **qualitativ** die Tumordurchblutung (*Vaskularisation* vorhanden oder nicht vorhanden) beurteilt. Außerdem wurde die annähernde *Gefäßzahl* geschätzt und die *Gefäßverteilung* (zentral oder peripher gelegene Gefäße) dokumentiert.

Die **quantitative** dopplersonographische Bewertung wurde an Gefäßen vorgenommen, die innerhalb eines soliden Befundes oder in papillären Strukturen (zentral), an dessen Rand bzw. in einem Umfeld von bis zu 2 cm um ihn herum mit nachgewiesener Beziehung zum Tumor, in der Zystenwand oder in Septen (peripher) dargestellt werden konnten. Mit dem gepulsten Doppler wurden die maximale systolische Flussgeschwindigkeit (V_{max}), die maximale diastolische Flussgeschwindigkeit (V_{min}) und die gemittelte maximale Flussgeschwindigkeit (V_{mean}) gemessen und anhand dieser die Indizes für den Gefäßwiderstand (Resistance-Index, Pulsatilitätsindex, AB-Ratio) berechnet. Alle Parameter wurden getrennt für zentrale, periphere und für die Gesamtheit aller nachgewiesenen, in Beziehung zum Tumor stehenden Gefäße erfasst. Außerdem wurden fünf Gefäße der kontralateralen Brust mit einbezogen. Beim Vorhandensein mehrerer Gefäße wurden sowohl der Mittelwert als auch der Median, der Maximalwert und der Minimalwert analysiert.

Die Ergebnisse wurden in ein Protokollblatt eingetragen und mit Angaben aus der Patientenakte zu Menopausenstatus, Zyklusphase, Hormoneinnahme und Umgebungsuntersuchungen vervollständigt. Die Daten wurden postoperativ mit der histologisch gesicherten Diagnose korreliert und ausgewertet (siehe Kap. 3.3.5. und 3.3.6.).

3.3.2 HISTOPATHOLOGIE

Nach operativer Entfernung der 219 Tumoren wurden die histologischen Präparate am Institut für allgemeine Pathologie und pathologische Anatomie der TU München feingeweblich untersucht. 153 der Befunde waren histologisch maligne, darunter 58 in der retrospektiven und 94 in der prospektiven Gruppe. Dies entspricht einem Anteil maligner Prozesse von 70 % (66 % retrospektiv, 72 % prospektiv). Die Klassifikation der malignen Tumoren erfolgte nach der TNM-Stadieneinteilung (**Tab. 5**).

66 der Tumoren waren benigne Läsionen (30 in der als retrospektiv klassifizierten, 36 der als prospektiv klassifizierten Gruppe).

Die histologische Analyse der benignen Tumoren zeigte 22 Mastopathie-Knoten, drei Fibrosen, 26 Fibroadenome und 15 andere, benigne Tumoren (**Tab. 6**). Unter den malignen Tumoren waren drei Carcinomata in situ, 129 invasiv duktale Karzinome, 15 invasiv lobuläre Karzinome und sechs Mammakarzinome anderer Histologie (**Tab. 6**). 65 Karzinome wurden im Stadium T1 entfernt. 56 der bösartigen Tumoren wurden postoperativ als T_2 klassifiziert, 16 Neoplasien befanden sich im Stadium T_3 , 13 im Stadium T_4 . In 68 Fällen war keiner der histologisch untersuchten Lymphknoten befallen (N_0), in 60 Fällen lag ein Lymphknoten befall vor (N_1 oder N_2). Bei 25 Patientinnen war nicht bekannt, ob das Karzinom in Lymphknoten gestreut hatte, weil z.B. die Lymphknoten nicht entfernt worden waren (**Tab. 7**).

T _{is}	Carcinoma in situ
T_1	Tumor kleiner 2 cm
T ₂	Tumor größer als 2 cm, aber kleiner als 5 cm

T ₃	Tumor größer als 5 cm
T_4	Tumor jeder Größe mit Ausdehnung auf Brustwand oder
	Brusthaut
N ₀	Kein LK-Befall
N_1	Bewegliche ipsilaterale axilläre LK
N_2	Fixierte ipsilaterale axilläre LK
N_3	LK entlang der ipsilateralen A. mammaria interna

Tab. 5: Stadieneinteilung von Mammakarzinomen nach der TNM-Klassifikation, vereinfacht nach [93].

	retro- spektiv	pro- spektiv		retro- spektiv	Ś
Benigne Prozesse	30	36	Tumorstadium (T)	Pat.zahl]
	Pat.zahl	Pat.zahl	· · ·		
Mastopathie	8	14	T _{is}	2	
Fibrose	1	2	T ₁	25	
Fibroadenom	17	9	T ₂	23	
Andere benigne	4	11	T ₃	4	
L			Τ ₄	5	
Maligne Prozesse	59	94	Lymphknoten (N)	Pat.zahl	P
	Pat.zahl	Pat.zahl	N _x	5	
Carcinoma in situ	2	1	N ₀	34	
Invasiv duktal	50	79	N ₁	17	
Invasiv lobulär	5	10	N ₂	3	
Andere maligne	2	4	N ₃	0	

Tab. 6: Histologie der Mammatumoren (n = 219)

Tab. 7: Klassifikation der malignen Mammatumoren (n = 153)

Das Durchschnittsalter der Patientinnen mit Karzinomen betrug 64 Jahre (retrospektive Gruppe 63 Jahre, prospektive Gruppe 64 Jahre) mit einem Minimum bei 35 Jahren und einem Maximum bei 92 Jahren; der Median ergab 62 Jahre. Bei den gutartigen Tumoren lag das Durchschnittsalter ebenso wie der Median bei 53 Jahren (Durchschnittsalter retrospektiv 50 Jahre, prospektiv 55 Jahre), die jüngste Patientin war 24 Jahre, die älteste 82 Jahre alt.

	Gesamt	Benigne retrosp.	Benigne prospektiv	Maligne retrosp.	Maligne prospektiv
Anzahl	n = 217	n = 28	n = 36	n = 59	n = 94
	-				
Prämenopausal	72	15	19	12	26
Postmenopausal	145	13	17	47	68

Tab. 8: Menopausenstatus der Patientinnen mit Mammatumor (n=217). Retrosp. = retrospektiv.

Von den 217 Patientinnen befanden sich 72 in prämenopausalem Alter, 145 in der Postmenopause (**Tab. 8**). In der Gruppe der Frauen in der Postmenopause befanden sich sowohl retro- als auch prospektiv signifikant mehr maligne Tumoren (p = 0,015 im Fisher's Exact Test). Eine bösartige Geschwulst hatten 38 der Patientinnen in der Prämenopause (53 %), dagegen 115 der Frauen in der Postmenopause (79 %). Gutartige Tumoren hatten 34 Frauen in der Prämenopause und 30 in der Postmenopause (47 % versus 21 %).

3.3.3 DATENANALYSE MIT DER MULTIPLEN LOGISTISCHEN REGRESSION

Die statistischen Auswertungen für die retrospektive Patientengruppe waren analog der Auswertung der Adnex-Tumoren (Kap. 3.2.5.) am Institut für medizinische Statistik und Epidemiologie der TU München (Direktor: Prof. Dr. A. Neiß) erfolgt und wurden bereits an anderer Stelle veröffentlicht 0.

Dabei hatten von den 89 als retrospektiv klassifizierten Befunden zwei kein sonomorphologisches Korrelat aufgewiesen. Histologisch handelte es sich in beiden Fällen um mastopathische Herde. In der Analyse der Einzelkriterien hatten sich die sonomorphologischen Kriterien Volumen, laterale Schallauslöschung, dorsale Schallverstärkung, Echodensität und Infiltration als nicht signifikant erwiesen 0. Die Parameter Septenabbruch, sagittale Ausrichtung des Tumors, dorsale Schallauslöschung, unregelmäßige Randkontur, unscharfe Begrenzung und inhomogene Struktur erwiesen sich dagegen mit einem p-Wert von 0,0001 als signifikante Malignitätskriterien. Als bester sonomorphologischer Parameter erreichte die Begrenzung eine Treffsicherheit von 86 % (Sensitivität 81 %, Spezifität 96 %). Durch Kombination der sonomorphologischen Parameter Begrenzung, Randkontur und Schallauslöschung mittels multipler logistischer Regression in einem Entscheidungsbaum konnte die Treffsicherheit der sonographischen Untersuchung auf 91 % gesteigert werden (Sensitivität 88 %, Spezifität 96 %).

Bei den dopplersonographischen Parametern erwiesen sich die Vaskularisation, V_{max} aller Gefäße und die Resistance- und Pulsatilitätsindizes sowohl der Gesamtvaskularisation als auch der peripheren und der zentralen Vaskularisation als signifikant. Als bester dopplersonographischer Parameter erreichte der Mittelwert des PI der Gesamtvaskularisation eine Treffsicherheit von 82 % (Sensitivität 84 %, Spezifität 70 %).

In der multivariaten logistischen Regression erzielten fünf voneinander unabhängige Parameter in Kombination miteinander das beste Resultat: Als sonographische Kriterien die Begrenzung und die dorsale Schallauslöschung. Außerdem als quantitative dopplersonographische Parameter jeweils der Mittelwert des PI aller Gefäße und der systolischen Maximalgeschwindigkeit V_{max} aller Gefäße sowie die Vaskularisation als qualitativer dopplersonographischer Parameter (**Abb. 5**). Durch diese Kombination konnte die Treffsicherheit bei den als retrospektiv klassifizierten Patientinnen, anhand derer die multivariate Analyse erstellt worden war, auf 95 % erhöht werden. Die Spezifität betrug dabei 98 %, die Sensitivität 89 %. Angesichts der Tatsache, dass die besten Einzelkriterien sonographisch 86 % (Begrenzung) und dopplersonographisch 79 % Validität (PI gesamt) erreichten, zeigte sich so ein erheblicher Zugewinn an diagnostischer Treffsicherheit durch Kombination der Kriterien mittels multipler logistischer Regression. Insgesamt wurden mit der Entscheidungsmatrix damit nur drei Tumoren falsch diagnostiziert, histologisch zwei gutartige Befunde (eine Mastopathie, ein Fibroadenom) und ein muzinös invasives Karzinom.



Abb. 5: Diagnosebaum für die kombinierte sonographische und dopplersonographische Bewertung von Mammatumoren (n = 89). Die grauen Kästchen geben die Wahrscheinlichkeit für die Malignität eines Prozesses an (erstellt durch [91]).

Der Entscheidungsbaum wurde anhand der prospektiven Patientengruppe auf seine Wertigkeit hin untersucht. Um Veränderungen bei den Validitätswerten zu erläutern, wurden auch die Einzelparameter der prospektiven Patientengruppe auf ihre Sensitivität und Spezifität hin untersucht und mit der retrospektiven Patientengruppe verglichen.

Die Auswertung mit statistischen Methoden sollte einer Auswertungsweise mit künstlicher Intelligenz gegenübergestellt werden. Beide Auswertungsverfahren sollten hinsichtlich ihrer Treffsicherheit und der klinischen Anwendbarkeit miteinander verglichen werden.

3.3.4 AUSWERTUNG MIT NEURONALEN NETZEN

Für die Auswertung mit künstlicher Intelligenz wurde ein einfaches neuronales Netz verwendet (MedNet, Version 1.0, entwickelt an der Universität Stuttgart), das aus einer Schicht von Eingangsneuronen und einem Ausgangsneuron besteht.

Die Eingänge ei führen in diesem Netz zu den Neuronen der Eingangsschicht, deren Aufgabe darin liegt, die Information in numerische Werte umzuwandeln, z.B. "sagittal" = 1, "nicht sagittal" = 0. Um den Fehler durch nicht erbrachte Informationen bei nicht exakt beurteilbaren Kriterien möglichst klein zu halten. wurden die einzelnen Informationsinhalte auf verschiedene Neuronen verteilt, ein Neuron steht also z.B. für "sagittal" (ja = 1, nein = 0), eines für "transversal" (ja = 1, nein = 0) und eines für "rund" (ja = 1, nein = 0). Bei Prozessen, die weder als sagittal noch als transversal noch als rund eingestuft werden konnten, stehen die Eingänge in allen drei Neuronen demzufolge auf Null und liefen keinen Beitrag zur Berechnung des Ausgangswertes.

In der Lernphase wurde für die Ausgabewerte bei histologisch malignen Prozessen der numerische Wert 1 gesetzt, bei benignen der Wert 0. Als Lerndatensätze wurden in den meisten Versuchen die als "retrospektiv" klassifizierten Daten der Patientinnen mit Mammatumoren verwendet, als Testdatensätze die "prospektiven" Daten eingesetzt. Für die Testdaten wurde ein Schwellenwert von 0,5 festgelegt: Alle Ereignisse größer oder gleich dieser Schwelle wurden als maligne, alle Ereignisse unterhalb als benigne angenommen.

Das neuronale Netz wurde zum Vergleich mit der Auswertungsmethode der multiplen logistischen Regression auf mehrere Arten trainiert und getestet. In den meisten Versuchen wurden die retrospektiven Daten als Lernsätze verwendet. Als Eingabedaten dienten in verschiedenen Schritten

- ausschließlich die sonomorphologischen Kriterien,
- ausschließlich die dopplersonographischen Kriterien,
- alle sonomorphologischen und dopplersonographischen Kriterien,
- alle sonomorphologischen Kriterien in Kombination mit den qualitativen Gefäßparametern,

- alle sonomorphologische Kriterien in Kombination mit den qualitativen Gefäßparametern und RI und PI,
- alle sonomorphologische Kriterien in Kombination mit den qualitativen Gefäβparametern und den quantitativen Parametern RI, PI und v_{max},
- nur die in die Entscheidungsmatrix eingeflossenen Kriterien (Vaskularisation, Begrenzung, Mittelwert PI gesamt, Mittelwert v_{max} gesamt, dorsale Schallauslöschung).

Um etwaige Fehler aus den ersten Versuchen mit eventuell noch nicht ganz ausgefeilter Methodik und noch nicht ganz routinierten Untersuchern zu eliminieren, wurden zwei Trainingsrunden mit Lerndatensätzen aus gemischt retro- und prospektiven Patientinnen durchgeführt. Die Auswahl in Trainings- und Testgruppe erfolgte dabei nach dem Zufallsprinzip (zunächst die mit ungeraden Datensätzen nummerierten Patientinnen als Trainings- und mit geraden Zahlen nummerierten als Testgruppe, anschließend umgekehrt). Außerdem wurde ein Lerndurchgang mit allen retrospektiven und prospektiven Daten gestartet, bei dem die Eingabeparameter, beginnend bei den Eingängen mit den schwächsten Gewichten, sukzessive reduziert wurden. Dadurch sollte festgestellt werden, wie viele Eingangsparameter das Netz benötigt, um noch alle Datensätze korrekt lernen zu können.

4 ERGEBNISSE

4.1 Reproduzierbarkeit von dopplersonographischen Blutflussanalysen

Bei allen Messungen betrug der durchschnittliche, durch Auslitern ermittelte Pumpfehler -0,4 ml/h bei Abweichungen von maximal -5,5 bis +2,3 ml/h. Somit lag die Abweichung der Strömungsgeschwindigkeit bei durchschnittlich 0,4 % (maximal 5,5 %). Da die Schallausbreitungsgeschwindigkeit in Ultraschallgel niedriger ist als die durchschnittliche Schallausbreitungsgeschwindigkeit von 1540 m/s im menschlichen Gewebe, auf die die Doppler-Geräte geeicht sind, ist eine falsch zu niedrige Messung der Doppler-Frequenzen anzunehmen. Der Fehler bei Durchführen der Versuche in einem Wasserbad betrüge bei einer Schallausbreitungsgeschwindigkeit von 1495 m/s in Wasser 3 %, bei dem dichteren Ultraschallgel ist er geringer als 3 % anzunehmen, so dass er vernachlässigt wurde.

4.1.1 GEPULSTE DOPPLERSONOGRAPHIE

Geschwindigkeitsmessung bei verschiedenen Pumpvolumina

Während die mit dem Doppler gemessenen Werte bei höheren Geschwindigkeiten das 10-%-Konfidenzintervall kaum überschritten und der Mittelwert aus allen Messungen nahezu die errechnete Sollgeschwindigkeit erreichte, wichen sie unterhalb eines Pumpvolumens von 50 ml/h (0,018 m/s Gesamtgeschwindigkeit im Rohr, entsprechend 0,035 m/s Maximalgeschwindigkeit in der Flussmitte) deutlich von der Sollgeschwindigkeit ab. Mit Ausnahme der Flussgeschwindigkeit von 5 ml/h (0,0035 m/s Maximalgeschwindigkeit), bei denen in sieben von acht Versuchen kein Fluss mehr angezeigt wurde, wurden die niedrigen Flussgeschwindigkeiten vom Gerät sonst tendenziell deutlich überschätzt. So lagen auch die Mittelwerte außerhalb des 90-%-Konfidenzintervalles (**Abb. 6**).



Pumpvolumen [ml/h]

Abb. 6: Genauigkeit der dopplersonographischen Geschwindigkeitsmessung in acht Versuchsreihen. Die durchgezogene Linie stellt die Mittelwerte, die grob gestrichelte Linie die errechneten Soll-Werte dar. Fein gestrichelt das 90-%-Konfidenzintervall. Gleiche Werte der Einzelmessungen (Δ) sind sich überlagernd eingezeichnet.

Winkelabhängigkeit der Geschwindigkeitsmessung

Die im Gerät eingebaute Korrekturfunktion für den Winkel zwischen Schallstrahl und Gefäßachse ist eine Cosinus-Funktion. Ihre Ableitung, ausgehend von einer Geschwindigkeit von 0,07 m/s und 50 Grad, wurde in einem Diagramm aufgetragen (**Abb. 7**).

Die dopplersonographische Geschwindigkeitsmessung bei Simulation verschiedener Winkel unter In-Vivo-Bedingungen entsprach unterhalb von 60 Grad praktisch der Cosinus-Funktion. Der Bereich der exaktesten Messungen lag zwischen 35 und 55 Grad: Hier überbzw. unterschritt keine Geschwindigkeit 0,08 bzw. 0,06 m/s. Der maximale prozentuale Messfehler betrug 14,3 %, der durchschnittliche Messfehler betrug 0,8 %. Bei Winkeln unterhalb von 35 Grad tendierten die Messungen, angelehnt an die Cosinus-Funktion, zur leichten Unterschätzung der Geschwindigkeit (minimal 0,05 m/s, entsprechend einem maximalen Messfehler von 28,6 %), der durchschnittliche prozentuale Messfehler betrug 17,5 %. Bei Winkeln oberhalb 55 Grad wurden die Geschwindigkeiten leicht überschätzt (maximal 0,09 m/s, durchschnittlicher prozentualer Messfehler 16,3 %). Oberhalb von 60 Grad wich die Geschwindigkeitsmessung - wie die Cosinus-Funktion - exponentiell von der Soll-Gerade ab. Bei Winkeln größer als 80 Grad wurde der durchschnittliche prozentuale Messfehler größer als 100 % (**Abb. 8**). Die Werte der Messungen lagen oberhalb von 60 Grad jedoch deutlich unter denen der Cosinus-Funktion, also näher bei der Soll-Geraden. Die gemessenen Geschwindigkeiten bei Winkeln von 88 Grad schwankten zwischen 0,35 bis 0,53 m/s, der durchschnittliche Messfehler lag hier bei 540 %.

Die Ergebnisse der an verschiedenen Tagen durchgeführten zehn Messreihen unterscheiden sich bei Winkeln unterhalb von 60 Grad kaum. Oberhalb von 60 Grad waren sie in zunehmendem Maße Schwankungen unterworfen, die jedoch zufälliger und nicht systematischer Natur waren.



Abb. 7: Genauigkeit der dopplersonographischen Geschwindigkeitsmessung bei **Winkelvariation**. Durchgezogene Linie: Errechnete Soll-Geschwindigkeit bei 100 ml/h Pumpvolumen (0,07 m/s). o: Winkelkorrekturfunktion der Gerätes (Cosinus-Funktion). x: Dopplersonographische Einzelmessungen. Ergebnisse aus zehn Versuchsreihen.



Abb. 8: Prozentualer Messfehler bei der dopplersonographischen Geschwindigkeitsmessung mit verschiedenen Winkeln, Mittelwerte aus zehn Messreihen. Negative Prozentzahlen: Abweichung der gemessenen Geschwindigkeit nach unten.

Variieren der Empfangsverstärkung

Bei einer Empfangsverstärkung von 50 % (Drehregler in mittlerer Position) wurde nach Mittelung der Ergebnisse der acht Messreihen die Sollgeschwindigkeit von 0,14 m/s erreicht. Die Mittelwerte der Messungen mit höherer Empfangsverstärkung waren in Vergleich leicht erhöht, die der Messungen mit niedrigerer Empfangsverstärkung leicht erniedrigt (Abb. 9). Dabei fiel die prozentuale Abweichung vom Soll-Wert bei den niedrigeren Empfangsverstärkungen etwas größer aus: Der Mittelwert bei der niedrigsten Verstärkungsstufe (12,5 %) betrug 0,1075 m/s, was einer prozentualen Abweichung von 23,2 % entspricht, der Mittelwert bei der höchsten Empfangsverstärkung (87,5 %) war 0,1638 m/s (prozentuale Abweichung 17 %) (Abb. 10).



Abb. 9: Dopplersonographische Geschwindigkeitsmessung bei unterschiedlich hoher **Empfangsverstärkung**. Messergebnisse aus je acht Versuchsreihen. Auf der x-Achse sind die Positionen des Empfangverstärkungs-Drehreglers aufgetragen. Gleiche Werte werden nur einfach aufgeführt. Durchgezogene Linie: Errechnete Sollgeschwindigkeit.



Doppler-Empfangsverstärkung (Position des Drehreglers)

Abb. 10: Mittlere prozentuale Abweichung der dopplersonographischen Geschwindigkeitsmessung bei unterschiedlich hoher Empfangsverstärkung (x-Achse: Position des Drehreglers). Negative Prozentzahlen: Abweichung der Geschwindigkeit nach unten.

Dopplerfenster

Bei dem das ganze Gefäß umfassenden Dopplerfenster entsprachen sechs der acht Messungen dem Soll-Wert von 0,14 m/s, zwei Werte wichen um 0,01 m/s davon ab. Der Mittelwert aus acht Messungen ergab somit genau 0,14 m/s. Bei der Dopplerableitung allein in der Flussmitte waren die Werte deutlich erhöht, der Mittelwert aus allen Messungen betrug 0,1638 m/s. Bei der Cursorposition am Rand wurden die Geschwindigkeiten mit einem Mittelwert von 0,1163 tendenziell zu niedrig gemessen (Abb. 11).



Abb. 11: Dopplersonographische Geschwindigkeitsmessung bei unterschiedlich gesetztem **Dopplerfenster**. Messergebnisse aus je acht Versuchsreihen bei Ableitung der Dopplerflusskurve mit je nur zentralem, nur randständigem und den ganzen Schlauch umfassendem Cursor. Gleiche Werte sind nicht doppelt aufgeführt. Durchgezogene Linie: Errechnete Sollgeschwindigkeit.

Genauigkeit der Geschwindigkeitsmessung in unterschiedlicher Tiefe

Der Mittelwert der dopplersonographisch abgeleiteten maximalen Geschwindigkeit in 1 cm Tiefe lag bei 0,15 m/s, in 4 cm Tiefe bei 0,1413 m/s. Beide waren damit gegenüber dem errechneten Soll-Wert von 0,14 m/s leicht erhöht. Der Unterschied zwischen beiden Werten war nicht signifikant. Die Einzelwerte der acht Messungen in 1 cm Tiefe streuten etwas weniger (Abweichung vom Soll-Wert maximal -0,01 bis 0,04 m/s) als bei den Messungen in 4 cm Tiefe (Abweichung vom Soll-Wert maximal -0,03 bis 0,04 m/s) (Abb. 12).



Abb. 12: Tiefenabhängigkeit der dopplersonographischen Geschwindigkeitsmessung. Messergebnisse aus acht Versuchsreihen bei Ableitung der Dopplerflusskurve in 1 cm und in 4 cm Tiefe, gleiche Werte sind nicht doppelt aufgeführt. Durchgezogene Linie: Errechneter Sollwert.

Dopplersonographische Geschwindigkeitsmessung bei Simulation untereinander liegender Gefäße

Die Geschwindigkeit bei der dopplersonographischen Messung betrug in sechs Versuchsreihen bei einem Pumpvolumen von 100 ml/h konstant 0,07 m/s sowohl im oberen, in 1 cm Höhe vom Schallkopf entfernten, als auch im unteren, in 2 cm Entfernung vom Schallkopf liegenden Gefäß. Es bestand damit kein Unterschied zwischen der Ableitung im oberen und der Ableitung im unteren Gefäß.

Einfluss des Hämatokritwertes

Angefangen von unverdünntem Erythrozytenkonzentrat mit einem Hämatokrit von 82,7 % in der ersten (Hämoglobin 24,5 g/dl) und 79,6 % in der zweiten Messreihe (Hämoglobin 24,4 g/dl) bis hinunter zu einem Hämatokrit von 0,00001 % betrug die abgeleitete Geschwindigkeit in beiden Messreihen 0,07 m/s. Damit entsprach sie immer der realen maximalen Geschwindigkeit von 0,07 m/s. Das Doppler-Frequenzspektrum war unterhalb eines Hämatokrits von 0,19 % weniger dicht. Unterhalb von 0,01 % war es deutlich weniger dicht, aber noch gut sichtbar und differenzierbar. Im Farbduplex stellte sich bis zu einem Hämatokrit von 0,01 % das Gefäß noch durchgehend dar. Bei geringeren Hämatokritwerten konnte nur noch das Aufleuchten einzelner Farbpunkte gesehen werden, die mit weiter steigender Verdünnung immer kleiner wurden und sich weiter voneinander entfernten.

4.1.2 FARBKODIERTE DOPPLERSONOGRAPHIE

Farbkodierte Dopplersonographie bei Simulation untereinander liegender Gefäße

In der Farbduplexableitung imponierte das obere Gefäß konstant dicker als das untere Gefäß. Die Messungen des Farbspektrums ergaben eine deutliche Überschätzung des Innendurchmessers sowohl im oberen als auch im unteren Gefäß (Abb. 13).



Abb. 13: Schlauchdurchmesser bei Farbduplex-Darstellung untereinander liegender Schläuche. In sechs Versuchsreihen wurde der Durchmesser des Farbstrahles gemessen, gleiche Werte sind nicht doppelt aufgeführt. Durchgezogene Linie: Tatsächlicher Innendurchmesser des Silikonschlauches.

Variieren der maximalen Geschwindigkeitsskala (CDI-Modus)

Je höher die maximale Geschwindigkeitsskala eingestellt wurde, um so höhere Geschwindigkeiten wurden benötigt, um einen Blutfluss darzustellen: Bei einer Geschwindigkeitsskala von 0,023 m/s reichte schon ein Pumpvolumen von durchschnittlich 3 ml/h aus (entsprechend einer Gesamtgeschwindigkeit von 0,0011 m/s bzw. einer Spitzengeschwindigkeit in Flussmitte von 0,0021 m/s), um einzelne Farbduplexsignale zu gewinnen; der Gefäßverlauf ließ sich ab 5 ml/h (0,0018 m/s) rekonstruieren. Bei einer maximalen Geschwindigkeitsskala von 0,09 m/s wurden dafür 13,5 bzw. 18,5 ml/h benötigt (0,0048 bzw. 0,0065 m/s), bei einer Geschwindigkeitsskala von 0,16 m/s sogar 27 bzw. 36 ml/h (0,0095 bzw. 0,0127 m/s) (Abb. 14). Eine lückenlose Darstellung des ganzen Schlauches gelang bei einer maximalen Geschwindigkeitsskala von 0,023 m/s ab 6,5 ml/h (0,0023 m/s), bei 0,09 m/s ab 23,5 ml/h (0,0083 m/s) und bei 0,16 m/s ab 44,5 ml/h (0,0157 m/s).



Abb. 14: Darstellung eines Blutflusses bei unterschiedlichen maximalen Geschwindigkeitsskalen (CDI-Modus). Schraffierte Fläche: Pumpvolumina, bei denen kein Blutfluss dargestellt werden konnte. Karierte Fläche: Punktuelles Aufleuchten des farbkodierten Dopplersignals. Ausgemalte Fläche: Rekonstruktion eines Gefäßverlaufes möglich. Darüber: Deutlicher Fluss im ganzen Schlauchlumen.

Winkelabhängige Grenze der Gefäßdarstellung bei unterschiedlicher maximaler Geschwindigkeitsskala (CDI-Modus)

Je höher die maximale Geschwindigkeitsskala gestellt wurde, umso steilere Winkel wurden benötigt, um bei einem Pumpvolumen von 100 ml/h (durchschnittliche Flussgeschwindigkeit im Gefäß 0,035 m/s, maximale Geschwindigkeit in der Flussmitte 0,07 m/s) einen Fluss in der farbkodierten Dopplersonographie darzustellen. War bei einer maximalen Geschwindigkeitsskala von 0,023 m/s bereits bei Winkeln von durchschnittlich 85,1 Grad (83 bis 87 Grad) ein Fluss sichtbar, musste bei einer Geschwindigkeitsskala von 0,09 m/s ein Winkel von durchschnittlich 80,1 Grad (74 bis 85 Grad) und bei einer Geschwindigkeitsskala von 0,17 m/s ein Winkel von durchschnittlich 75,9 Grad (71 bis 81 Grad) eingehalten werden (Abb. 15).



Abb. 15: Grenze der Farbdarstellung bei flachen Winkeln und variierter maximaler Geschwindigkeitsskala (CDI-Modus). Karierte Fläche: Winkel, bei denen kein Blutfluss darstellbar war. Ausgemalte Fläche: Winkel, bei dem Blutfluss zur Darstellung kam.

Variieren des Dynamikbereiches der Energie (CDE-Modus)

Mittels Color Doppler Energy konnten noch geringere Geschwindigkeiten in dem Schlauch sichtbar gemacht werden als mit der konventionellen farbkodierten Dopplersonographie. Bereits ab einem Pumpvolumen von 1 ml/h (durchschnittliche Geschwindigkeit im Rohr 0,00035 m/s, maximale Geschwindigkeit in Flussmitte 0,0007 m/s) zeigten sich bei einem Dynamikbereich der Energie von 40 dB einzelne Farbpunkte. Ab einem durchschnittlichen Pumpvolumen von 2,2 ml/h konnte der Schlauchverlauf rekonstruiert werden, und ab einem Pumpvolumen von 4 ml/h wurde das ganze Schlauchlumen lückenlos dargestellt (Abb. 16). Bei niedrigeren Dynamikbereichen stiegen die Geschwindigkeiten leicht, bei 20 dB konnten ab durchschnittlich 4 ml/h (0,0014 m/s) erste Farbpunkte gesehen werden; ein lückenloser Verlauf konnte ab 9 ml/h (0,0032 m/s) dargestellt werden.



Abb. 16: Grenze der Darstellung eines Blutflusses bei unterschiedlichen Dynamikbereichen der Energie (CDE-Modus). Schraffierte Fläche: Pumpvolumina, bei denen kein Blutfluss dargestellt werden konnte. Karierte Fläche: Punktuelles Aufleuchten des farbkodierten Dopplersignals. Ausgemalte Fläche:

Rekonstruktion eines Gefäßverlaufes möglich. Darüber: Deutlicher Fluss im ganzen Schlauchlumen.

Variieren der Farbintensität (CDI, CDE)

Die Breite des farbdopplersonographisch aufgezeichneten und anschließend abgemessenen Farbspektrums im CDI-Modus entsprach bei einer maximalen Geschwindigkeitsskala von 0,09 m/s und bei mittlerer Position des Farbintensitäts-Drehreglers ziemlich genau dem tatsächlichen Innendurchmesser (Mittelwert aus vier Messungen 1,05 mm, **Abb. 17**). Der Schlauchinnendurchmesser bei der minimalen Einstellung des Farbintensitäts-Drehreglers wurde bei der Geschwindigkeitsskala von 0,09 m/s im Schnitt unterschätzt (Mittelwert 0,55 mm), bei der maximalen Einstellung um knapp das Doppelte überschätzt (Mittelwert 1,95 mm).



Farbintensität (Position des Drehreglers)

Abb. 17: Relation der Breite des Farbspektrums im CDI-Modus zum tatsächlichen Innendurchmesser des Schlauches (durchgezogene Linie) in Abhängigkeit von der eingestellten Farbintensität und der maximalen Geschwindigkeitsskala. Auf der x-Achse sind die Positionen des Farbintensitäts-Drehreglers aufgetragen. Je vier Versuchsreihen bei einer Geschwindigkeitsskala von 0,09 m/s (o = Einzelwerte, x = Mittelwert) und bei 0,023 m/s ($\Delta =$ Einzelwerte, + = Mittelwert).

Bei einer maximalen Geschwindigkeitsskala von 0,023 m/s dagegen lagen die Werte schon bei der minimalen Drehregler-Position von 12,5 % mit 1,2 mm zu hoch. Bei der Geschwindigkeitsskala von 0,023 m/s zeigte der Mittelwert bei mittlerer Einstellung des Drehreglers (50 %) mit 2,35 mm mehr als das Doppelte und bei maximaler Einstellung (87,5 %) mit 3,4 mm über das Dreifache des tatsächlichen Innendurchmessers an (**Abb. 17**).

Im CDE-Modus wurde bei der üblichen Einstellung eines "dynamic range of energy" von 35 dB der Innendurchmesser bei mittlerer Drehregler-Position (50 %) nur leicht überschätzt (Mittelwert aus vier Messungen 1,2 mm). Bei minimaler Position des Drehreglers (12,5 %) wurde er mit durchschnittlich 0,5 mm deutlich unterschätzt, bei maximaler Position (87,5 %) mit durchschnittlich 1,7 mm deutlich überschätzt.



Abb. 18: Relation der Breite des Farbspektrums im **CDE-Modus** zum tatsächlichen Innendurchmesser des Schlauches (durchgezogene Linie) in Abhängigkeit von der eingestellten **Farbintensität**. Werte und Mittelwerte aus je vier Versuchsreihen. Auf der x-Achse sind die Positionen des Farbintensitäts-Drehreglers aufgetragen.

4.2 Ovarialtumoren

Es wurden 63 Patientinnen mit Ovarialtumoren untersucht. Bei allen konnte ein sonomorphologisches Korrelat gefunden werden, so dass alle Befunde in die Auswertung einfließen konnten.

4.2.1 SONOGRAPHISCHE DIGNITÄTSBEURTEILUNG

In die vorliegende Studie wurde ein möglichst großes Spektrum der von anderen Arbeitsgruppen beschriebenen sonomorphologischen Parameter einbezogen, soweit sie zur Dignitätsbeurteilung von Adnex-Tumoren sinnvoll erschienen. Die Signifikanzgrade der einzelnen Kriterien wurden im Fisher's Exact Test ermittelt, außerdem wurden die jeweilige Sensitivität, Spezifität, Treffsicherheit und die positiven und negativen Voraussagewahrscheinlichkeiten berechnet. Fünf der 14 Kriterien stellten sich mit einer Irrtumswahrscheinlichkeit von p < 0,05 als hochsignifikant heraus (**Tab. 9**).

Mehrere der anhand von ROC-Kurven ermittelten Grenzwerte für das *Tumorvolumen* zeigten mit einem Signifikanzgrad von p < 0,01 signifikante Ergebnisse (Abb. 19). Die beste Sensitivität und Spezifität ergab sich für einen Grenzwert des Volumens von 111 ml. Wurde für alle größeren Tumoren Malignität angenommen, konnten 73 % der malignen und 64 % der benignen Tumoren richtig beurteilt werden. Insgesamt konnten 70 % der Befunde bei diesem Grenzwert richtig eingestuft werden. Vergleichbare Treffsicherheiten wurden mit 67 % bzw. 69 % auch bei Cut-off-Volumina von 85 ml (Sensitivität 68 %, Spezifität 69 %) und 97 ml (Sensitivität 63 %, Spezifität 76 %) erreicht.

Auch das Vorhandensein von *papillären Strukturen* in Zysten (p < 0,01), außerdem die *Irregularität von Zystenwänden* (p < 0,001)und die *Infiltration in Nachbarorgane* (p < 0,01) erwiesen sich als zuverlässige Kriterien für das Vorliegen eines malignen Prozesses. Allerdings konnten diese Parameter entsprechend der Grundstruktur des Tumors nicht in allen Fällen beurteilt werden. Als bestes Malignitätskriterium stellte sich das Vorhandensein von *soliden Anteilen* eines Tumors mit einer Sensitivität von 95 %, einer Spezifität von 68 % und einer Validität von 78 % heraus (p = 0,00002).

Nicht signifikant waren in dieser Untersuchung der Nachweis von Aszites, die übrigen Parameter zur Beurteilung von Zysten (*Dicke der Zystenwand*, *Papillendicke*, *Septierung*, *Septendick*e; siehe **Tab. 9**) und alle Kriterien zur Beurteilung von soliden Prozessen (*Homogenitä*t, echodichter Randsaum, Randkontur, Begrenzung).

Parameter	Signifikanz-	Malignitätskrit.	n	Sensitivität	Spezifität	Treffsi-
	grad p	Grenzwerte		[%]	[%]	cherheit
						[%]

Allgemein

Volumen	< 0,01	> 85 ml	61	73	64	67
	< 0,01	> 97 ml	61	68	69	69
	< 0,01	>111 ml	61	63	76	70
Struktur	< 0,0001	solide Anteile	63	95	68	78
Aszites	n.s.	vorhanden	52	20	95	73
Infiltration	< 0,01	vorhanden	53	28	100	75

Zyste

_		

Zystenwand	< 0,001	irregulär	53	59	92	81
Dicke der	n.s.	> 3mm	24	77	55	67
Zystenwand						
Papilläre	< 0,01	vorhanden	48	67	81	77
Strukturen						
Papillendicke	n.s.	> 3mm	9	100	0	67
Septierung	n.s.	vorhanden	50	79	47	56
Septendicke	n.s.	> 3mm	27	82	44	59

Solider Anteil

Homogenität	n.s.	inhomogen	20	57	67	60
Echodichter	n.s.	vorhanden	16	27	60	38
Randsaum						
Randkontur	n.s.	unregelmäßig	20	64	67	65
Begrenzung	n.s.	unscharf	20	21	100	45

Tab. 9: Spezifitäten, Sensitivitäten und Treffsicherheiten der sonomorphologischen Parameter, die in die Untersuchung eingingen. Neben der Fallzahl n (maximal 63) Angabe des jeweiligen Malignitätskriteriums (Malignitätskrit.) bzw. des entsprechenden Grenzwertes, n.s. = nicht signifikant.



SPEZIFITÄT

Abb. 19: ROC-Kurve für die Bestimmung der optimalen Grenzwerte des **Volumens** für die Dignitätsbeurteilung von Adnex-Tumoren, n = 61. o = Markierung der drei besten Grenzwerte.

Im Anschluß an die univariate Beurteilung der einzelnen sonomorphologischen Parameter wurden die Werte in eine multivariate Untersuchung einbezogen, in der mit Hilfe der multiplen logistischen Regression die beste Kombination voneinander unabhängiger Variablen berechnet wurde. Die einzelnen Parameter wurden dabei, abhängig von ihrem Vorhersagewert in der univariaten Analyse, unterschiedlich stark gewichtet. Ein aufgrund der Ergebnisse der multivariaten Analyse angefertigter Diagnosebaum soll zur Entscheidungsfindung führen: Betrug die Wahrscheinlichkeit gleich oder mehr als 50 %, wurde ein maligner Tumor angenommen. Bei einer Wahrscheinlichkeit unter 50 % wurde der Tumor als benigne klassifiziert.

Für die Sonomorphologie konnte durch die multivariate Analyse keine Steigerung der Treffsicherheit im Vergleich zu den univariaten Ergebnissen erzielt werden. Das aussagekräftigste univariate Kriterium, das Vorhandensein eines soliden Anteils, erreichte eine Sensitivität von 95 %, eine Spezifität von 68 % und eine Validität von 78 % (**Tab. 11**, **Abb. 20**). Insgesamt konnte damit die Dignität von 49 der 63 Tumoren richtig beurteilt werden. Nur ein seröses Adenokarzinom wies keinen soliden Anteil auf. Dagegen wurden 13 benigne Tumoren aufgrund ihres soliden Anteils als falsch-positiv bewertet, darunter 4 Zysten (Corpus luteum-Zyste, Epithelzyste und zwei Inklusionszysten), eine Hämatosalpinx, ein Hämatom, zwei Teratome, eine Endometriosezyste und zwei seröse Zystadenome. Die Kombination mit anderen Kriterien erbrachte eine Erhöhung entweder der Sensitivität oder der Spezifität nur auf Kosten des jeweils anderen Parameters ohne Steigerung der Validität.



Abb. 20: Entscheidungsmatrix für die Anwendung der multiplen logistischen Regression zur sonographischen Dignitätsbeurteilung von Adnex-Tumoren (n = 63). Die Prozentzahlen geben die Wahrscheinlichkeit für das Vorliegen von Malignität bei Erfüllung des vorgeschalteten Kriteriums an (p < 0,001).

4.2.2 DOPPLERSONOGRAPHISCHE DIGNITÄTSBEURTEILUNG

Mit der farbkodierten Dopplersonographie wurde das Vorhandensein von Gefäßen im Tumor und im Tumorumfeld mit nachweislichem Zusammenhang zu diesem untersucht (Vaskularisation) und die annähernde Gefäßzahl ermittelt. An nachgewiesenen Gefäßen wurden mit der gepulsten Dopplersonographie Blutflusskurven zur Beurteilung der Flussgeschwindigkeiten (V_{max}, V_{min}, V_{mean}) und der Gefäßwiderstände (RI, PI, AB-Ratio) abgeleitet.

73 % der benignen und 90 % der malignen Tumoren wiesen Blutgefäße auf, so dass sich hinsichtlich der *Gesamtdurchblutung* kein signifikanter Unterschied ergab. Dagegen

zeigten 95 % der bösartigen Prozesse, aber nur 10 % der gutartigen Tumoren Gefäße in soliden Tumorarealen (*zentrale Vaskularisation*). Mit dem Nachweis zentraler Vaskularisation wurden 95 % Sensitivität, 90 % Spezifität und eine Validität von 92 % erreicht. Damit war die zentrale Vaskularisation der sensibelste Parameter (Irrtumswahrscheinlichkeit p < 0,00001 %) allerdings nur nach vorherigem Ausschluss der Patientinnen, die insgesamt keine Vaskularisation aufwiesen.

Auch die *Gefäßzahl* bei benignen und malignen Prozessen unterschied sich signifikant (p < 0,01). Wenn man für bösartige Tumoren einen optimierten Grenzwert von über 5 Gefäßen annimmt, war die Spezifität mit 86 % hoch, die Sensitivität mit 55 % dagegen nicht optimal. Bei der Wahl engerer Abstufungen (ein bis zwei, drei bis fünf, sechs bis zehn und über zehn Gefäße) ergab sich bei der relativ geringen Fallzahl kein signifikantes Ergebnis.

Der *Notch*, eine frühdiastolische Einkerbung im Verlauf der Blutflusskurve, zeigt sich klassischerweise bei Gefäßen mit einer starken Muskelschicht, wie etwa der A. uterina. Demzufolge deutet ein Fehlen des Notch auf eine geringe Zahl glatter Muskeln in der Arterie, wie dies bei neoplastischen Gefäßen der Fall ist. Ein Effekt, der sich auch in dieser Studie zeigt: Nur einer von 17 malignen Tumoren wies einen Notch auf (Sensitivität 94 %). Dem stehen allerdings auch nur 8 von 30 benignen Tumoren mit Notch gegenüber, so dass das Ergebnis mit einer Spezifität von 26 % nicht signifikant war (Abb. 21).

Bei den *quantitativen Blutflussparametern* wiesen die malignen Tumoren höhere Flussgeschwindigkeiten (V_{max} , V_{min} , V_{mean}) und niedrigere Gefäßwiderstände (RI, PI, AB-Ratio) als die benignen Tumoren auf.

Der beste Schwellenwert der einzelnen Blutflussparameter wurde anhand von ROC-Kurven (Abb. 22) ermittelt. Aus ihm wurde das Signifikanzniveau berechnet. Bei der Messung mehrerer Tumorgefäße wurden sowohl der Mittelwert als auch Median, Maximal- und Minimalwert getrennt analysiert, dadurch konnte jedoch im Allgemeinen keine Verbesserung der Aussagekraft erzielt werden. Ein Vergleich der ROC-Kurven für Mittelwert, Median, Höchstwert und Minimalwert am Beispiel des PI aller Gefäße zeigte nur geringe Abweichungen der einzelnen Kurven voneinander (Abb. 23).



Abb. 21: Nachweis eines Notch (spätsystolische Einkerbung) bei Gefäßen benigner und maligner Ovarialtumoren.

Die Aufteilung in periphere und zentrale Gefäße konnte zwar zum Teil zur Verbesserung der ROC-Kurven führen, bei der dadurch verminderten Fallzahl waren die entsprechenden Grenzwerte jedoch häufig nicht mehr signifikant (Abb. 24, Tab. 10). Die zentrale Vaskularisation zeigte sich dabei deutlich aussagekräftiger als die periphere Durchblutung. Bei letzterer streuten die Werte zum Teil sehr stark, so dass nur für zwei Parameter signifikante Grenzwerte errechnet werden konnten.

Als der anfälligste aller quantitativen Parameter erwies sich die *systolische Maximalge-schwindigkeit* (Abb. 22, Abb. 25, Tab. 10). Nur der jeweils höchste Wert (Maximum) mehrerer Gefäße zeigte bei der Gesamtvaskularisation und bei der zentralen Durchblutung signifikante Zahlen (p < 0.05). Der Grenzwert läge hier jeweils bei 0.2 m/s, oberhalb dieses Wertes ist die Wahrscheinlichkeit eines malignen Prozesses hoch.

Aussagekräftigere Werte wies die *maximale diastolische Flussgeschwindigkeit* V_{min} auf (**Tab. 10**). Sie erreichte bei der Gesamtdurchblutung immer signifikante Grenzwerte

(p < 0,05 bzw. p < 0,01), bei der zentralen Vaskularisation mit Ausnahme des niedrigsten Durchblutungswertes (Minimum). Die Signifikanzgrade liegen für den Mittelwert bei p < 0,05, für Median und Maximum bei p < 0,01. Als Grenzwert wurde für den Mittelwert der Gesamtvaskularisation 0,06 m/s errechnet (**Abb. 26**), bei der zentralen Vaskularisation 0,07 m/s Bei der peripheren Vaskularisation erreichte nur der Minimalwert signifikante Werte, der beste Cut-off-Wert lag bei 0,03 m/s (p < 0,05). Jeweils oberhalb dieser Werte besteht eine hohe Probabilität für einen malignen Tumor.

Auch die *mittlere Maximalgeschwindigkeit* V_{mean} erreichte bei der Gesamtdurchblutung durchweg signifikante Werte (p < 0,05 bzw. p < 0,01), der beste Grenzwert des Mittelwertes lag bei 0,13 m/s. Bei der zentralen Durchblutung konnten für den Mittelwert bei 0,09 m/s und den Maximalwert bei 0,14 m/s signifikante Cut-off-Werte gefunden werden (p < 0,05). Werte oberhalb dieser Grenzen sprechen für Malignität. Bei der peripheren Durchblutung ließ sich kein signifikanter Wert bestimmen.

Der *Pulsatilitäts-Index* PI der Gesamtvaskularisation wies immer signifikante Werte auf (p < 0,05 bzw. p < 0,01). Ein Cut-off des Mittelwertes der Gesamtdurchblutung lag bei 1,27, ein weiterer signifikanter Wert ergab sich für einen PI gesamt von 1,1 (**Abb. 27**). Bei der zentralen Durchblutung dagegen zeigte nur der kleinste Wert mehrerer Gefäße (Minimalwert) eine signifikante Grenze bei 0,43 (p < 0,05). Werte unterhalb dieser Grenze sprechen für Malignität.

Beim *Resistance-Index* RI wiesen sowohl Mittelwert als auch Median, Maximalwert und Minimum der Gesamtvaskularisation signifikante Grenzwerte auf (p < 0,05 bzw. p < 0,01, **Tab. 10**). Für den Mittelwert wurde anhand der ROC-Kurven eine Grenze von 0,63 errechnet (**Abb. 28**), Werte darunter sprechen für die Malignität eines Prozesses. Bei zentraler und peripherer Vaskularisation errechnete sich nur für das Minimum ein signifikanter Wert. Die Grenzen wurden mit Hilfe der ROC-Kurven bei 0,37 bzw. 0,66 gesetzt (p < 0,05).

Von der *AB-Ratio* (V_{max} dividiert durch V_{min}) der Gesamtdurchblutung konnten ebenfalls immer signifikante Grenzwerte errechnet werden (p jeweils < 0,05), Cut-off des Mittelwertes lag bei 3,33 m/s. Während bei der zentralen Vaskularisation nur das Minimum mehrerer Gefäße einen signifikanten Grenzwert bei 1,6 m/s aufwies (p < 0,05), ergab keiner der peripheren Werte eine signifikante Grenze. Die quantitativen Parameter zeigten sich, isoliert betrachtet, weniger aussagekräftig als die qualitativen Parameter. Der RI aller Gefäße als bester qualitativer Parameter der Gesamt-vaskularisation (Signifikanzgrad p < 0,01) konnte mit 80 % Sensitivität und 67 % Spezifität nur eine Validität von 72 % erzielen. Eine höhere Validität erreicht der Mittelwert von V_{min} der zentralen Gefäße mit 81 % (Sensitivität 79 %, Spezifität 100 %), bei der geringeren Fallzahl von nur 22 weist er jedoch einen schlechteren Signifikanzgrad (p < 0,05) auf.



Abb. 22: ROC-Diagramm für den Vergleich der Sensitivitäten und Spezifitäten von V_{max} , V_{min} , V_{mean} , PI, RI und AB bei der Dignitätsbeurteilung von Adnex-Tumoren. Mittelwerte aus allen Gefäßen (= Gesamtvaskularisation) der einzelnen Patientinnen.



SENSITIVITÄT

Abb. 23: ROC-Diagramm für den Vergleich der Sensitivitäten und Spezifitäten verschiedener Werte des **Pulsatilitäts-Index (PI)** bei der Dignitätsbeurteilung von Adnex-Tumoren. MW = **Mittelwert** aller Gefäße einer Patientin, **Max** = höchster Wert, **Min** = niedrigster Wert für den Pulsatilitäts-Index aller (zentraler und peripherer) Gefäße einer Patientin.



SPEZIFITÄT

Abb. 24: ROC-Diagramm für den Vergleich der Sensitivitäten und Spezifitäten des Mittelwertes für den **Resistance-Index (RI) zentraler, peripherer und aller Gefäße** (gesamt) bei der Dignitätsbeurteilung von Adnex-Tumoren. MW = Mittelwert verschiedener Gefäße einer Patientin.
Parameter	Signifikanz	ifikanz Malignitätskrit.		Sensitivität	Spezifität	Treffsicher
	grad p	Grenzwerte		[%]	[%]	heit [%]

Qualitative Parameter

Vaskularisation	n.s.	vorhanden	63	90	27	51
Gefäßzahl	< 0,01	>5	49	55	86	73
Zentrale Vask.	< 0,00001	vorhanden	49	95	90	92
"Notch"	n.s.	keiner	40	94	26	55

Gefäße gesamt (Mittelwert)

Vmax	n.s.	> 0,6 m/s	50	5	80	50
Vmin	< 0,01	> 0,06 m/s	50	80	60	68
Vmean	< 0,05	> 0,13 m/s	50	60	73	68
Vmean	< 0,05	> 0,09	50	80	53	64
PI	< 0,01	< 1,27	50	80	63	70
PI	< 0,01	< 1,1	50	95	47	66
RI	< 0,01	< 0,63	50	80	67	72
AB	< 0,05	< 3,33	50	75	63	68

Zentrale Gefäße (Mittelwert)

Vmax	n.s.	> 0,2	22	68	100	73
Vmin	< 0,05	> 0,07	22	79	100	82
Vmean	< 0,05	> 0,09	22	74	100	77
PI	n.s.	< 0,59	22	84	67	82
RI	n.s.	< 0,4	22	89	67	86
AB	n.s.	< 1,72	22	89	67	86

Periphere Gefäße (Mittelwert)

Vmax	n.s.	> 0,24	35	29	46	43
Vmin	n.s.	> 0,03	35	100	32	46
Vmean	n.s.	> 0,34	35	14	96	80
PI	n.s.	< 2,08	35	71	64	66
PI	n.s.	< 1,42	35	86	50	57
PI	n.s.	< 1,13	35	100	32	46
RI	n.s.	< 0,69	35	86	54	60
AB	n.s.	< 3,3	35	100	33	49

Tab. 10: Spezifitäten, Sensitivitäten und Treffsicherheiten der dopplersonographischen Parameter (qualitative Kriterien und Mittelwerte der quantitativen Kriterien) bei der Dignitätsbeurteilung von Adnex-Tumoren. Neben der Fallzahl n (maximal 63) Angabe des jeweiligen Malignitätskriteriums (Malignitätskrit.) bzw. des entsprechenden Grenzwertes. n.s. = nicht signifikant.



Abb. 25: Verteilung der systolischen Maximalgeschwindigkeit (V_{max}) bei benignen und malignen Adnex-Tumoren. Δ = einzelne Mittelwerte der V_{max} aller Gefäße einer Raumforderung.



Abb. 26: Verteilung der diastolischen Maximalgeschwindigkeit (V_{min}) bei benignen und malignen Adnex-Tumoren. Δ = einzelne Mittelwerte der V_{min} aller Gefäße einer Raumforderung.



Abb. 27: Verteilung des Pulsatilitäts-Index (PI gesamt) bei benignen und malignen Adnex-Tumoren. Δ = einzelne Mittelwerte des PI aller Gefäße einer Raumforderung.



Abb. 28: Verteilung des Resistance-Index (RI gesamt) bei benignen und malignen Adnex-Tumoren. Δ = einzelne Mittelwerte des RI aller Gefäße einer Raumforderung.

Um die Unabhängigkeit der dopplersonographischen Einzelparameter bei der Dignitätsbeurteilung von Adnex-Tumoren zu überprüfen, wurden die dopplersonographischen Werte ebenfalls in eine multivariate Untersuchung einbezogen. Mit Hilfe der multiplen logistischen Regression wurde zudem die beste Kombination berechnet. Wie bei den sonomorphologischen Untersuchungen soll eine Entscheidungsmatrix zu einer optimalen Verknüpfung der Untersuchungsergebnisse führen: Beträgt die Wahrscheinlichkeit über 50 %, wird ein maligner Tumor angenommen, liegt sie darunter, ein gutartiger.

Nach Kombination in der multivariaten Analyse erwies sich der Doppler als aussagekräftigeres Entscheidungsmittel als die Sonographie, allerdings ausschließlich mit der Analyse der qualitativen Parameter (Vaskularisation / zentrale Vaskularisation vorhanden). Durch die quantitativen Parameter (RI, PI, etc.) konnte kein Vorteil hinsichtlich der Dignitätsbeurteilung erzielt werden. Die Sensitivität ergab mit dem Entscheidungsbaum 86 %, die Spezifität 93 %, was insgesamt einer Validität von 90 % entspricht (**Tab. 11**, **Abb. 29**). Drei maligne Tumoren wurden mit diesem Modell als falsch benigne eingestuft, histologisch zwei seröse Adenokarzinome im FIGO-Stadium IIIc und ein Tumor niedrig maligner Potenz (FIGO Ia). Ein Adenokarzinom und der Tumor niedrig maligner Potenz waren nicht vaskularisiert, das zweite Adenokarzinom wies keine zentrale Vaskularisation auf. Drei benigne Tumoren wurden fälschlicherweise als maligne bezeichnet, histologisch eine Hämatosalpinx und zwei Inklusionszysten. Alle drei zeigten eine zentrale Vaskularisation.



Abb. 29: Diagnosematrix zur dopplersonographischen Dignitätsbeurteilung von Adnex-Tumoren. Die Prozentzahlen geben die Wahrscheinlichkeit für das Vorliegen von Malignität eines Prozesses bei Erfüllung des jeweils vorgeschalteten Kriteriums an.

4.2.3 KOMBINIERTE DIGNITÄTSBEURTEILUNG

Die sonographischen und dopplersonographischen Kriterien kombiniert wurden anschließend mit Hilfe der multiplen logistischen Regression einer multivariaten Analyse unterzogen (**Tab. 11**) und in eine Entscheidungsmatrix eingefügt.

Für die Verbindung der morphologischen Merkmale mit der Vaskularisation kristallisierten sich zwei Modelle mit identischen Validitätsparametern heraus. Das erste Modell besteht einer Synthese der sonomorphologischen aus und der dopplersonographischen Entscheidungsmatrix (Abb. 30). Während für Tumoren ohne soliden Anteil Benignität angenommen wurde, wurden die Tumoren mit solidem Anteil auf ihre Vaskularisation und anschließend auf die zentrale Vaskularisation geprüft. Während das Kriterium solider Anteil nur eine Validität von 78 % erreicht, konnte diese durch Einbeziehen der Vaskularisation auf 84 % und durch zusätzliches Einbeziehen der zentralen Vaskularisation auf 90 % gesteigert werden (Tab. 11). Mit diesem Modell wurden eben die sechs Tumoren falsch eingestuft, die auch bei der alleinigen dopplersonographischen Bewertung falsch klassifiziert worden waren. Die als falsch benigne eingestuften drei malignen Tumoren waren histologisch zwei seröse Adenokarzinome im FIGO-Stadium IIIc und ein Tumor niedrig maligner Potenz (FIGO Ia). Ein Adenokarzinom wies keinen soliden Anteil und keine Vaskularisation auf (Malignitätsprobabilität 3,5 %). Der Tumor niedrig maligner Potenz war nicht vaskularisiert (Malignitätsprobabilität 32%), das zweite Adenokarzinom wies keine zentrale Vaskularisation auf (Malignitätsprobabilität 2,2 %). Die drei fälschlicherweise als maligne bezeichneten benignen Tumoren waren histologisch eine Hämatosalpinx und zwei Inklusionszysten. Alle drei zeigten neben einem soliden Anteil eine zentrale Vaskularisation (Malignitätswahrscheinlichkeit 88,9 %).

Das zweite Modell bezog statt der zentralen Vaskularisation die qualitativen Charakteristika PI gesamt (Mittelwert) und V_{mean} gesamt (Mittelwert) in die Diagnosefindung ein (Abb. 31). Die Treffsicherheit konnte von 78 % (solider Anteil) und 84 % (solider Anteil plus Vaskularisation) mit dem PI der Gesamtvaskularisation und der mittleren Maximalgeschwindigkeit auf 90 % gesteigert werden. Mit diesem Modell wurden drei maligne Tumoren falsch negativ beurteilt, histologisch zwei Neoplasien niedrig maligner Potenz und ein seröses Adenokarzinom im FIGO-Stadium 3c. Das seröse Adenokarzinom zeigte sonographisch keinen soliden Anteil und keine Vaskularisation (Malignitätswahrscheinlichkeit 3,5 %). Einer der Tumoren niedrig maligner Potenz war nicht vaskularisiert (Malignitätswahrscheinlichkeit 32 %). Der zweite LMP-Tumor war zwar zentral vaskularisiert, die Malignitätswahrscheinlichkeit betrug aber bei einem Mittelwert des PI der Gesamtdurchblutung von 1,53 und einem Mittelwert der mittleren Maximalgeschwindigkeit V_{mean} von 0,06 ebenfalls nur 32 %. Drei benigne Tumoren wurden falsch positiv eingestuft, feingeweblich ein Teratom, eine Inklusionszyste und ein seröses Zystadenom. Die Inklusionszyste und das seröse Zystadenom wiesen neben einem soliden Anteil und einer Vaskularisation zu niedrige Mittelwerte des PI gesamt von 0,58 bzw. 0,45 auf (Malignitätswahrscheinlichkeit 88,9 %). Der PI des Teratoms war mit 4,31 hoch, dagegen war auch der Mittelwert der mittleren Maximalgeschwindigkeit V_{mean} mit 0,55 zu hoch (Malignitätswahrscheinlichkeit 75 %).



Abb. 30: Entscheidungsmatrix für eine kombinierte Anwendung der in der multiplen logistischen Regression ermittelten signifikanten sonographischen und qualitativen dopplersonographischen Parameter bei der Differenzierung unklarer Adnex-Tumoren (Modell I). Die Prozentzahlen geben die Wahrscheinlichkeit für das Vorliegen von Malignität eines Prozesses bei Erfüllung des jeweils vorgeschalteten Kriteriums an.

Die durch die kombinierte sonographische und dopplersonographische Untersuchung erreichte Sensitivität, Spezifität und Validität konnte mit jeweils 86 %, 93 % und 90 % die Ergebnisse der Sonomorphologie signifikant verbessern (p < 0,01). Vor allem die Spezifität ließ sich durch die Kombination steigern (93 % vs. 68 %). Die Treffsicherheit erhöhte sich damit von 78 % auf 90 %. Die Ergebnisse der kombinierten Entscheidungsmatrix waren der Dopplersonographie alleine allerdings nicht überlegen (**Tab. 11**). Sowohl mit der dopplersonographischen Entscheidungsmatrix als auch mit der kombinierten Entscheidungsmatrix wurden drei von 22 malignen und drei von 41 benignen Tumoren falsch bewertet.



Abb. 31: Entscheidungsmatrix für eine kombinierte Anwendung der in der multiplen logistischen Regression ermittelten signifikanten sonographischen sowie qualitativen *und* quantitativen dopplersonographischen Parameter bei der Differenzierung unklarer Adnex-Tumoren (Modell II). Die Prozentzahlen geben die Wahrscheinlichkeit für das Vorliegen von Malignität eines Prozesses bei Erfüllung des jeweils vorgeschalteten Kriteriums an.

SONOGRAPHIE								
Kriterien	riterien solider Anteil							
Sensitivität			95	%				
Spezifität			68	%				
Validität			78	%				
	DOPPLI	ERSO	NOGRAPH	IE				
Kriterien	Vaskula	risati	ion	Vas	kula	risation		
				zen	tral	e Vask.		
Sensitivität	90	%			86	%		
Spezifität	27	%			93	%		
Validität	51	%			90	%		
KOMBINATION S	KOMBINATION SONOGRAPHIE UND DOPPLERSONOGRAPHIE, MODELL I							
Kriterien	solider Ante	il	solider	Anteil	SC	olider Anteil		
			Vaskula	risation	Va	skularisation		
Sonsitivität	05 %		00	0/2	ze	86 %		
Sensitivität	93 /0 68 %		90 80	/0		03 %		
Validität			84	/0 0/0		90 %		
	10 /0	-		/0				
KOMBINATION S	ONOGRAPHIE U	JND	JOPPLERS	ONOGRAPH	HE,	MODELL II		
Kriterien	solider Ant.	sol	ider Ant.	solider An	ıt.	solider Ant.		
			Vask.	Vask.		Vask.		
				PI gesam	lt	PI gesamt		
						V _{mean} ges.		
Sensitivität	95 %		90 %	72 %		86 %		
Spezifität	68 %		80 %	95 %		93 %		
Validität	78 %		84 %	82 %		90 %		

Tab. 11: Vergleich der diagnostischen Wertigkeit der Entscheidungsmatrix für die Sonographie, die Dopplersonographie und die Kombination aus Sonographie mit der Dopplersonographie bei der Dignitätsbeurteilung von Adnex-Tumoren (n = 63). Modell I: nur qualitative dopplersonographische Kriterien, Modell II: qualitative und quantitative dopplersonographische Kriterien. V_{mean} = mittlere Maximalgeschwindigkeit, PI = Pulsatilitätsindex.

4.3 Mammatumoren

Um den Stellenwert der "neuronalen Netze" für die Optimierung einer kombinierten sonographischen und dopplersonographischen Dignitätsbeurteilung maligner Tumoren zu überprüfen, wurde dieses Verfahren bei insgesamt 219 Patientinnen mit Mammatumoren angewandt und mit der statistischen Analyse durch die multiple logistische Regression verglichen. Dabei wurde die Aussagekraft der mit Hilfe der multiplen logistischen Regression anhand von 89 Patientinnen erstellten Entscheidungsmatrix an einer prospektiven Patientengruppe von 130 Patientinnen (94 maligne, 34 benigne Tumoren) überprüft.

4.3.1 KOMBINIERTE SONOGRAPHISCHE UND DOPPLERSONOGRAPHISCHE DIGNITÄTSBEURTEILUNG (MULTIPLE LOGISTISCHE REGRESSION)

Die anhand der retrospektiven Analyse entwickelte Entscheidungsmatrix wurde für die kombinierte sonographische und dopplersonographische Dignitätseinstufung prospektiv anhand der 130 Patientinnen mit einem Mammatumor (94 maligne, 36 benigne) überprüft. Die Tumoren wurden präoperativ ohne Kenntnis der Histologie befundet. In der anschließenden Auswertung wurden die Einstufungen auf ihre Treffsicherheit bezüglich des histologischen Ergebnisses überprüft.

Die Anwendung des Entscheidungsbaumes an den prospektiven Datensätzen zeigt eine deutlich verschlechterte Spezifität von 65 % bei einer nur geringfügig verschlechterten Sensitivität von 95 %. Die Treffsicherheit sank dadurch auf 87 % (**Tab. 12**).

ENTSCHEIDUNGSMATRIX	Sensiti- vität	Spezi- fität	Validitä t	Sensiti- vität	Spezi- fität	Validität
	retrospektiv			prospektiv		
	98 %	89 %	95 %	95 %	65 %	87 %

Tab. 12: Vergleich der mit der Entscheidungsmatrix erreichten Sensitivitäten, Spezifitäten und Validitäten bei retro- und prospektiven Datensätzen der Patientinnen mit Mammatumor (n retrospektiv = 89, n prospektiv = 130).

Von den 36 benignen Tumoren konnten zwei nicht mit dem Entscheidungsbaum beurteilt werden, da sie zwar eine Vaskularisation zeigten, jedoch bei vermutlich für das Gerät zu langsamen Blutflussgeschwindigkeiten keine Dopplerflusskurven ableitbar waren. Von den übrigen 34 Tumoren wurden zwölf falsch eingestuft, so dass die Spezifität 65 % beträgt. Nur fünf von 94 Karzinomen waren mit dem Entscheidungsbaum fälschlicherweise als "benigne" beurteilt worden, woraus sich eine Sensitivität von 95 % errechnet. Insgesamt wurden damit 111 von 128 Tumoren richtig diagnostiziert.

Die fünf falsch eingestuften malignen Tumoren waren histologisch ein Carcinoma in situ und vier invasiv duktale Karzinome, davon drei im Stadium T₁ und eines im Stadium T₄. Alle Tumoren einschließlich des einen retrospektiv falsch eingeordneten waren zwar vaskularisiert, waren aber durch eine scharfe Begrenzung, eine fehlende dorsale Schallauslöschung und zu niedrige Werte des PI oder der systolischen Maximalgeschwindigkeit V_{max} gekennzeichnet. In drei Fällen erwies sich V_{max} als zu niedrig, einmal der Pulsatilitäts-Index und zweimal sowohl V_{max} als auch PI. Häufig legten sonographische Kriterien, unabhängig vom Entscheidungsbaum, einen starken Verdacht auf ein Karzinom nahe, wie z.B. im Fall des T₄-Karzinoms. Die Infiltration der Haut und damit die hohe Wahrscheinlichkeit für Malignität war dort auch sonographisch festgestellt worden, da hier jedoch die Bewertung alleine anhand des Baumes geprüft werden soll, fällt dieser Tumor trotzdem in die Kategorie der als falsch negativ eingeschätzten Befunde.

Unter den 15 benignen Tumoren, die falsch erkannt worden waren (3 retrospektiv, 12 prospektiv) befanden sich sechs mit "Vaskularisation vorhanden" in Kombination mit "unscharfe Begrenzung" (errechnete Wahrscheinlichkeit für Malignität 100 %), fünf mit "keine Vaskularisation" in Kombination mit "unscharfe Begrenzung" (errechnete Wahrscheinlichkeit für Malignität 50 %) und vier mit vorhandener Vaskularisation, scharfer Begrenzung und zu hohen Pulsatilitätsindizes und systolischen Maximalgeschwindigkeiten.

Die Treffsicherheit zeigt sich damit in der prospektiven Gruppe nur leicht erhöht gegenüber den Validitäten der Einzelparameter. So erreicht etwa in der Dopplersonographie der Mittelwert des PI aller Gefäße 85 % Validität. Das sonomorphologische Kriterium "Septen" erzielt sogar 86 % Validität, die "Randkontur" 84 %. Allerdings entziehen sich hier einige Tumoren der Beurteilbarkeit: Dopplersonographisch waren 18 Tumoren nicht vaskularisiert, und zweimal war das Ableiten der Dopplerflusskurve bei vaskularisierten Tumoren nicht möglich, so dass hier kein PI bestimmt werden konnte. Bei 46 Tumoren konnte nicht beurteilt werden, ob ein Septenabbruch oder eine Verdrängung vorlag, was dieses Einzelkriterium für die Beurteilung aller Tumoren nicht mehr in Frage kommen lässt. Als bester Einzelparameter bleibt somit in der prospektiven Gruppe die Randkontur, die bei allen 130 prospektiven Tumoren beurteilbar gewesen war und die mit 84 % eine nur um drei Prozentpunkte schlechtere Validität aufweist als der Entscheidungsbaum.

4.3.2 Kombinierte sonographische und dopplersonographische Dignitätsbeurteilung (Neuronale Netze)

In die Auswertung mit neuronalen Netzen gingen als Datensätze die 217 Mammatumoren mit sonomorphologischem Korrelat ein. Die Informationsinhalte der einzelnen Parameter wurden aufgesplittet, also z.B. die Information "Vaskularisation vorhanden" auf ein Eingangsneuron und die Information "Vaskularisation nicht vorhanden" auf ein anderes Eingangsneuron gelegt. Auf diese Weise entstanden insgesamt 142 Eingänge (Eingansneuronen), davon 30 aus den sonomorphologischen, zwei aus allgemeinen Kriterien (Menopausenstatus). Der Rest setzte sich aus dopplersonographischen Parametern zusammen: Zehn aus allgemeinen Gefäßparametern (Vaskularisation vorhanden oder nicht vorhanden, Gefäßanzahl, Gefäßverteilung) und 100 aus den Dopplerflusskurven, darunter jeweils 25 Werte aus der gesamten, zentralen, und peripheren Vaskularisation und 25 aus den Vaskularisationsdaten der Kontrollseite.

Das Netz wurde in den meisten Versuchen anhand der als **retrospektiv** klassifizierten Datensätze trainiert (87 "Lerndatensätze"). Beim Einbeziehen nur der *sonomorphologischen Kriterien* (30 Eingangsdaten) konnte das Netz dabei nicht völlig korrekt trainiert werden (**Tab. 13**). Als Endergebnis ordnete es drei maligne und einen benignen Tumor dem jeweils falschen Kollektiv zu. Die Sensitivität liegt damit bei 95 %, die Spezifität bei 96 % und die Validität bei 95 %. Mit dem Lernen nur anhand der *dopplersonographischen Werte* (110 Eingangsdaten) wurde ein maligner Tumor fälschlicherweise als benigne eingeschätzt (Sensitivität 98 %), die benignen Tumoren konnten alle korrekt trainiert werden (Spezifität 100 %, Validität 99 %). Bei der der *Kombination aller sonographischen mit allen dopplersonographischen Kriterien*, also dem Einbeziehen aller Parameter (142 Eingänge), lernte das Netz alle Datensätze richtig (Sensitivität, Spezifität und Validität 100 %). Anschließend wurden Eingänge mit niedrigen Gewichten, also die nur wenig zum Netzausgang beitragenden Kriterien, sukzessive eliminiert. Erst unterhalb einer Zahl von 13 Eingängen lernte das Netz nicht mehr alle retrospektiven Datensätze richtig. Durch Eingabe alleine der Parameter, aus denen sich die in der multiplen logistischen Regression erstellte *Entscheidungsmatrix* (Baum) zusammensetzt (7 Eingänge), wurden ein maligner und zwei benigne Tumoren falsch trainiert, so dass sich eine Validität von 96 % ergibt (Sensitivität 98 %, Spezifität 93 %). Bei der Kombination der sonomorphologischen Parameter mit den allgemeinen, qualitativen Gefäßparametern (40 Eingänge) lernte das Netz fehlerfrei alle Datensätze (Sensitivität 100 %, Spezifität 100 %, Validität 100 %), ebenso durch zusätzliches Einbeziehen von RI und PI (72 Eingänge) sowie dem V_{max} (88 Eingänge, **Tab. 13**).

Lerndaten- sätze	Eingangs- parameter	Ein- gänge	Lern- daten	Sensi- tivität	Spezi- fität	Vali- dität	Test- daten	Sensi- tivität	Spezi- fität	Vali- dität
		(11)	(11)	[/•]		[/•]	(11)	[/•]	[/ V] Toston	[/ •]
					Leinen				1051011	
Retrosp.	Nur Sono	30	89	95	96	95	130	83	44	72
Retrosp.	Nur Doppler	110	89	98	100	99	130	80	64	75
Retrosp.	Alle	142	89	100	100	100	130	97	58	82
Retrosp.	Baum	7	89	98	93	97	130	90	69	85
Retrosp.	Sono + VP	40	89	100	100	100	130	88	56	79
Retrosp.	Sono + VP + RI + PI	72	89	100	100	100	130	96	61	86
Retrosp.	Sono + VP + RI + PI + V _{max}	88	89	100	100	100	130	96	67	88
Gemischt retro u. pro (ungerade)	Alle	142	109	100	100	100	108	94	82	90
Gemischt retro u. pro (gerade)	Alle	142	108	100	100	100	109	87	88	87

Tab. 13: Mittels neuronaler Netze erzielte Sensitivitäten, Spezifitäten und Validitäten bei der Auswertung der sonographischen und dopplersonographischen Dignitätsbeurteilung von Mammatumoren. VP = qualitative Vaskularisationsparameter. Retro = retrospektive Patientengruppe (n = 89), pro = prospektive Patientengruppe (n = 130).

Das Training an gemischt retro- und prospektiven Datensätzen unter Einbeziehen aller Parameter (142 Eingänge, 108 bzw. 109 Lerndatensätze) meisterte das Netz ebenfalls fehlerlos (Sensitivität, Spezifität und Treffsicherheit 100 %, **Tab. 13**). Nach dem Lernen wurde das trainierte Netz anhand der Daten der **prospektiven** Patientengruppe getestet (130 "Testdatensätze", **Tab. 13**). In den zwei Versuchen mit gemischt retro- und prospektiven Patientinnen diente jeweils der Hälfte der Patientendaten, die nicht in die Trainingsrunde mit einbezogen worden waren, als Testdaten (109 bzw. 108 Testdatensätze).

Mit den *sonomorphologischen Kriterien* wurden beim Testen 16 maligne und 20 benigne Tumoren falsch eingeschätzt (Sensitivität 83 %, Spezifität 44 %, Validität 72 %). Beim Testen der *dopplersonographischen Kriterien* wurden 19 maligne und 13 benigne Tumoren vom Netz falsch eingestuft, was einer Sensitivität von 80 %, einer Spezifität von 64 % und einer Validität von 75 % entspricht. Beim Testen der *Kombination aller sonographischen mit allen dopplersonographischen Kriterien* wurden drei maligne und 15 benigne Tumoren falsch eingeschätzt (Sensitivität 97 %, Spezifität 58 %, Validität 86 %). Die drei falsch eingestuften malignen Tumoren waren histologisch ein Carcinoma in situ und zwei invasiv duktale Karzinome im Stadium T₁. Bei den falsch benignen Tumoren handelte es sich in erster Linie um mastopathische Herde (n = 9). Das Testen der in den *Entscheidungsbaum* eingeflossenen Kriterien erzielte ebenfalls ein gutes Ergebnis: Neun maligne und elf benigne Tumoren wurden falsch eingeschätzt, die Validität liegt somit bei 85 % (Sensitivität 90 %, Spezifität 69 %).

Das Hinzufügen der allgemeinen, qualitativen Gefäßparameter konnte die Treffsicherheit der sonomorphologischen Parameter alleine deutlich steigern: Nur noch elf maligne und 16 benigne Tumoren wurden so falsch erkannt (Validität 79 %, Sensitivität 88 %, Spezifität 56 %). Das zusätzliche Einbeziehen von RI und PI konnte sie auf 86 % (Sensitivität 96 %, Spezifität 61 %) erhöhen, das zusätzliche Hinzufügen der aus der systolischen Maximalgeschwindigkeit V_{max} abgeleiteten Eingangsdaten sogar auf 88 %. Bei der letztgenannten Kombinationsmöglichkeit wurden nur 4 maligne und 12 benigne Tumoren vom Netz falsch beurteilt (Sensitivität 96 %, Spezifität 67 %).

Sehr gute Resultate erzielten die Testdurchläufe an den aus retrospektiven und prospektiven Patientinnen gemischten Datensätzen: Der Test an den mit geraden Nummern gekennzeichneten Patientinnen nach Training an den mit ungeraden Zahlen gekennzeichneten Patientendaten lieferte eine Sensitivität von 94 % und eine Spezifität von 82 %, die Validität beträgt damit 90 %. Nur fünf maligne und sechs benigne von insgesamt 108 Tumoren

wurden dem jeweils falschen Kollektiv zugeordnet. Umgekehrt ergab sich immerhin eine Validität von 87 % mit zehn falsch erkannten malignen und vier falsch erkannten benignen Tumoren von insgesamt 109 Tumoren (Sensitivität 87 %, Spezifität 88 %). Allerdings trägt bei diesen Versuchen auch die relativ höhere Zahl an Trainingsdaten und die relativ geringere Zahl an Testdaten zu der guten Treffsicherheit bei.

Insgesamt zeigt sich, dass die Treffsicherheit der Testdatensätze (Validität 72 bis 90 %) deutlich unter den sehr guten Ergebnissen der Trainingsdurchläufe (Validität zwischen 95 % und 100 %) lag. Vor allem eine schlechte Beurteilbarkeit der benignen Tumoren mit nicht ganz optimalen Spezifitätswerten fiel hier ins Gewicht. Diese mag vor allem durch eine insgesamt zu geringe Zahl benigner Tumoren bedingt sein.

4.3.3 VERGLEICH DER NEURONALEN NETZE MIT STATISTISCHEN VERFAHREN

Beim Vergleich der computergestützten Auswertung mittels neuronaler Netze mit der statistischen Auswertung zeigt sich eine deutliche Überlegenheit der neuronalen Netze bei der retrospektiven Analyse (bei den Netzen auch als "Lernphase" bezeichnet). Bei Kombination allein der *sonomorphologischen Kriterien* ordnete das Netz drei maligne und einen benignen Tumor dem jeweils falschen Kollektiv zu. Die Sensitivität lagt damit bei 95 %, die Spezifität bei 96 % und die Validität bei 95 %. Die Entscheidungsmatrix, die mittels multipler logistischer Regression alleine anhand sonomorphologischer Kriterien erstellt worden war, stufte sieben maligne und einen benignen Tumor falsch ein (Validität 91 %, Sensitivität 88 %, Spezifität 96 %), also doppelt so viele wie die neuronalen Netze. Sowohl die neuronalen Netze als auch die multiple logistische Regression sind dem besten Einzelkriterium (Begrenzung: 81 % Sensitivität, 96 % Spezifität, 86 % Validität) deutlich überlegen.

Mit dem Lernen nur anhand der *dopplersonographischen Werte* konnte die Treffsicherheit der neuronalen Netze weiter gesteigert werden, nur ein maligner Tumor wurde fälschlicherweise als benigne klassifiziert (Sensitivität 98 %, Spezifität 100 %, Validität 99 %). Ein Vergleich mit der multiplen logistischen Regression kann nicht gezogen werden, da keine Entscheidungsmatrix alleine anhand dopplersonographischen Parametern erstellt worden war. Der beste dopplersonographische Einzelparameter, der Mittelwert des PI aller Gefäße, erreichte eine Treffsicherheit von 82 % (Sensitivität 84 %, Spezifität 70 %) und ist damit der Kombination aller dopplersonographischer Kriterien in den neuronalen Netzen deutlich unterlegen. Bei der *Kombination aller sonographischer mit allen doppler-sonographischen Kriterien* lernte das Netz alle Datensätze richtig (Sensitivität, Spezifität und Validität 100 %). Damit ist sie der aus dopplersonographischen und sonomorphologischen Daten erstellten Entscheidungsmatrix überlegen, die eine Validität von 97 % erreichte (Sensitivität 98 %, Spezifität 93 %). Durch Eingabe alleine der Parameter, aus denen sich die in der multiplen logistischen Regression erstellte *Entscheidungsmatrix* zusammensetzt (Vaskularisation, Begrenzung, Mittelwert PI gesamt, Mittelwert V_{max} gesamt, dorsale Schallauslöschung), wurden ein maligner und zwei benigne Tumoren falsch trainiert, so dass sich - etwas besser als beim Entscheidungsbaum selbst - eine Validität von 96 % ergab (Sensitivität 98 %, Spezifität 93 %).

	Retrospekti	ve Analyse	Prospektive Analyse		
	MLR	NN	MLR	NN	
Sonomorphologie	91 %	95 %	79 %	72 %	
Dopplersonographie	-	99 %	-	75 %	
Kombination Sono und	97 %	100 %	87 %	82 %	
Doppler					
Beste erreichbare Treffsi-	97 %	100 %	87 %	88 %	
cherheit					

Tab. 14: Vergleich der Treffsicherheiten bei der Auswertung sonographischer und dopplersonographischer Untersuchungen von Mammatumoren (n = 219) mittels multipler logistischer Regression (MLR) und neuronaler Netze (NN).

In der Testphase lagen die Validitäten - ähnlich der mittels multipler logistischer Regression erzeugten Ergebnisse - unter den bei den Lerndurchgängen erzielten Werten, wobei ihre Höhe in etwa den mittels des Entscheidungsbaumes erzeugten Resultaten entspricht bzw. ihnen zum Teil leicht überlegen ist.

Anhand alleine der *sonomorphologischen Kriterien* wurden beim Testen mit neuronalen Netzen 16 maligne und 20 benigne Tumoren falsch eingeschätzt (Sensitivität 83 %, Spezifität 44 %, Validität 72 %). Mit der Entscheidungsmatrix wurden 13 Karzinome und 14 benigne Befunde falsch eingestuft, so dass die multiple logistische Regression mit einer Validität von 79 % (Sensitivität 84 %, Spezifität 67 %) hier den neuronalen Netzen leicht überlegen war. Etwas bessere Vorhersagewerte erreichten die neuronalen Netze durch Eingabe alleine der *dopplersonographischen Kriterien*: 19 maligne und 13 benigne Tumoren stufte das Netz falsch ein, was einer Sensitivität von 80 %, einer Spezifität von 64 % und einer Validität von 75 % entspricht.

Durch Kombination der sonomorphologischen mit den dopplersonographischen Parametern wurde mittels multipler logistischer Regression eine Sensitivität von 95 %, eine Spezifität von 65 % und eine Treffsicherheit von 87 % erzielt. Fünf maligne und zwölf benigne Tumoren wurden damit falsch beurteilt, insgesamt wurden damit 17 von 128 Tumoren falsch eingestuft. Zwei benigne Tumoren konnten anhand der Entscheidungsmatrix nicht beurteilt werden, da sie zwar vaskularisiert waren, aufgrund zu niedriger Flussgeschwindigkeiten aber keine Dopplerflusskurve abgeleitet werden konnte. Die neuronalen Netze erreichten durch Eingabe aller sonomorphologischer und dopplersonographischer Parameter (142 Eingänge) eine Treffsicherheit von 82 % (Sensitivität 97 %, Spezifität 58%). Drei maligne und 21 benigne Tumore stufte das Netz damit falsch ein. Durch Eingabe nur der Parameter, die auch für die Entscheidungsmatrix verwendet wurden, steigerte sich die Treffsicherheit auf 85 % (Sensitivität 90 %, Spezifität 69 %). Drei bzw. neun maligne und 21 bzw. elf benigne Tumoren stufte das Netz damit falsch ein. Die Treffsicherheit liegt in beiden Fällen leicht unter der bei der multivariaten Analyse. Allerdings konnten durch die neuronalen Netze auch die zwei benignen Tumoren beurteilt werden, die bei der Entscheidungsmatrix ausgeschlossen werden mussten.

Durch das Verwenden anderer Eingangsparameter konnte die Wertigkeit der Analyse mit neuronalen Netzen aber noch deutlich gesteigert werden: Durch die Kombination der sonomorphologischen Parameter mit den qualitativen Doppler-Parametern sowie den aussagekräftigsten Widerstandswerten, dem RI und dem PI, erreichte die Validität 86 %. Durch das zusätzliche Einbeziehen des Mittelwertes der systolischen Maximalgeschwindigkeit V_{max} erreichten die neuronalen Netze eine Treffsicherheit von 88 % (Sensitivität 96 %, Spezifität 67 %). Nur vier maligne und zwölf benigne Tumoren (von 130) wurden so falsch eingestuft, die Treffsicherheit ist damit der multiplen logistischen Regression überlegen.

5 DISKUSSION

Die transvaginale Sonographie und die Dopplersonographie stellen in der Diagnostik gynäkologischer Tumoren einen festen Bestandteil dar. Durch eine Kombination beider Methoden mit Hilfe der multiplen logistischen Regression oder der neuronalen Netze, einem Verfahren der künstlichen Intelligenz, kann die diagnostische Wertigkeit der Verfahren anhand standardisierter Parameter deutlich gesteigert werden. In der vorliegenden Arbeit konnte die Treffsicherheit des besten sonographischen Einzelkriteriums bei unklaren Adnex-Tumoren durch Kombination mit der Dopplersonographie in einer diagnostischen Entscheidungsmatrix von 78 % auf 90 % gesteigert werden. Mit dem ausschließlichen Einsatz der B-Mode-Sonographie konnten nach multivariater Analyse der Parameter und Kombination in einer Entscheidungsmatrix 91 % aller unklaren Mammatumoren bezüglich ihrer Dignität richtig eingestuft werden. Die Kombination mit der farbkodierten und gepulsten Dopplersonographie brachte noch einen Zugewinn von etwa 5 %.

Im Gegensatz zu Schnittbildverfahren mit "festen" Bildern wie etwa der Kernspintomographie, die ohne Informationsverlust auch retrospektiv beurteilt werden können, ist die Sonographie untersucher- und erfahrungsabhängig [110] (S. 564). Die Abmessung der Dimensionen, die detaillierte Beurteilung von Schallphänomenen wie etwa der Begrenzung solider Tumoren oder der Infiltration in Nachbarorgane setzen ein hohes Maß an Erfahrung voraus. Noch stärker von solchen Einflussfaktoren abhängig ist die dopplersonographische Beurteilung, deren Informationsgehalt ganz wesentlich auf der richtigen Anwendung der Methode basiert, angefangen mit der korrekten Einstellung des Gefäßwinkels [44] und dem Erzeugen einer sauberen Blutflusskurve.

5.1 Reproduzierbarkeit dopplersonographischer Untersuchungen

Die Güte sonographischer und dopplersonographischer Messungen ist im wesentlichen drei Einflussfaktoren ausgesetzt: Dem Untersuchungsobjekt mit seinen unterschiedlichen Streuungseigenschaften und hormon- und altersabhängigen Schwankungen in Durchblutung und Morphologie, der Qualität des Dopplergerätes und der Erfahrung des Untersuchenden mit seiner Gerätekenntnis [66] (S. 35 - 36).

Der letzte Aspekt hat dabei zentrale Bedeutung. Bei der semiquantitativen Flussdiagnostik verändern insbesondere Fehler bei der Winkeleinstellung und ein nicht gefäßdeckend platziertes Dopplerfenster die absoluten Geschwindigkeiten. Im Farbmodus kann eine nicht sensibel genug gewählte Geräteeinstellung unter anderem gravierende Fehler in der Gefäßdiagnostik bewirken.

Durch das Gerät bedingte Fehlmessungen spielen nur eine marginale Rolle, da sich die Gerätehersteller durch zahlreiche Techniken bemühen, diese möglichst gering zu halten. Bei "gerätebedingten" Fehlern spielen häufig auch physikalische Effekte wie die schlechte Signalableitung bei zu flachen Winkeln eine entscheidende Rolle.

Beim Überprüfen der *Genauigkeit der dopplersonographischen Geschwindigkeitsmessung* mittels **gepulster Dopplersonographie** bei verschiedenen Pumpvolumina zeigte sich eine gute Korrelation zwischen den dopplersonographisch gemessenen Geschwindigkeiten und den mit der Infusiomatenpumpe geförderten Flussmengen bei Pumpvolumina über 50 ml/h, entsprechend 0,035 m/s Maximalgeschwindigkeit in der Flussmitte. Die Werte über- oder unterschritten bei diesen Flussvolumina das 90 %-Konfidenzintervall praktisch nicht. Die beste Näherung ergab der Mittelwert aus mehreren Messungen. Dies entspricht den Ergebnissen anderer In-Vitro-Versuche mit deutlich höheren Flussgeschwindigkeiten (größer 100 ml/min), die eine exzellente Korrelation zwischen dopplersonographisch gemessenem und errechnetem Fluss berichten [64].

Unterhalb von 50 ml/h (0,035 m/s errechnete Maximalgeschwindigkeit) wurden die Werte konstant überschätzt. Der Mittelwert aus mehreren Messungen überschreitet hier leicht das 90 %-Konfidenzintervall. Bei 5 ml/h wurde kein Fluss mehr gemessen und mit Ausnahme eines Versuches kein Signal mehr abgeleitet. Dies beruht auch auf der Gerätefunktion, die nur zwei Stellen hinter dem Komma darstellt: Da sich bei 5 ml/h eine Maximalgeschwindigkeit von 0,0035 m/s errechnet, wird dies vom Gerät auf 0,00 m/s abgerundet.

Als Ursache der leichten Abweichungen bei schnelleren Geschwindigkeiten können Winkel- oder Rundungsfehler angenommen werden. Langsame Geschwindigkeiten sind insofern um einiges anfälliger, da Rundungsfehler viel mehr ins Gewicht fallen [94] (S. 52). Außerdem wirkt sich der Einsatz von Wandbewegungsfiltern deutlicher aus; die konstante Überschätzung der Messergebnisse wird in der Literatur vor allem dieser Tatsache zugeschrieben [94] (S. 52). Die Reproduzierbarkeit der Messergebnisse kann als gut angesehen werden. Die Werte bei den an verschiedenen Tagen durchgeführten acht Messreihen unterschieden sich nicht deutlich.

Die dopplersonographische Geschwindigkeitsmessung bei verschiedenen Winkeln wurde unter Simulation von In-Vivo-Bedingungen durchgeführt, der Schallkopf also von Hand gehalten. Damit sollten leichte Fehler von bis zu 5 Grad Winkelabweichung, die in der Realität nicht zu vermeiden sind, simuliert werden. Die rechnerische Abweichung der Werte bei Fehlmessungen von maximal 5° zeigt Tab. 15. Zwischen 35 und 55 Grad erreichten die Messungen die besten Resultate mit kaum einer Abweichung von der errechneten Geschwindigkeit. Die Unterschätzungen unterhalb von 35 und die Überschätzungen zwischen 55 und 60 Grad waren mit maximal 0,02 m/s gering. Die Werte der einzelnen Messreihen waren hier keinen groben Schwankungen unterworfen. Bei Winkeln oberhalb von 60 Grad nahm die prozentuale Messungenauigkeit steil zu und überschritt mit mittleren prozentualen Fehlwerten von über 20 %, oberhalb von 80 Grad von über 100 %, den tolerierbaren Rahmen. Die Schwankungen der Werte zwischen einzelnen Messreihen, die nicht einem systematischen, sondern einem zufälligen Muster folgten, ergeben sich durch die hohe prozentuale Abweichung bereits bei leichten Fehlschätzungen in diesem Winkelbereich. Dieses Ergebnis bestätigt die theoretischen Berechnungen nach der Doppler-Gleichung (Formeln 4 und 5): Bei der Berechnung der Strömungsgeschwindigkeit führt aufgrund der Cosinus-Funktion im Nenner bei Winkeln oberhalb von 60 Grad ein Fehler bei Über- oder Unterschätzung von 5 Grad zu hohen prozentualen Abweichungen der quantitativen Flussmessung (Tab. 15).

Die Ergebnisse zeigen, dass es unumgänglich ist, auch unter In-Vivo-Bedingungen, bei denen die exakte Winkelbestimmung aufgrund von biologischen Faktoren noch weitaus schwieriger ist, den Winkel bei jeder Messung mit der Winkelkorrekturfunktion des Gerätes zu berichtigen und Messungen mit einem Winkel über 60 Grad wegen der zu hohen Fehlerrate zu verwerfen.

Doppler-Winkel	Abweichung bei 5° Unterschätzung	Abweichung bei 5° Überschätzung
0° - 10°	1,1 %	1,9 %
15° - 30°	2,0 % - 4,4 %	2,8 % - 5,7 %
35° - 55°	5,4 % - 11,0 %	7,0 % - 11,0 %
60° - 70°	12,8 % - 19,0 %	18,0 % - 32,0 %
75° - 80°	24,0 % - 33,0 %	50,0 % - 100,0 %
85°	50,0 %	nicht bestimmbar

Tab. 15: Prozentuale Abweichung dopplersonographischer Messungen der Blutflussgeschwindigkeit bei einem Winkelfehler von 5 Grad (nach [94], S. 23).

Eine bedeutende Rolle für die quantitative Geschwindigkeitsmessung spielt die *Doppler-Empfangsverstärkung*: Bei zu hoch eingestellter Empfangsverstärkung wurden die Geschwindigkeiten leicht überschätzt, bei zu niedrig eingestellter Empfangsverstärkung leicht überschätzt. Die Überschätzung lässt sich mit einer größeren Artefaktanfälligkeit erklären; so werden etwa Bewegungsartefakte bei höherer Empfangsverstärkung mit in das Doppler-Frequenzspektrum einbezogen. Bei der höchsten Doppler-Empfangsverstärkung (87,5 %) füllte das Hintergrundrauschen das ganze Bild aus, so dass die Dopplerkurve nur undeutlich davon abzugrenzen war. Bei zu niedriger Empfangsverstärkung werden, insbesondere im Bereich sehr niedriger Strömungsgeschwindigkeiten, die höheren Frequenzen zum Teil abgeschnitten, was zu einem Unterschätzen der Geschwindigkeit führt. Bei der niedrigsten Stufe der Empfangsverstärkung (12,5 %) war fast kein Doppler-Frequenzspektrum mehr ableitbar.

Ein korrektes Platzieren des *Dopplerfensters* ist für die Messung von Flussgeschwindigkeiten extrem wichtig: Bei einer laminaren Strömung fließt eine Flüssigkeit am schnellsten in der Mitte und wird zum Rand hin langsamer [66] (S. 20). Legt man den Cursor über das ganze Rohr bzw. Gefäß, spiegelt das Doppler-Frequenzspektrum das gesamte Geschwindigkeitsprofil wider. Bei Ableitung nur in der Mitte müssen sich demzufolge falsch zu hohe, bei Ableitung nur am Rand falsch zu niedrige Geschwindigkeiten ergeben. Die Ergebnisse dieses Versuchs entsprachen den theoretischen Überlegungen: Die Geschwindigkeiten bei einem Dopplerfenster, das den ganzen Schlauch einschloss, stimmten nahezu exakt mit der Sollgeschwindigkeit überein. Der nur in der Flussmitte platzierte Cursor ergab dagegen deutlich zu hohe Geschwindigkeiten, durchschnittlich waren sie um 17 % erhöht. Bei am Rand platziertem Cursor dagegen waren die Geschwindigkeiten durchschnittlich um 16,9 % zu niedrig.

Bei der *Geschwindigkeitsmessung in unterschiedlicher Tiefe* zeigte die oberflächennahe Geschwindigkeitsmessung (1 cm Tiefe) einen Trend zur geringgradigen Überschätzung der realen Geschwindigkeit. Die Messungen in 4 cm Tiefe entsprachen relativ genau der errechneten Sollgeschwindigkeit. Wiederholte Messungen in 4 cm Tiefe wiesen allerdings eine größere Variabilität der Werte auf als die oberflächlich gelegenen Messungen. Dieses Phänomen wurde auch von anderen Arbeitsgruppen beim Durchführen von In-Vitro-Versuchen beobachtet [88]. Es kann mit einem steigenden Signalverlust in größerer Tiefe erklärt werden, durch den die Rekonstruktion der Flusskurven durch das Doppler-Gerät verfälscht werden kann.

Untereinander liegende Gefäße haben keinen Einfluss auf die dopplersonographische Ableitung: Sie wird nicht verändert, wenn sie durch ein anderes Gefäß hindurch erfolgt. Dieses Ergebnis war zu erwarten, solange das Dopplerfenster nur ein Gefäß erfasst. Bei versehentlichem Erfassen zweier Gefäße durch den Cursor können falsch zu hohe oder zu niedrige Werte resultieren [94] (S. 35). Dieser Effekt sollte aber nicht Gegenstand der durchgeführten Versuchsreihe sein.

Der Hämatokritwert bzw. der Hämoglobingehalt des Blutes hat innerhalb von physiologischen Bereichen keinerlei Einfluss auf die dopplersonographische Geschwindigkeitsmessung und auf die Blutflussdarstellung im Farbdoppler. Erst weit mehr als eine Zehnerpotenz unterhalb der noch mit dem Leben vereinbaren niedrigsten Hämatokritwerte zeigte sich eine leichte Reduktion des Farbsignales und der Dichte des Doppler-Frequenzspektrums bei noch korrekter Geschwindigkeitsmessung. Ein annäherndes Verschwinden des Farbsignales und ein deutliches Ausdünnen des Frequenzspektrums war erst bei Hunderterpotenzen unterhalb der noch mit dem Leben vereinbaren niedrigsten Hämatokritwerte zu sehen. Die bei Karzinompatienten häufig vorhandene Anämie wirkt sich also weder auf die Darstellung der Vaskularisation des Tumors im Farbdoppler noch auf die semiquantitative Geschwindigkeitsbestimmung Tumorgefäßen von in der Dopplersonographie aus.

Einen wesentlichen Einfluss hat die Geräteeinstellung auf die **farbkodierte Dopplersonographie**. Je sensibler die Geräteeinstellung, bei umso geringeren Geschwindigkeiten - und gegebenenfalls auch bei ungünstigem Winkel - kann ein Fluss dargestellt werden und desto breiter ist der Farbstrahl des sichtbaren Frequenzspektrums. Da eine gesteigerte Vaskularisation ein Hauptkriterium für Malignität ist, kann dieser Punkt entscheidende Auswirkungen auf die sonographische Dignitätsbeurteilung gynäkologischer Tumoren haben.

In der farbkodierten Dopplersonographie bei Simulation zweier untereinander liegender Gefäße wies das oben liegende Gefäß konstant einen größeren Innendurchmesser auf. Dieser kann durch die geringere Entfernung zum Schallkopf bedingt sein. Der Innendurchmesser wurde aber auch im unten liegenden Gefäß deutlich zu groß gemessen. Die Überschätzung des Innendurchmessers auch in anderen Versuchen zeigt, dass die farbkodierte Dopplersonographie zum Ausmessen des Durchmessers von Gefäßen nicht verwendet werden sollte. da diese Bestimmung von zu vielen unterschiedlichen Einflüssen (Geräteeinstellung, Tiefe des Gefäßes etc.) abhängt.

Bei höherer maximaler Geschwindigkeitsskala im CDI-Modus werden höhere Geschwindigkeiten zur Darstellung eines Blutflusses mittels farbkodierter Dopplersonographie benötigt. Die zu messenden Minimalgeschwindigkeiten bewegen sich dabei jedoch in sehr niedrigen Bereichen zwischen 0,0011 m/s bei einer Geschwindigkeitsskala von 0,023 m/s und 0,0157 m/s bei einer Geschwindigkeitsskala von 0,16 m/s. Damit ist eine lückenlose Darstellung des Blutflusses im Farbmodus gerade noch bei tatsächlichen Flussgeschwindigkeiten möglich, die ca. zehnfach niedriger als die eingestellte maximale Geschwindigkeitsskala sind. Bei einer Geschwindigkeit von 0,0011 m/s wäre keine Ableitung eines gepulsten Dopplersignales möglich, da das Gerät auf zwei Stellen hinter dem Komma rundet. In sehr kleinen Tumorgefäßen mit sehr langsamem Blutfluss kann in diesen Bereichen jedoch die Entscheidung "Vaskularisation vorhanden" bzw. "nicht vorhanden" durch eine wenig sensible Geräteeinstellung verfälscht werden. Bei der In-vivo-Dignitätsdiagnostik wurde manchen Tumoren eine Vaskularisation festgestellt, das Ableiten bei eines Dopplersignales war jedoch nicht möglich, so dass die Flussgeschwindigkeiten bei diesen Fällen vermutlich in sehr niedrigen Bereichen lagen.

Die winkelabhängige Grenze der Gefäßdarstellung bei unterschiedlicher maximaler Geschwindigkeitsskala (CDI-Modus) zeigt, dass umso steilere Winkel benötigt wurden, je höher die maximale Geschwindigkeitsskala eingestellt war. Bei einer niedrigen maximalen Geschwindigkeitsskala von 0,023 m/s konnte bei einem Pumpvolumen von 100 ml/h (durchschnittliche Flussgeschwindigkeit im Gefäß 0,035 m/s, maximale Geschwindigkeit in der Flussmitte 0,07 m/s) ein Blutfluss bei Winkeln von durchschnittlich 85 Grad (minimal 87 Grad) nachgewiesen werden. Bei einer Geschwindigkeitsskala von 0,09 m/s musste der Winkel durchschnittlich 80 Grad, bei einer Geschwindigkeitsskala von 0,17 m/s durchschnittlich 76 Grad betragen, um noch einen Fluss darzustellen.

Mittels *Color Doppler Energy (CDE-Modus)* konnten etwas geringere Geschwindigkeiten detektiert werden als mit der konventionellen farbkodierten Dopplersonographie. Durch das *Variieren des Dynamikbereiches der Energie* reichten die Grenzen von einer minimalen durchschnittlichen Geschwindigkeit von 0,00035 m/s bei der sensitivsten Einstellung (40 dB) bis zu 0,0014 m/s bei einem Dynamikbereich der Energie von 20 dB. Somit kann der energiekodierte Farbdoppler die Zahl der gut- oder bösartigen Neubildungen erhöhen, bei denen eine Vaskularisation nachgewiesen werden kann, während diese mit dem farbkodierten Doppler (CDI-Modus) aufgrund langsamer Flussgeschwindigkeiten als "nicht vaskularisiert" eingestuft wurden.

Auch das *Variieren der Farbintensität* sowohl im *CDI*- als auch im *CDE-Modus* kann zu einer sensibleren Gefäßdarstellung beitragen: Je höher die Farbintensität eingestellt wird, umso dicker imponiert das Gefäß. Bei einer sehr niedrig eingestellten maximalen Geschwindigkeitsskala (0,023 m/s) im CDI-Modus und sehr niedriger Farbintensität wurde die Schlauchdicke gering überschätzt, mit weiterem Aufdrehen des Farbintensitäts-Reglers nahm die Überschätzung deutlich zu. Sie erreichte das Dreifache des tatsächlichen Innendurchmessers bei maximal hoher Farbintensität. Im CDE korrelierte auch bei sensitiv eingestellter Energieskala (35 dB) eine mittlere Farbintensität gut mit dem tatsächlichen Schlauchdurchmesser, während niedrigere Farbintensitäten zur Unterschätzung und höhere Farbintensitäten zur Unterschätzung führten. Ähnlich verhält sich der CDI-Modus bei einer mittleren maximalen Geschwindigkeitsskala (0,09 m/s).

Eine korrekte Bestimmung des Gefäßdurchmessers scheint damit aufgrund der zahlreichen Einflussfaktoren auf das Farbdopplersignal (tatsächliche Flussgeschwindigkeit, Einstellung der maximalen Geschwindigkeitsskala, Winkel zwischen Schallstrahl und Gefäßachse, Wahl der Farbintensität) nicht praktikabel. **Zusammenfassend** lassen sich aus dem Strömungsmodell folgende Aussagen über die Genauigkeit und Reproduzierbarkeit von dopplersonographischen Blutflussmessungen treffen:

- 1. Flussgeschwindigkeiten lassen sich mit der gepulsten Dopplersonographie mit ausreichender Reproduzierbarkeit messen.
- Oberhalb einer tatsächlichen Maximalgeschwindigkeit von 0,035 m/s weisen die dopplersonographischen Messungen eine hohe Genauigkeit auf. Unterhalb einer Maximalgeschwindigkeit von 0,035 m/s werden die tatsächlichen Geschwindigkeiten mit der gepulsten Dopplersonographie leicht überschätzt.
- 3. Bei einem Winkel zwischen Schallstrahl und Gefäßachse zwischen 35 und 55 Grad ist eine hohe Genauigkeit der dopplersonographischen Flussgeschwindigkeiten zu erzielen. Insbesondere bei flachen Winkeleinstellungen (> 60 Grad) zeigt sich eine exponentielle Zunahme der Messfehler bezüglich Richtigkeit und Reproduzierbarkeit der Messungen.
- Die dopplersonographische Beurteilung der Flussgeschwindigkeit bei tief gelegenen Gefäßen ist im Vergleich zu oberflächennahen Gefäßen einer höheren Schwankungsbreite unterworfen.
- 5. Für eine exakte duplexsonographische Geschwindigkeitsmessung sind folgende Voraussetzungen erforderlich:
 - Wahl der Empfangsverstärkung bei ca. 50 % für eine klare Kontrastierung des Frequenzspektrums
 - Gefäßdeckende Größe und Position des Echofensters
- 6. Die biologische Einflussgröße des Hämatokritwertes, bzw. der Anzahl der korpuskulären Reflektoren, scheint keinen Einfluss auf die Genauigkeit oder Sensitivität dopplersonographischer Flussmessungen aufzuweisen.
- Die Bestimmung des Gefäßdurchmessers mit der farbkodierten Dopplersonographie ist aufgrund unterschiedlicher Einflussfaktoren nicht praktikabel. Bei einer in der Regel bevorzugten sensitiven Geräteeinstellung werden die Gefäßdurchmesser deutlich überschätzt.

- 8. Für eine lückenlose Blutflusserfassung im farbkodierten Doppler (CDI-Modus) darf die Einstellung der maximalen Geschwindigkeitsskala höchstens das Zehnfache der tatsächlichen Flussgeschwindigkeit betragen. Je steiler dabei der Winkel zwischen Schallstrahl und Gefäßachse verläuft, desto besser wird der tatsächliche Blutfluss registriert.
- 9. Mit dem Powerdoppler (CDE-Modus) gelingt eine deutlich sensitivere Erfassung von langsamen Flüssen gegenüber dem farbkodierten Doppler (CDI-Modus).

5.2 Bedeutung der Kombination von Sonographie und Dopplersonographie für die Diagnostik von Ovarialtumoren

Die transvaginale Sonographie und die Dopplersonographie sind ein unverzichtbares Instrument für die Entdeckung und präoperative Evaluation von Ovarialtumoren. In der vorliegenden Arbeit konnte gezeigt werden, dass die dopplersonographische Beurteilung der Vaskularisation von Adnex-Tumoren die B-Mode-sonographische Diagnostik verbessert.

Durch die Verbesserung der Ultraschallgeräte wurden in den letzten Jahren neue Aspekte sichtbar und dadurch neue Kriterien entwickelt, um Scores zu verbessern oder typische Malignitätskriterien zu finden. Viele dieser Kriterien haben sich aber als weniger aussagekräftig erwiesen. So war etwa das Vorhandensein eines echodichten Randsaumes in dieser Studie häufig schwer nachzuvollziehen, es erwies sich bei der dadurch stark reduzierten Fallzahl als nicht signifikant. Den höchsten Signifikanzgrad unter den morphologischen Kriterien wies in dieser Arbeit das von Fleischer bereits 1978 genannte Unterscheidungsmerkmal der Tumorstruktur (rein zystisch, zystisch-solide und rein solide) auf [28]. Der Nachweis solider Anteile zeigte sich als zuverlässigstes Kriterium mit der bei entsprechend hoher Fallzahl (n = 63) bestmöglichen Treffsicherheit von 78 %. Daneben zeigte das Volumen, ebenfalls eines der bereits in sehr frühen Veröffentlichungen untersuchten Merkmale, hohe Validitätsraten von 67 % bis 71 %. Da jedoch auch gutartige Prozesse große Ausdehnung erreichen können (in der vorliegenden Untersuchung hatten zwei seröse Zystadenome, ein muzinöses Zystadenom und eine einfache Ovarialzyste Volumina von über 2000 ml), andererseits maligne Tumoren, besonders bei Erfassung in frühen Stadien, sehr klein sein können, kann dieser Parameter aber nicht als alleiniges Unterscheidungskriterium dienen.

Die morphologische Unterscheidung "zystisch" und "solide" und das getrennte Untersuchen der jeweiligen Anteile kann die Treffsicherheit erhöhen; bei den dadurch produzierten geringeren Fallzahlen waren einige Parameter jedoch nicht mehr signifikant. So erreichte keiner der Parameter zur Beurteilung solider Anteile (*Homogenität*, *echodichter Randsaum, Randkontur, Begrenzung*) signifikante Zahlen. Ein erneutes Überprüfen an einem größeren Kollektiv könnte hier jedoch Erfolg versprechend sein. Bei den zystischen Formationen mussten vor allem die schwierig erfassbaren und vom Untersucher abhängigen Punkte gestrichen werden: Alle *papillären Strukturen* maßen über 3 mm, eine Unterscheidung zwischen benigne und maligne war hier also nicht möglich. Besser, aber ebenfalls nicht signifikant waren die *Dicke der Zystenwand* und die *Septendicke*, auch hier spielt die geringe Fallzahl sicherlich eine Rolle. Auch das Vorhandensein von *papillären Strukturen* und die *Begrenzung der Zystenwand* sehr sensible Parameter dar (Validität 77 % bzw. 81 %).

Die von der Untereinteilung in "zystisch" und "solide" unabhängigen Merkmale "*Aszites*" und "*Infiltration*" bieten insofern Probleme, als das Vorhandensein bei der Infiltration - sofern vom Untersucher richtig beurteilt - eindeutig, beim Aszites mit hoher Wahrscheinlichkeit für Malignität spricht. In beiden Fällen schließt die Abwesenheit aber einen malignen Prozess nicht aus. Da auch manche benigne Ovarialtumoren, etwa bestimmte Zysten, mit Aszites einhergehen können oder der Aszites durch andere Krankheiten wie Leberzirrhose oder Peritonealmetastasen nicht-ovarieller Karzinome hervorgerufen werden kann, erwies sich dieses Kriterium im Gegensatz zur Infiltration als nicht signifikant.

In der *multivariaten Analyse der sonographischen Parameter* konnte das zuverlässigste Kriterium, das Vorhandensein eines soliden Anteils, durch keinen Parameter verbessert werden. Die diagnostische Zuverlässigkeit konnte somit nicht auf über 78 % erhöht werden. Durch das Zurückgreifen auf ein größeres Patientengut könnte bei erhöhter Fallzahl und damit eventuell verbesserter Signifikanz der einzelnen Parameter ein größeres Kombinationsspektrum erzielt werden, um eine verbesserte multivariate Bewertung von zystischen und soliden Arealen zu erlauben.

In der aktuellen Literatur wird die diagnostische Treffsicherheit der B-Mode-Sonographie bei der Beurteilung von Adnex-Tumoren zwischen 75 % und 95 % angegeben [68], [106]

(S. 430). Eine große Studie an insgesamt 754 Ovarialtumoren beschreibt eine hohe Treffsicherheit der transvaginalen Sonographie von 86,6 % (Patientinnen in der Postmenopause) bis 92 % (Patientinnen in der Prämenopause). Dabei waren die sonographischen Parameter ebenfalls retrospektiv mit einer multivariaten Analyse ausgewertet worden, wobei sich der solide Anteil ebenfalls als validester Parameter vor Kriterien wie dem Nachweis von Aszites und der Bewertung der Tumorgröße erwies [87]. In vielen Arbeiten zeichnet sich die Sonographie bei einer guten Treffsicherheit zwar durch sehr hohe Sensitivitäten von bis zu 100 % aus, erreicht aber vergleichsweise deutlich schlechtere Spezifitäten von nur rund 70 % [12], [106] (S. 425). Die vorliegende Studie konnte diesen Unterschied mit einer Sensitivität von 95 %, aber nur einer Spezifität von 68 % bestätigen.

Nur ein seröses Adenokarzinom zeigte in der vorliegenden Arbeit im sonographischen Bild keinen soliden Anteil. Als falsch-positiv wurden neben benignen Tumoren mit soliden Arealen (Teratome, Zystadenome) die zystischen Prozesse bewertet, in denen Einblutungen oder eine Eindickung des Zysteninhaltes sonographisch eine solide Struktur erscheinen ließen. Neben einer Corpus luteum-Zyste und einer Epithelzyste waren dies zwei Inklusionszysten und eine Endometriosezyste.

Die "pseudo-soliden" Anteile dieser Zysten lassen sich durch eine dopplersonographische Beurteilung der Durchblutung besser differenzieren. Die stetige Verbesserung der Dopplergeräte in den vergangenen Jahren führten neben einer Erfassung auch langsamerer Blutflussgeschwindigkeiten zu einer empfindlicheren Darstellung des funktionellen Durchblutungsverhaltens von Tumoren. Dadurch musste die diagnostische Aussage der Dopplersonographie jedoch auch neu interpretiert werden, so dass die Euphorie der Anfänge immer mehr kritischen Stimmen wich [12]. So musste die These, dass sich maligne Tumoren im Gegensatz zu benignen durch die Anwesenheit von Blutgefäßen auszeichnen, teilweise korrigiert werden. In einer der früheren Studien wurde bei 93 % aller malignen und nur bei 35 % aller benignen Tumoren eine Gefäßversorgung nachgewiesen [62]. Folgestudien zeigten jedoch auch eine Durchblutung der weniger gut vaskularisierten benignen Prozesse an, die früher als "nicht vaskularisiert" klassifiziert wurden [107], [109]. Die Durchblutung des Ovars und ovarieller Prozesse ist zudem Menopausenstatus, Zyklusphase (periovulatorisches abhängig von Ovar), Hormoneinnahme und Tumorart. Entzündliche Vorgänge sind beispielsweise sehr stark durchblutet. Auch hämorrhagische Zysten, Corpora Lutea, Dermoid-Tumoren und Endometriosen können durch ihre hohen Vaskularisationsraten und die niedrigen Gefäßwiderstände einen bösartigen Tumor imitieren [31].

In dem vorliegenden Kollektiv wurde zwar bei 90 % der malignen Tumoren eine Vaskularisation nachgewiesen, aber auch bei 73 % aller benignen Tumoren, der Unterschied ist somit nicht signifikant. In manchen Veröffentlichungen werden sogar noch höhere Vaskularisationswerte sowohl benigner als auch maligner Prozesse genannt [85].

Die Zahl maligner Tumoren, bei denen auch mit der heutigen Gerätetechnik keine Vaskularisation darstellbar ist, ist relativ gering. Sie betrifft aber vor allem kleine Neoplasien in frühen Stadien, also gerade diejenigen, die noch eine bessere Heilungschance aufweisen. In der vorliegenden Studie waren ein seröses Adenokarzinom und ein Tumor niedrig maligner Potenz (LMP) nicht vaskularisiert. Das seröse Adenokarzinom befanden sich zwar im FIGO-Stadium III c, wies also bereits regionäre Lymphknotenmetastasen bzw. Peritonealmetastasen im Größenbereich über 2 cm auf, war jedoch von der lokalen Ausdehnung mit einem Volumen von 14 ml eher klein.

Bislang selten diskutiert wird in der Literatur ein Parameter, der sich in dieser Studie als der sensibelste ($p < 10^{-5}$) zur Erkennung maligner Prozesse erwiesen hat: das Vaskularisationsmuster [17], [56]. Während sich bei 95 % der malignen Tumoren Gefäße direkt im Tumorzentrum fanden (zentrale Vaskularisation), zeigte das Durchblutungsmuster gutartiger Befunde eher eine periphere Anordnung, die Blutgefäße standen also bevorzugt am Tumorrand beziehungsweise in der Zystenwand oder in Septen. Lediglich zwei Luteinzysten und eine Hämatosalpinx waren zentral vaskularisiert und senkten damit geringfügig die Spezifität auf 90 %. Ein seröses Adenokarzinom im FIGO-Stadium III c mit einer nur geringen Lokalausdehnung (Volumen 97 ml) zeigte keine zentrale Vaskularisation. Die Validität betrug damit 92 %. Man nimmt an, dass peripher lokalisierte Gefäße Teil der normalen, bereits vor der Tumorentwicklung vorhandenen Vaskularisation des Ovars sind, wogegen sich zentrale Gefäße aufgrund eines Proliferationsreizes, des Angiogenese-Faktors, gebildet haben [61]. Allerdings ist auch hier fraglich, inwieweit die Weiterentwicklung der Geräte diesen Punkt beeinflussen wird: Durch die Entwicklung sensitiverer Geräte und neuer Doppler-Verfahren wie der Color Doppler Energy (CDE) können immer kleinere Gefäße mit langsameren Blutströmungen erfasst werden, dadurch ist zu erwarten, dass in der Zukunft bei immer mehr benignen Tumoren Vaskularisation sichtbar gemacht wird. Zudem wird die Güte dieses Parameters dadurch getrübt, dass nicht alle Patientinnen in die Untersuchung mit einbezogen werden können: Die Tumoren, die insgesamt keine Vaskularisation aufweisen, müssen für die Frage "zentral oder rein peripher vaskularisiert" natürlich ausgeschlossen werden.

Die *Gefäßanzahl* erwies sich bei einem optimierten Grenzwert von 5 Gefäßen zwar als signifikantes Kriterium (p < 0,01). Mit einer guten Spezifität von 86 %, aber nur 55 % Sensitivität ist sie für die Dignitätsbeurteilung ein eher schwacher Parameter. Sie ist in der Praxis manchmal schwierig zu messen, da etwa stark gewundene Gefäße leicht doppelt gezählt werden können, und einzelne Schnittebenen oft kein repräsentatives Tumorareal darstellen.

Ein "*Notch*", die frühdiastolische Einkerbung der Blutflusskurve als Zeichen einer muskelstarken Gefäßwand, zeigte sich häufiger bei benignen als bei malignen Gefäßen. Da der Anteil der benignen Gefäße mit Notch allerdings auch relativ gering war, erwies sich dieses Kriterium als nicht signifikant. Die Rate an benignen Tumoren mit Notch betrug 26 %, während nur einer von 77 malignen Tumoren dieses Gefäßphänomen aufwies (Sensitivität 6 %). Diese Zahlen entsprechen den in der Literatur zum Teil genannten Größenordnungen [106] (S. 83).

Aufgrund der methodischen Probleme bei den qualitativen dopplersonographischen Kriterien stützen sich die Arbeitsgruppen in letzter Zeit zunehmend auf die quantitativen Parameter. Vor allem RI und der PI werden favorisiert [43], [62], [92], [95], da sich bestimmte Messfehler bei der Bestimmung der systolischen Maximalgeschwindigkeit V_{max} und der mittleren Maximalgeschwindigkeit V*mean*, etwa durch falsche Winkeleinstellung, durch den Bruch herauskürzen. Der PI müsste theoretisch noch aussagekräftiger sein als der RI, weil er durch das Einbeziehen der gemittelten maximalen Geschwindigkeit mehr Information beinhaltet. Es hat sich aber gezeigt, dass beide ähnliche Resultate erlangen [18], was sich unter anderem darin begründet, dass v_{mean} eine höhere Artefaktanfälligkeit aufweist als v_{max}. V_{mean} muss zudem an den meisten Dopplergeräten von Hand ausgemessen werden, wodurch zusätzliche Ungenauigkeiten entstehen. Mit beiden Indizes konnten in der vorliegenden Studie hohe Treffsicherheiten von bis zu 86 % erzielt werden, wobei sich der RI als geringfügig überlegen erwies. Der Mittelwert des RI und des PI aller Gefäße etwa besaßen mit 80 % die gleiche Sensitivität, der RI hatte aber mit 67 % Spezifität eine geringfügig bessere Treffsicherheit als der PI mit 63 % Spezifität (Validität 72 % gegenüber 70 % beim PI). Sie liegen damit in vergleichbarem Rahmen mit den von anderen Studiengruppen erzielten Sensitivitäten und Spezifitäten (**Tab. 16**). Der anhand von ROC-Kurven in dieser Arbeit ermittelte beste Grenzwert für den PI lag mit 1,27 gering über dem höchsten in der Literatur verwendeten Grenzwert von 1,25 [55], [56]. Der beste Grenzwert für den RI lag mit 0,63 in der Mitte der in der Literatur verwendeten Grenzwerte, deren Spannbreite von 0,4 bis 0,8 reichen [12], [27], [43], [62], [86] (S. 408), [92], [105] (**Tab. 16**). Allerdings zogen manche Studiengruppen den niedrigsten Wert mehrerer Gefäße zur Auswertung heran. Die Streuung innerhalb der malignen und der benignen Gruppe war aber relativ hoch (**Abb. 27** und **Abb. 28**), so dass der RI und der PI nicht isoliert für die Dignitätsdiagnostik herangezogen werden sollte.

Veröffentlichung	<u>Σ mal.</u> n	Grenz- wert PI	Sens. / Spez. [%]	Grenz- wert RI	Sens / Spez. [%]
Hata et al., 1992 [43]	27/63			0,72	93/53
Kurjak et al., 1992 [62]	29/83			0,4	96/99
Weiner et al., 1992 [107]	17/53	1,0	100/83		
Schneider et al., 1993 [92]	16/55			0,8	94/59
Timor-Tritsch et al., 1993 [105]	16/94	0,62	88/97	0,46	94/99
Bromley et al., 1994 [12]	12/33			0,6 0,8	66/81 91/52
Carter et al., 1994 [17]	88/228	1,0	57/78	0,4	13/97
Jain, 1994 [49]	10/50			0,4	70/82
Kawai et al., 1994 [56]	40/109	1,25	75/79		
Prömpeler et al., 1994 [86]	45/129	0,7	83/77	0,5	85/77
Franchi et al., 1995 [34]	37/129			0,65	76/72
Fleischer et al., 1996 [27]	26/126	1,0	92/88		

Tab. 16: Vergleich verschiedener in der Literatur angegebener Grenzwerte für Pulsatilitäts- und Resistance-Index bei der dopplersonographischen Untersuchung von Adnex-Tumoren mit Sensitivitäten (Sens.) und Spezifitäten (Spez.). Zahl maligner Tumoren zur gesamten Tumorzahl = Σ mal. / n.

Auffällig ist, dass in der Literatur die ersten Veröffentlichungen sehr hohe Sensitivitäten und Spezifitäten von zum Teil über 90 % erreichten [62], [105], [107]. In späteren Arbeiten lagen die beiden Beurteilungswerte zum Teil deutlich niedriger [12], [17], [34]. Die Trennung in zentrale und periphere Vaskularisation erfolgte in den meisten Veröffentlichungen nicht. In der vorliegenden Studie konnte vor allem durch Heranziehen der zentralen Durchblutungswerte eine deutliche Verbesserung der Validitäten mit 84 % Sensitivität und 67 % Spezifität für den PI sowie 89 % Sensitivität und 67 % Spezifität für den RI erzielt werden. Da die Lokalisation der Tumorgefäße den Flusswiderstand und die systolische Maximalgeschwindigkeit beeinflusst [30], wirkt sich dies auch auf unterschiedliche Grenzwerte zentraler und peripherer Gefäße aus: Die Grenzwerte der errechneten Parameter PI, RI und AB-Ratio lagen bei den peripheren Gefäßen deutlich über den entsprechenden Werten der zentralen Gefäße. Die Grenzwerte für die zentrale Vaskularisation (0,59 für den PI; 0,4 für den RI) lagen deshalb auch deutlich niedriger als für die Gesamtvaskularisation.

Die systolische Maximalgeschwindigkeit V_{max} aller Gefäße zeigte in der untersuchten Patientengruppe sehr starke Schwankungen, so dass sie sich häufig als nicht signifikant erwies. Lediglich der Höchstwert der zentralen und Gesamtvaskularisation war mit einem Grenzwert von jeweils 0,2 m/s signifikant. Somit stellte sich die systolische Maximalgeschwindigkeit als der anfälligste Parameter heraus, was zum Teil auch den Ergebnissen anderer Studiengruppen entspricht [17]. Allerdings kontrastiert dieses Resultat mit den deutlich besseren Ergebnissen von V_{mean} und V_{min} in dieser Arbeit. Die Grenzen des Mittelwertes lagen bei 0,06 m/s für V_{min} gesamt und bei 0,13 m/s für V_{mean} gesamt. Oberhalb dieser Werte wird, wie bei V_{max}, Malignität angenommen. Die AB-Ratio zeigte sich etwas instabiler als RI und PI, so dass in der vorliegenden Studie, analog zu vielen Veröffentlichungen [12], [17], [34], [43], [49], [56], [62], [86], [92], [105], [107], diesen der Vorzug gegeben wurde. In die multivariate Analyse wurde der AB zwar mit einbezogen, war jedoch dort ebenfalls dem RI und PI unterlegen.

In der multivariaten Analyse der dopplersonographischen Untersuchungen erwiesen sich die *Gesamtdurchblutung* und die *zentrale Vaskularisation* als die sensibelsten Kriterien. Durch die Kombination dieser beiden qualitativen Parameter in einer Entscheidungsmatrix konnte die Validität der Dopplersonographie damit auf 90 % gesteigert werden bei einer Sensitivität von 86 % und einer Spezifität von 93 %.

Mit der Bestimmung der quantitativen Blutflussparameter konnte keine Verbesserung der Treffsicherheit erzielt werden. Lokalisation und Ausmaß der Vaskularisation scheinen die Abläufe der Neoangiogenese und die Tumorbiologie besser darzustellen als die quantitativen Flussparameter. Zudem ist die Ovarialdurchblutung stark von funktionellen Faktoren beeinflusst, von denen die quantitativen Dopplerparameter besonders stark abhängig sind. Bereits bei der Ausbildung eines Corpus luteum lässt sich neben der Zunahme der Durchblutung auch eine Abnahme des Widerstandes in ovariellen Gefäßen beobachten [78]. Auch in der Follikelphase kann sich das physiologische Flussmuster stark verändern [101]. Die Reproduzierbarkeit der Widerstands- und Geschwindigkeitsmessungen ist unter anderem deshalb, vor allem aber aufgrund methodischer Probleme sehr beschränkt [100]. Diese methodische Problematik entfällt zu einem Großteil bei der Beurteilung der intratumoralen Gefäßverteilung.

In den meisten bisher veröffentlichten Studien wurde die dopplersonographische Untersuchung von Adnex-Tumoren isoliert betrachtet. Die gepulste und farbkodierte Dopplersonographie ist aber keine konkurrierende, sondern eine ergänzende Methode zur Transvaginalsonographie. Sie sollte daher nur in Zusammenhang mit der sonomorphologischen Diagnostik durchgeführt und beurteilt werden. Erst in letzter Zeit bestehen Versuche, beide Methoden zu kombinieren [3], [39], [42]. Aufgrund unterschiedlicher Beurteilungs- und Auswertungsverfahren schwanken die Angaben über die Treffsicherheiten jedoch noch stark.

In der vorliegenden Arbeit wurde mit Hilfe der multiplen logistischen Regression eine Entscheidungsmatrix entwickelt, welche die sonographischen und dopplersonographischen Parameter nach ihrer diagnostischen Bedeutung und ihrer Unabhängigkeit voneinander gewichtet. Dadurch konnte die Zahl richtig eingestufter Tumoren im Vergleich zur isolierten B-Mode-Sonographie um acht Tumoren auf 57 von 63 erhöht werden, was einen Zugewinn an Treffsicherheit von 12 % bedeutet. Durch die Dopplersonographie konnte vor allem die Spezifität der B-Mode-Sonographie deutlich verbessert werden. Der sonographische Nachweis soliden Tumorarealen lässt sich durch die von dopplersonographische Analyse von Gesamtdurchblutung (Vaskularisation) und Gefäßanordnung weiter differenzieren. Nur ein benigner Tumor mit einem soliden Anteil wies eine zentrale Durchblutung auf. Damit erreichte die kombinierte sonographische und dopplersonographische Untersuchung eine Treffsicherheit von 90 %. Durch das Einbeziehen von qualitativen dopplersonographischen Kriterien konnte die Treffsicherheit nicht weiter gesteigert werden. Allerdings erzielt eine aus den Parametern solide Anteile, Vaskularisation, Mittelwert der PI aller Gefäße und Mittelwert des V_{mean} aller Gefäße kombinierte Entscheidungsmatrix die gleiche Sensitivität (86 %), Spezifität (93 %) und Validität.

Der Entscheidungsbaum bietet den Vorteil, dass die einzelnen Charakteristika je nach ihrer statistischen Treffsicherheit verschieden starke Gewichtungen erhalten und nicht, wie im System der Punkte-Scores üblich, gleich behandelt werden ohne Rücksicht auf ihren Stellenwert. Zudem können Verknüpfungen zwischen verschiedenen Parametern in der sinnvollen Reihenfolge berücksichtigt werden, so etwa die Abfolge Vaskularisation und zentrale Vaskularisation. Die Auswahl erfolgt durch schrittweise Selektion der einzelnen Kriterien, nach jeder Aufzweigung kann eine erneute Berechnung der Validitäten erfolgen. Etwaige Abhängigkeiten von Parametern untereinander können durch die statistische Analyse aufgedeckt und vor dem Erstellen des Entscheidungsbaumes eliminiert werden. Die Treffsicherheit ist damit deutlich höher als in der früher angewandten Kombination der Kriterien mittels Punkte-Scores (Tab. 17), die sich zudem meist nur an sonomorphologischen Kriterien orientieren, ohne die Dopplersonographie einzubeziehen [12], [22], [69], [90]. Der von Sassone et al. 1991 entwickelte Punktescore, der initial mit einer maximalen Sensitivität von 100 % und guten Spezifität von 83 % beschrieben wurde [90], wurde von einige Studiengruppen untersucht. Diese erzielten damit jedoch weit schlechtere Ergebnisse [43], [92], [105]. Kombinierte Score-Systeme existieren nur wenige, und erst seit kurzem bestehen Versuche, das starre Punktesystem durch die multiple logistische Regression zu ersetzen [99], mit deutlich besseren Sensitivitäten und Spezifitäten.

Allerdings ist die Aussagekraft der multivariaten logistischen Regression bei nicht sehr hohen Fallzahlen nur als relativ anzusehen. Durch die Prüfung, Weiterentwicklung und eventuell Differenzierung des Entscheidungsbaums an einem größeren Patientenkollektiv können sicher noch weitaus bessere Ergebnisse erreicht werden. Alternativ hat sich auch gezeigt, dass die Auswertung mit neuronalen Netzen, einer Methode der künstlichen Intelligenz, zu exzellenten Ergebnissen führt [98] (S. 21), [103] (S. 17). Die Ultraschalluntersuchung wird aber immer von der individuellen Erfahrung des Untersuchers abhängig bleiben [102], was die Qualität der verschiedenen Auswertungsmethoden immer etwas trüben wird.

Veröffentlichung	<u>Σ mal.</u> n	Art des Scores	Kompo- nenten	Sens. [%]	Spez. [%]
Sassone et al., 1991 [90]	18/136	Punktescore	US	100	83
Hata et al., 1992 [43]	27/63	Sassone-Sc.	US	85	69
Kurjak et al., 1992 [62]	29/83	Punktescore	US + FD	48	98
Schneider et al., 1993 [92]	16/55	Sassone-Sc.	US	88	75
Timor-Tritsch et al., 1993 [105]	16/94	Sassone-Sc.	US	94	87
Bromley et al., 1994 [12]	12/33	Punktescore	US	91	52
DePriest et al., 1994 [22]	44/213	Punktescore	US	89	73
Lerner et al., 1994 [69]	42/350	Punktescore	US	97	77
Tailor et al., 1997 [99]	15/67	MLR	US + FD	87	98

Tab. 17: Vergleich der Sensitivitäten und Spezifitäten verschiedener Score-Systeme und Modelle multipler logistischer Regression zur Beurteilung von Adnex-Tumoren. US = Ultraschall, FD = Farbdoppler, MLR = multiple logistische Regression. Zahl maligner Tumoren zur gesamten Tumorzahl = Σ mal. / n.

5.3 Einsatz neuronaler Netze und der multiplen logistischen Regression

zur Auswertung sonographischer und dopplersonographischer

Untersuchungen

In den letzten Jahren hat die Rolle der Sonographie und Dopplersonographie zur präoperativen Dignitätsbeurteilung von unklaren Befunden der Mamma an Bedeutung gewonnen, besonders bei radiologisch "dichten" Brüsten. Trotz der in manchen Veröffentlichungen beschriebenen hervorragenden Sensitivitäten und Spezifitäten [19], [96], die der Mammographie zum Teil nicht mehr nachstehen, konnte sie sich jedoch noch nicht als alternatives Screening-Instrument durchsetzen. Dies ist sicherlich unter anderem durch eine Vielzahl noch bestehender diagnostischer Probleme und Unsicherheiten bedingt [59], [67]. Durch die gemeinsame Anwendung von B-Mode-Sonographie und gepulster und farbkodierter Dopplersonographie sowie die Kombination beider mittels statistischer Methoden wie der multiplen logistischen Regression oder Methoden der künstlichen Intelligenz wie den neuronalen Netzen kann die diagnostische Wertigkeit deutlich gesteigert werden.

In der vorliegenden Studie konnten an dem als retrospektiv klassifizierten Patientengut sehr hohe Sensitivitäten und Spezifitäten erzielt werden. Die mittels multipler logistischer Regression erstellte Entscheidungsmatrix erreichte eine Treffsicherheit von 97 % (Sensitivität 98 %, Spezifität 93 %), die neuronalen Netze eine Treffsicherheit von 100 %. In der prospektiven Analyse konnten mit der Entscheidungsmatrix immerhin Validitätswerte von 87 %, mit den neuronalen Netzen von 88 % erzielt werden. Die mit multivariater logistischer Regression oder mit neuronalen Netzen erzeugte Validität war somit zwar im Vergleich zum retrospektiven Patientengut gesunken, sie kann dennoch als gut bezeichnet werden. Die Treffsicherheit der prospektiven Gruppe im Vergleich zu dem besten Einzelkriterium, der Randkontur, konnte mit statistischen Methoden und mit den neuronalen Netzen leicht um zwei bis drei Prozentpunkte erhöht werden.

Dass die fast durchweg exzellenten Ergebnisse der Lerndurchgänge bei den Testdurchläufen nicht gehalten werden konnten, ist ein Phänomen, das bei der Anwendung neuronaler Netze nicht selten angetroffen wird [82]. Häufig liegt dies an einem zu starken Ummodellieren der Daten. Das Netz speichert dann unsinnige Informationen, die nicht zur Unterscheidung beitragen. Dem kann durch striktes Einhalten bestimmter Verhältnisse von eingegebenen Datensätzen zu Gewichten entgegengesteuert werden, die in der Literatur mit mindestens 3:1 bis 10:1 angegeben werden [2] (S. 173), [75] (S. 548). Bei den in der vorliegenden Studie eingegebenen 142 Gewichten müsste das Netz also als unterstes Minimum an 426 Datensätzen trainiert werden. Die in dieser Arbeit tatsächlich vorhandenen Patientendaten lagen deutlich darunter: Das Lernen erfolgte hier, auch zur besseren Vergleichbarkeit mit der multiplen logistischen Regression, meist nur an 87 Datensätzen. Eine Erhöhung der Datensätze erbrachte auch tatsächlich eine deutliche Verbesserung der prospektiven Ergebnisse. Eine andere Möglichkeit wäre die Reduktion der Gewichte, z.B. durch Streichen schwacher Gewichte, die wenig zur Dignitätsfindung beitragen. Dies führte in der vorliegenden Studie aber zu keiner wesentlichen Verbesserung. Gute prospektive Ergebnisse lieferte das Training anhand der in die Entscheidungsmatrix eingeflossenen Kriterien (Vaskularisation, Begrenzung, Mittelwert PI gesamt, Mittelwert V_{max} gesamt, retrotumorale Schallauslöschung), das zwar beim Lernen einen etwas höheren Fehlerwert ergab, beim Testen aber nur wenige Fälle falsch erkannte. Da die im Entschei-
dungsbaum verwerteten Kriterien nur sieben Gewichte ergaben, ist hier ein Verhältnis von 10:1 gewahrt.

Noch bessere Resultate könnten sicherlich mit einer höheren Fallzahl erzielt werden. Als Eingangsdatensätze könnte z.B. eine Kombination aus sonographischen Parametern mit qualitativen Gefäßparametern, Resistance- und Pulsatilitätsindizes und Maximalgeschwindigkeiten verwendet werden. Höhere Fallzahlen würden es zudem erlauben, ein Netz mit mehreren, auch verborgenen, Schichten zu verwenden. Hierbei werden noch weitaus größere Patientenzahlen benötigt, da sich mit dem Einfügen neuer Neuronen und ihrer Verbindung untereinander die Zahl der Gewichte erhöht. Bei Netzen mit verborgenen Schichten sind auch nicht-lineare Diskriminierungen möglich. Ein Netz ohne Zwischenschichten kann dagegen nur lineare Grenzen ziehen, legt also etwa im zweidimensionalen Raum Geraden zwischen zu Urterscheidende Punkte. Dies kann bei mehreren Punkten mit unterschiedlichen Ergebnissen zu Problemen führen, z.B. wenn A in dieselbe Gruppe wie C gehört, auf der Wertegeraden zwischen beiden Punkten aber das zu einer anderen Gruppe gehörende B liegt. Ein Netz mit mehreren Schichten kann insofern der multiplen logistischen Regression deutlich überlegen sein, die nur dichotome Entscheidungen zulässt.

Auch das statistische Verfahren der multiplen logistischen Regression erzielte beim prospektiven Patientengut schlechtere Treffsicherheiten als in der retrospektiven Patientengruppe. Für eine höhere Zuverlässigkeit der Ergebnisse würden auch hier größere Fallzahlen in den Untergruppen benötigt. Vor allem die Zahl benigner Tumoren der retrospektiven Gruppe war mit insgesamt 28 Fällen sowohl für das Erstellen des Entscheidungsbaumes als auch für das Training der Netze sehr niedrig. Dies bedingt die weit deutlichere Verschlechterung der Spezifität: bei der Entscheidungsmatrix um 24 Prozentpunkte auf nur noch 65 %, bei den neuronalen Netzen um 24 bis 44 Prozentpunkte auf 58 % bis maximal 69 %. Bei den als prospektiv klassifizierten Patientinnen waren zwei Tumoren zwar vaskularisiert, es konnte jedoch keine Dopplerflusskurve abgeleitet werden, so dass diese beiden Fälle mit der vorliegenden Entscheidungsmatrix nicht beurteilt werden konnten.

Die Entscheidungsmatrix wird anhand der optimalen miteinander kombinierbaren Kriterien des jeweiligen Kollektivs erstellt, er bedeutet also die beste Lösung für das retrospektive Kollektiv. Durch die verschobene Gewichtung der Einzelkriterien in dem neuem Kollektiv, also bei den prospektiven Patientinnen, verschlechterte sich auch das Resultat des Entscheidungsbaumes. Besonders die Abnahme der Spezifität fiel ins Gewicht. Vor allem die schlechte Reproduzierbarkeit der Spezifität bei dem Einzelkriterium "Begrenzung", aber auch die Verschlechterung der Spezifität des Einzelkriteriums "PI" verringerten die Gesamt-Spezifität.

Die Erstellung der Entscheidungsmatrix müsste erneut mit höheren Fallzahlen, zum Beispiel an dem um die prospektive Gruppe vergrößerten Kollektiv oder idealerweise an einer noch größeren Patientengruppe, erfolgen. Tendenziell zeigt sich jedoch, dass die Kombination der Einzelkriterien mit der multiplen logistischen Regression einen erheblichen Zugewinn an Treffsicherheit bringen kann. Wichtig ist hierbei das Einbeziehen sowohl sonomorphologischer als auch dopplersonographischer Kriterien, da eine Matrix allein mit sonographischen oder allein mit dopplersonographischen Kriterien nicht alle wesentlichen Informationen beinhalten kann.

Insgesamt zeigte die Auswertung mit neuronalen Netzen eine leichte Überlegenheit gegenüber der multiplen logistischen Regression, die sich durch das Einbeziehen großer Fallzahlen noch erhöhen kann. Die von der Statistik geforderte Exaktheit kann mit neuronalen Netzen gelockert werden. Bei einer großen Fülle von Eingabeparametern, wie in dieser Arbeit untersucht, bieten die Netze sicherlich Vorteile, da alle diese Informationen in die Auswertung mit einbezogen werden können. Limitierend für die Anwendung ist die schwierige Nachvollziehbarkeit der Ergebnisse: Für die juristische Absicherung von Entscheidungen, z.B. ob eine Operation erfolgen soll oder nicht, müssen die Rechenwege bis in jedes Detail nachprüfbar sein, was bei der künstlichen Intelligenz nicht der Fall ist.

Die Qualität der Untersuchungen und die Ergebnisse der Einzelkriterien beschränken sowohl die Anwendung der neuronalen Netze als auch die der multiplen logistischen Regression sehr stark: je besser die Untersuchung und je schlüssiger die Werte der Einzelkriterien, umso treffsicherer werden die Netze oder die Statistik entscheiden. Zum Teil ist aber auch der Tumor selbst der limitierende Faktor, z.B. unterscheiden sich einige wenige maligne Tumoren vom tumorbiologischen und morphologischen Aspekt nicht von benignen Prozessen. Ein maligner Tumor in dem prospektiven Patientengut zeigte z.B. exakt die gleichen sonographischen Kriterien wie ein benigner Tumor. Liegt eine intratumorale Vaskularisation vor und sind Dopplerflusskurven ableitbar, ist diese Koinzidenz aufgrund der Datenfülle sehr unwahrscheinlich. Bei einer großen Patientenzahl werden sich die neuronalen Netze hier als überlegen erweisen, da sie alle Daten mit einbeziehen können.

Die Dignitätseinstufung von Tumoren mit Hilfe der künstlichen Intelligenz ist zwar zurzeit, besonders für Entscheidungen mit möglichen juristischen Folgen, noch nicht in vollem Umfang einsetzbar. Dies gilt jedoch auch für die Auswertung von Sonographieund Dopplerergebnissen mit statistischen Methoden und die Beurteilung anhand von Score-Systemen oder Entscheidungsbäumen. Letztlich entscheiden die klinische Erfahrung sowie das Einbeziehen zahlreicher Untersuchungsmethoden über die endgültige diagnostische Einstufung. Eine Entscheidungshilfe können derzeit jedoch sowohl statistische Methoden als auch die künstliche Intelligenz bieten. Für die Zukunft bietet sich wohl eher die künstliche Intelligenz an. Auf verschiedenen Gebieten wird die künstliche Intelligenz die herkömmlichen statistischen Methoden in absehbarer Zeit verdrängen, z.B. bei der Krankheitsfindung anhand von Symptomenkatalogen. In vielen Bereichen der Medizin wird sie bereits jetzt schon mit Erfolg eingesetzt [6], [21], [36] (S. 205), [46] (S. 512), [84] (S. 159). In Verbindung mit dem Ultraschall ist besonders ein Anwendungsbereich interessant, der sich zurzeit in nicht-medizinischen Bereichen etabliert: Die Bilderkennung. In weiten Bereichen unabhängig vom Untersucher, der nur noch die Bilder und Dopplerflusskurven ableitet, ohne ein zeitaufwendiges und mit Fehlern behaftetes Einschätzen und Übertragen von Daten, könnten das erstellte Bild idealerweise dreidimensional - und die zugehörigen Dopplerflusskurven an den Computer weitergeleitet und in die Netze eingespeist werden. Diese könnten sofort das Ergebnis ermitteln. Solche Vorstellungen sind zwar momentan noch nicht in die Realität umsetzbar. Die Verwendung neuronaler Netze dagegen sollte - und könnte - bald an allen größeren Kliniken Standard sein.

6 ZUSAMMENFASSUNG

Neoangiogenetische Faktoren vermitteln das Einsprossen von neuen Blutgefäßen. Dies ist eine wesentliche Voraussetzung für Wachstum und Ausbreitung maligner Tumoren. Eine erhöhte Vaskularisation eines Tumors weist daher meist auf ein malignes Geschehen hin. Mit hoch auflösenden Dopplergeräten gelingt die Darstellung von Blutgefäßen auch im Flussgeschwindigkeiten kleiner Gefäßdurchmesser. Bereich niedriger und Die Dopplersonographie ist deshalb als Ergänzung zur konventionellen B-Mode-Sonographie Dignitätsdiagnostik von Mammaund Ovarialkarzinomen mittlerweile in der unverzichtbar. Die in der Literatur genannten Sensitivitäten und Spezifitäten der Sonographie und der farbkodierten und gepulsten Dopplersonographie unterscheiden sich zum Teil jedoch erheblich. Die zum dopplersonographischen Examen herangezogenen qualitativen und quantitativen Parameter variieren zudem sehr stark, was einen Vergleich der einzelnen Ergebnisse und eine einheitliche Dignitätsbeurteilung erschwert.

Ziel dieser Studie war es daher, mit geeigneten statistischen Verfahren die relevanten Parameter für eine Kombination der Sonographie und der Farbdoppler-Sonographie bei Ovarial- und Mammatumoren herauszufinden und den Zugewinn durch einen additiven Einsatz der Dopplersonographie zu überprüfen.

Zur Auswertung wurde einerseits ein Modell der Kombination beider Methoden mit Hilfe spezieller statistischer Verfahren entworfen. Mit der Entwicklung einer diagnostischen Entscheidungsmatrix konnten die relevanten Kriterien beider Verfahren ermittelt und gewichtet und damit die Sensitivitäten und Spezifitäten verbessert werden. Weiterhin wurde anhand der Mammatumoren ein Vergleich zwischen klassischen statistischen Verfahren und einer Methode der künstlichen Intelligenz, den neuronalen Netzen, durchgeführt. Es sollte untersucht werden, ob dadurch ein Zugewinn an diagnostischer Sicherheit zu erzielen ist. In einem experimentellen Modell wurden außerdem mögliche Fehlerquellen der Durchführung sonographischer und dopplersonographischer Untersuchungen aufgezeigt.

Bei der Analyse von 63 Adnex-Tumoren (histologisch 41 benigne, 22 maligne Befunde) konnte durch eine Kombination von Sonographie und Dopplersonographie in einer Entscheidungsmatrix eine signifikante Steigerung der sonographischen Treffsicherheit erreicht werden. In der Gruppe der histologisch malignen wie auch der benignen Tumoren wurden jeweils nur noch drei Patientinnen falsch eingestuft, die Treffsicherheit stieg damit um 12 % auf 90 % (Sensitivität 86 %, Spezifität 93 %, p < 0,01). Die relevante Zusatzinformation der Dopplersonographie beruhte dabei nicht auf der komplexen Analyse von Flussmustern, sondern auf dem Nachweis von Vaskularisation in soliden Tumoranteilen (*zentrale Vaskularisation*). Die isolierte sonomorphologische Beurteilung erreichte mit ihrem sensibelsten Parameter, dem Vorhandensein eines *soliden Anteils*, eine Sensitivität von 95 % und eine Spezifität von 68 % (Validität 78 %). Der RI und der PI aller Gefäße wiesen 80 % Sensitivität und 67 % bzw. 63 % Spezifität auf. Die besten Grenzwerte fanden sich bei 1,27 für den PI und bei 0,63 für den RI. Die durch den additiven Einsatz der Dopplersonographie und die Kombination beider Methoden deutlich verbesserte Dignitätsbeurteilung kann künftig die klinische Entscheidung, ob Ovarialtumoren konservativ beobachtet oder laparoskopisch bzw. laparotomisch abgeklärt werden sollen, erleichtern.

Bei 89 Brustbefunden (histologisch 59 maligne, 30 benigne Tumoren) ergaben sowohl die Kombination der Parameter mit Hilfe der multiplen logistischen Regression als auch die Auswertung mit neuronalen Netzen eine Verbesserung gegenüber den Einzelkriterien: Mit Hilfe der Erscheinungsmatrix wurden in der als retrospektiv klassifizierten Gruppe nur noch ein maligner und drei benigne Tumoren falsch eingeschätzt (98 % Sensitivität, 89 % Spezifität, 95 % Validität). Mit neuronalen Netzen konnten in der retrospektiven Gruppe alle Tumoren richtig eingestuft werden (100 % Validität). Das neuronale Netz wurde auf verschiedene Arten trainiert, z.B. anhand aller dopplersonographischen und sonographischen Parameter. In einem prospektiven Kollektiv (n = 130) von Patientinnen mit Mammatumor lagen die Vorhersagewerte bei Anwendung der künstlichen Intelligenz mit einer Validität von 88 %, einer Sensitivität von 96 % und einer Spezifität von 67 % etwas besser als bei der statistischen Analyse mittels multipler logistischer Regression (Validität 87 %, Sensitivität 95 %, Spezifität 65 %).

Durch das experimentelle Strömungsmodell ließ sich zeigen, dass die Ergebnisse der qualitativen und quantitativen dopplersonographischen Untersuchungen durch die Geräteeinstellung stark beeinflusst werden können. Vor allem Fehler bei der Winkelbestimmung und eine nicht gefäßdeckende Größe und Position des Echofensters fallen bei der dopplersonographischen Bestimmung der Blutflussgeschwindigkeiten ins Gewicht. Der Hämatokritwert besitzt keinen Einfluss auf die Genauigkeit oder Sensitivität dopplersonographischer Flussmessungen. Die Bestimmung des Gefäßdurchmessers mit der farbkodierten Duplexsonographie ist aufgrund unterschiedlicher Einflussfaktoren nicht praktikabel. Mit dem Powerdoppler (CDE-Modus) gelingt eine deutlich sensitivere Erfassung von langsamen Flüssen gegenüber dem farbkodierten Doppler (CDI-Modus).

Der kombinierte Einsatz der Sonographie und der Farbdoppler-Sonographie erreichte in dieser Studie eine hohe Validität zur präoperativen Evaluation von Ovarial- und Mammatumoren. Die Kombination und Gewichtung der statistisch signifikanten Kriterien mittels multipler logistischer Regression und der Entwurf einer Entscheidungsmatrix steigern dabei deutlich die diagnostische Treffsicherheit. Zukunftsweisend ist die Anwendung neuronaler Netze zur Auswertung, deren prospektive Wertigkeit in dieser Arbeit der statistischen Analyse mittels multipler logistischer Regression leicht überlegen war.

7 LITERATUR

- Adler, D.D., Carson, P.L., Rubin, J.M., Ouinn-Reid, D.: Doppler ultrasound color flow imaging in the study of breast cancer; preliminary findings. Ultrasound Med. Biol. 16 (1990) 553-559.
- [2] Ahmad, S., Tesauro, G.: Scaling and generalization in neural networks: a case study. Proceedings of neural information and proc. systems conference (NIPS) 1 (1988) 160-176.
- [3] Alcazar, J.L., Jurado, M.: Using a logistic model to predict malignancy of adnexal masses based on menopausal status, ultrasound morphology, and color Doppler findings. Gynecol. Oncol. 69 (1998) 146-150.
- [4] Arenson, J., Allison, J.: New technology shows promise of improving the Doppler image. Medical imaging international. Sonderdruck Firma Acuson.
- [5] Baum, G.: Ultrasound mammography. Radiology 122 (1977) 199-205.
- [6] Baxt, W.G.: Application of artificial neural networks to clinical medicine. Lancet 346 (1995) 1135-1138.
- [7] Bernaschek, G., Deutinger, J., Kratochwil, A.: Endosonography in Obstetrics and Gynecology. Springer-Verlag; Berlin, Heidelberg, 1990, 164-182.
- [8] Blohmer, J.U., Oellinger, H., Schmidt, C., Lichtenegger, W.: Präoperative Prognoseeinschätzung von Patientinnen mit einem Mammakarzinom durch die farbkodierte Dopplersonographie. Geburtsh. Frauenheilk. 58 (1998) 6-12.
- [9] Boring, C.C., Squires, T.S., Tong, T.: Cancer statistics. CA Cancer J. Clin. 44 (1994) 7-26.
- [10] Bourne, T., Campbell, S., Steer, C., Whitehead, M., Collins, W.: Transvaginal coulor flow imaging: a possible new screening technique for ovarian cancer. BMJ 299 (1989) 1367-1370.
- [11] Bourne, T., Whitehead, M., Campbell, S., Royston, P., Bhan, V., Collins, W.: Ultrasound screening for familial ovarian cancer. Gynecol. Oncol. 43 (1991) 92-97.
- [12] Bromley, B., Goodman, H., Benacerraf, B.: Comparison between sonographic morphology and Doppler waveform for the diagnosis of ovarian malignancy. Obstet. Gynecol. 83 (1994) 434-437.
- [13] Brown, D.L., Frates, M.C., Laing, F.C., DiSalvo, D.N., Doubilet, P.M., Benson, C.B., Waitzkin, E.D., Muto, M.G.: Ovarian Masses: can benign and malignant lesions be differentiated with color and pulsed Doppler US? Radiology 190 (1994) 333-336.
- [14] Campbell, S., Bhan, V., Royston, P., Whitehead, M., Collins, P.: Transabdominal ultrasound screening for early ovarian cancer. BMJ 299 (1989) 1363-1367.
- [15] Campbell, S., Bourne, T.H., Collins, W.P.: Detection of early ovarian cancer. Lancet 336 (1990) 436.

- [16] Campbell, S., Royston, P., Bhan, V., Whitehead, M.I., Collins, W.P.: Novel screening strategies for early ovarian cancer by transabdominal ultrasonography. Br J Obstet. Gynecol. 97 (1990) 304-311.
- [17] Carter, J., Saltzman, A., Hartenbach, E., Fowler, J., Carson, L., Twiggs, L.B.: Flow characteristics in benign and malignant gynecologic tumors using transvaginal color flow Doppler. Obstet. Gynecol. 83 (1994) 125-130.
- [18] Carter, J.R., Lau, M., Fowler, J.M., Carlson, J.W., Carson, L.F., Twiggs, L.B.: Blood flow characteristics of ovarian tumors: Implications for ovarian cancer screening. Am. J. Obstet. Gynecol. 172 (1995) 901-907.
- [19] Cosgrove, D.O., Bamber, J.C., Davey, J.B., McKinna, J.A., Sinnett, H.D.: Color Doppler signals from breast tumors. Radiology 176 (1990) 175-180.
- [20] Creasman, W., DiSaia, P.: Screening in ovarian cancer. Am. J. Obstet. Gynecol. 165 (1991) 7-10.
- [21] DaPonte, J.S., Sherman, P.: Classification of ultrasonic image texture by statistical discrimination analysis of neural networks. Computer. Med. Imaging Graph. 15 (1991) 3-9.
- [22] DePriest, P. D., Varner, E., Powell, J., Fried, A., Puls, L., Higgins, R., Shenson, D., Kryscio, R., Hunter, J. E., Andrews, S. J., van Nagell, J. R.: The efficacy of a sonographic morphology index in identyfying ovarian cancer: a multiinstitutional investigation. Gynecol. Oncol. 55 (1994) 174-178.
- [23] DePriest, P.D., Nagell, J.R. van, Gallion, H.H., Shenson, D., Hunter, J.E., Andrews, S.J., Powell, D.E., Pavlik, E.J.: Ovarian cancer screening in asymptomatic postmenopausal women. Gynecol. Oncol. 51 (1993) 205-209.
- [24] Dixon, J.M., Walsh, J., Patterson, D., Chetty, U.: Colour Doppler ultrasonography studies of benign and malignant breast lesions. Br. J. Surg. 79 (1992) 259-260.
- [25] Droegemueller, W.: Screening for ovarian carcinoma: hopeful and wishful thinking. Am. J. Obstet. Gynecol. 170 (1994) 1095-1098.
- [26] Faust, J: Die Doppler-Sonographie in der Gefäßdiagnostik. Kranzbühler; Solingen, 1992.
- [27] Fleischer, A.C., Cullinan, J.A., Peery, Ch.V., Jones, H.W.: Early detection of ovarian carcinoma with transvaginal color Doppler ultrasonography. Am. J. Obstet. Gynecol. 174 (1996) 101-106.
- [28] Fleischer, A.C., Everette James, A., Millis, J.B., Julian, C.: Differential diagnosis of pelvic masses by gray scale sonography. Am. J. Roentgenol. 131 (1978) 469-476.
- [29] Fleischer, A.C., McKee, M.S., Gordon, A.N.: Transvaginal sonography of postmenopausal ovaries with pathologic correlation. J. Ultrasound. Med. 9 (1990) 637-644.
- [30] Fleischer, A.C., Rodgers, W.H, Kepple, D.M., Williams, L.L., Jones, H.W.: Color Doppler sonography of ovarian masses: a multiparameter analysis. J. Ultrasound Med. 12 (1993) 41-48.
- [31] Fleischer, A.C., Rodgers, W.H., Rao, B.K.: Assessment of ovarian tumor vascularity with transvaginal color Doppler sonography. J. Ultrasound Med. 10 (1991) 563-568.

- [32] Folkman, J., Klagsburn, M.: Angiogenic factors. Science 235 (1987) 442-447.
- [33] Folkman, J., Merler, E., Abernathy, C., Williams, GH.: Isolation of a tumor factor responsible for angiogenesis. J. Exp. Med. 133 (1971) 275-288.
- [34] Franchi, M., Beretta, P., Ghezzi, F., Zanaboni, F., Goddi, F., Salvatore, S.: Diagnosis of pelvic masses with transabdominal color Doppler, CA 125 and ultrasonography. Acta Obstet. Gynecol. Scand. 74 (1995) 734-739.
- [35] Fu, L.M.: Neural networks in computer intelligence. McGraw-Hill, New York, 1994.
- [36] Gebbinck, M.S., Verhoeven, J.T.M., Thijssen, J.M., Schouten, T.E.: Application of neural networks for the classification of diffuse liver disease by quantitative echography. Ultrasonic Imaging 15 (1993) 205-217.
- [37] Gerthsen, Ch., Kneser, H.O., Vogel, H.: Physik. Springer-Verlag; Berlin, Heidelberg, New York, Tokyo, 1986, 95-129 u. 149-192.
- [38] Goswamy, R.K., Campbell, S., Royston, J.P., Bhan, V., Battersby, R.H., Hall, V.J., Whitehead, M.I., Collins, W.P.: Ovarian size in postmenopausal women. Br. J. Obstet. Gynecol. 95 (1988) 795-801.
- [39] Guerriero, S., Ajossa, S., Risalvato, A., Lai, M.P., Mais, V., Angiolucci, M., Melis, G.B.: Diagnosis of adnexal malignancies by using color Doppler energy imaging as a secondary test in persistent masses. Ultrasound Obstet. Gynecol. 11 (1998) 277-282.
- [40] Hackelöer, B.J., Duda, V., Lauth, G.: Möglichkeiten der Ultraschalldiagnostik. Gynäkologe 22 (1989) 238-241.
- [41] Hamper, U., Sheth, S., Abbas, F., Rosenshein, N., Aronson, D., Kurman, R.: Transvaginal color Doppler sonography of adnexal masses: Differences in blood flow impedance in benign and malignant lesions. AJR 160 (1993) 1225-1228.
- [42] Hata, K., Akiba, S., Hata, T., Miyazaki, K.: A multivariate logistic regression analysis in predicting malignancy for patients with ovarian tumors. Gynecol. Oncol. 68 (1998) 256-262.
- [43] Hata K., Hata T., Manabe A., Sugimura, K., Kitao, M.: A critical evaluation of transvaginal Doppler studies, transvaginal sonography, magnetic resonance imaging, and CA 125 in detecting ovarian cancer. Obstet. Gynecol. 80 (1992) 922-926.
- [44] Hata, K., Collins, W.: Color Doppler imaging and ovarian tumor angiogenesis: the Janus approach. Ultrasound Obstet. Gynecol. 9 (1997) 297-299.
- [45] Higgins, R., van Nagell, J., Donaldson, E. : Transvaginal Sonography as a screening method for ovarian cancer. Gynecol. Oncol. 34 (1989) 402-406.
- [46] Itchhaporia, D., Snow, P.B., Almassy, R.J., Oetgen, W.J.: Artificial neural networks: current status in cardiovascular medicine. J. Am. Coll. Cardiol. 28 (1996) 512-521.
- [47] Jackson, V.P.: Duplex sonography of the breast. Ultrasound Med. Biol. 14-Suppl. (1988) 1131-1137.
- [48] Jacobs, I., Oram, D., Fairbanks, J., Turner, J., Frost, C., Grudzinskas, J.G.: A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the

accurate preoperative diagnosis of ovarian cancer. Br. J. Obstet. Gynecol. 97 (1990) 922-929.

- [49] Jain, K.A.: Prospective evaluation of adnexal masses with endovaginal gray-scale and duplex and color Doppler US: correlation with pathologic findings. Radiology 191 (1994) 63-67.
- [50] Jellins, J., Kossoff, G., Reeve, T.S., Barraclough, B.H.: Ultrasonic gray-scale visualisation of breast disease. Ultrasound Med. Biol. 1 (1975) 393-404.
- [51] Kaiser, W.A., Diedrich, K., Reiser, M., Krebs, D.: Moderne Diagnostik der Mamma. Geburtsh. u. Frauenheilk. 53 (1993) 1-14.
- [52] Karlan, B., Platt, L.: The current status of ultrasound and color Doppler imaging in screening for ovarian cancer. Gynecol. Oncol. 55 (1994) S28-S33.
- [53] Karlan, B., Raffael, L.J., Crvenkovic, G., Smrt, C., Chen, M.D., Lopez, E., Walla, C., Garber, C., Cane, P., Sarti, D., Rotter, J., Platt, L.D.: A multidisciplinary approach to early detection of ovarian carcinoma: rationale, protocol design, and early resultats. Am. J. Obstet. Gynecol. 169 (1993) 494-501.
- [54] Kato, K., Callies, R.: Diagnostische Sicherheit und Grenzen der Mammasonographie. Geburtsh. u. Frauenheilk. 58 (1998) 1-5.
- [55] Kawai, M., Kano, T., Kikkawa, F., Maeda, O., Oguchi, H., Tomoda, Y.: Transvaginal Doppler ultrasound with color flow imaging in the diagnosis of ovarian cancer. Obstet. Gynecol. 19 (1992) 163-167.
- [56] Kawai, M., Kikkawa, F., Ishikawa, H., Tamakoshi, K., Maeda, O., Hasegawa, N., Mizuno, K., Suzuki, A., Itakura, A., Nakashima, K., Tomoda, Y.: Differential diagnosis of ovarian tumors by transvaginal color-pulse Doppler sonography. Gynecol. Oncol. 54 (1994) 209-214.
- [57] Klews, P.-M.: Einführung in die farbkodierte Duplexsonographie (FKDS). In: "Farbkodierte Duplexsonographie", Wolf, K.-J., Fobbe, F., Thieme-Verlag; Stuttgart, New York, 1993, 1-13.
- [58] Kobayashi, T.: Review: ultrasonic diagnosis of breast cancer. Ultrasound Med. Biol. 1 (1975) 383-391.
- [59] Kopans, D., Meyer, J., Sadowsky, N.: Breast imaging. N. Engl. J. Med. 310 (1984) 960-966.
- [60] Kossof, G., Garrett, W.J., Carpenter, D.A., Jellins, J., Dadd, M.J.: Classification of soft tissues by gray-scale echography. Ultrasound Med. Biol. 2 (1976) 89-111.
- [61] Kurjak, A., Predanic, M., Kupesic-Urek, S., Jukic, S: Transvaginal color and pulsed Doppler assessment of adnexal tumor vascularity. Gynecol. Oncol. 50 (1993) 3-9.
- [62] Kurjak, A., Schulman, H., Sosic, A., Zalud, I., Shalan, H.: Transvaginal ultrasound, color flow, and Doppler waveform of the postmenopausal adnexal mass. Obstet. Gynecol. 80 (1992) 917-921.
- [63] Kurjak, A., Zalud, I., Jurkovic, D., Alfirevic, Z., Miljan, M.: Transvaginal color Doppler for the assessment of pelvic circulation. Acta Obstet. Gynecol. Scand. 68 (1989) 131-135.

- [64] Labovitz A., Anthonis D., Cravens, T., Kern, M.: Validation of volumetric flow measurements by means of a Doppler-tipped corony angioplasty guide wire. Am. Heart J. 126 (1993) 1456-1461.
- [65] Laing, F.C.: US analysis of adnexal masses: The art of making the correct diagnosis. Radiology 191 (1994): 21-22.
- [66] Landwehr, P.: Hämodynamische Grundlagen. In: "Farbkodierte Duplexsonographie", Wolf, K.-J., Fobbe, F., Thieme-Verlag; Stuttgart, New York, 1993, 19-36.
- [67] Lee, S.K., Lee, T., Lee, K.R., Su, Y.G., Liu, T.J.: Evaluation of breast tumors with color Doppler imaging: a comparison with image-directed Doppler ultrasound. J. Clin. Ultrasound 23 (1995) 367-373.
- [68] Leeners, B., Schild, R.L., Funk, A., Hauptmann, S., Kemp, B., Schroder, W., Rath, W.: Colour Doppler sonography improves the pre-operative diagnosis of ovarian tumours made using conventional transvaginal sonography. Eur. J. Obstet. Gynecol. Reprod. Biol. 64 (1996) 79-85.
- [69] Lerner, J.P., Timor-Tritsch, I.E., Federman, A., Abramovich, G.: Transvaginal ultrasonographic characterization of ovarian masses with an improved, weighted scoring system. Am. J. Obstet. Gynecol. 170 (1994) 81-85.
- [70] Leucht, W.: Beurteilungskriterien sonographischer Herdbefunde. In: "Lehratlas der Mammasonographie", Leucht, W., Thieme-Verlag; Stuttgart, New York, 1989, 20-31.
- [71] Lunt, M.J.: The optimum frequency and angle of insonation for a Doppler ultrasonic flow velocity meter. Phys. Med. Biol. 25 (1980) 839-848.
- [72] Madjar, H., Prömpeler, H., Wolfahrt, R., Bauknecht, T., Pfleiderer, A.: Farbdopplerflussdaten von Mammatumoren. Ultraschall in Med. 15 (1994) 69-76.
- [73] Marquet, K.L., Funk, A., Fendel, H., Handt, S.: Der echodichte Randsaum und hyperreflexive Ausläufer: Sensible Kriterien maligner Prozesse in der Mammasonographie. Geburtsh. u. Frauenheilk. 53 (1993) 20-23.
- [74] McNicholas M. M. J., Mercer P. M., Miller J. C. et al: Color Doppler Sonography in the Evaluation of Palpable Breast Masses. AJR 161 (1993) 765-771.
- [75] Mehrotra, K.G., Chilukuri, K.M., Ranka, S.: Bounds on the number of samples needed for neural learning. IEEE Trans. Neural Networks 2 (1991) 548-558.
- [76] Meire H. B., Farrant P., Guha T.: Distinction of benign from malignant ovarian cysts by ultrasound. Br. J. Obstet. Gynecol. 85 (1978) 893-899.
- [77] Merz, E.: Physikalische Grundlagen und technische Ultraschallverfahren. In: "Sonographische Diagnostik in Gynäkologie und Geburtshilfe", Merz, E., Thieme-Verlag; Stuttgart, New York, 1988, 4-11.
- [78] Miyazaki, T., Tanaka, M., Miyakoshi, K., Minegishi, K., Kasai, K., Yoshimura, Y.: Power and colour Doppler ultrasonography for the evaluation of the vasculature of the human corpus luteum. Hum. Reprod. 13 (1998) 2836-2841.
- [79] Morley, P., Barnett, E.: The use of ultrasound in the diagnosis of pelvic masses. Br. J. Radiol. (1970) 602-616.

- [80] Nagell, J.R. van, DePriest, P.D., Puls, L.E., Donaldson, E.S., Gallion, H.H., Pavlik, E.J., Powell, D.E., Kryscio, E.J.: Ovarian cancer screening in asymptomatic postmenopausal women by transvaginal sonography. Cancer 68 (1991) 458-462.
- [81] Nagell, J.R. van, Higgins, R.V., Donaldson, E.S., Gallion, H.H., Powell, D.E., Pavlik, E.J., Woods, C.H., Thompson, E.A.: Transvaginal Sonography as a screening method for ovarian cancer. Cancer 65 (1990) 573-577.
- [82] Newey, V.R.: Classical versus artificial neural network analysis. Ultrasound Obtet. Gynecol. 10 (1997) 5-8.
- [83] Possover M., Morawski A., Müller E., Hettenbach, A.: Neue Perspektiven des Farb-Ultraschalls in der Mamma-Diagnostik. Geburtsh. u. Frauenheilk. 54 (1994) 432-436.
- [84] Prater, J., Richard, W.D.: Segmenting ultrasound images of the prostate using neural networks. Ultrasonic Imaging 14 (1992) 159-185.
- [85] Prömpeler, H.J., Madjar, H., Sauerbrei, W., Lattermann, U., Pfleiderer, A.: Kombinierte Beurteilung von Ovarialtumoren mit transvaginaler B-Bild und Farbdoppler-Sonographie. Geburtsh. u. Frauenheilk. 56 (1996) 345-350.
- [86] Prömpeler, H.J., Madjar, H., Sauerbrei, W., Lattermann, U., Pfleiderer, A.: Quantitative flow measurements for classification of ovarian tumors by transvaginal color Doppler sonography in postmenopausal patients. Ultrasound Obstet. Gynecol. 4 (1994) 406-413.
- [87] Prömpeler, H.J., Madjar, H., Sauerbrei, W., Lattermann, U., Pfleiderer, A.:Diagnostic formula for the differentiation of adnexal tumors by transvaginal sonography. Obstet. Gynecol. 89 (1997) 428-433.
- [88] Ranke, C., Hendrickx, P., Roth, U., Brassel, F., Creutzig, A., Alexander, K.: Color and conventional image-directed Doppler ultrasonography: accuracy and sources of error in quantitative blood flow measurements. J. Clin. Ultrasound 20 (1992) 187-193.
- [89] Sahin-Akyar, G., Sumer, H.: Color Doppler ultrasound and spectral analysis of tumor vessels in the differential diagnosis of solid breast masses. Inv. Radiol. 31 (1996) 72-79.
- [90] Sassone, A.M., Timor-Tritsch, I., Artner, A., Westhoff, C., Warren, W.: Transvaginal sonographic characterization of ovarian disease: Evaluation of a new scoring system to predict ovarian malignancy. Obstet. Gynecol. 78 (1991) 70-76.
- [91] Schelling, M., Gnirs, J., Braun, M., Busch, R., Maurer, S., Kuhn, W., Schneider, K.T.M., Graeff, H.: Optimized differential diagnosis of breast lesions by combined B-mode and color Doppler sonography. Ultrasound Obstet. Gynecol. 10 (1997) 48-53.
- [92] Schneider, V., Schneider, A., Reed, K., Hatch, K.: Comparison of Doppler with two-dimensional sonography and CA 125 for prediction of malignancy of pelvic masses. Obstet. Gynecol. 81 (1993) 983-988.
- [93] Schünemann, H.: Gynäkologische Malignome. In: "Malignomtherapie", Beaufort, F., Schünemann, H. (Hrsg.), Zuckschwerdt-Verlag; München, 1987, 162-164.
- [94] Seitz, K., Kubale, R.: Duplexsonographie der abdominellen und retroperitonealen Gefäße. VCH Verlagsgesellschaft; Weinheim, 1988, 1-55.

- [95] Sohn, C.: Sonographische Durchblutungsdiagnostik in der Gynäkologie. In: "Ultraschall in Gynäkologie und Geburtshilfe", Sohn, C., Holzgreve, W. (Hrsg.), Thieme-Verlag; Stuttgart, 1995, 719-748.
- [96] Sohn, C., Grischke, M., Wallwiener, M., Kaufmann, M., Fournier, D. von, Bastert, G.: Die sonographische Durchblutungsdiagnostik gut- und bösartiger Brust-Tumoren. Geburtsh. u. Frauenheilk. 52 (1992) 397-403.
- [97] Sohn, C., Stolz, W., Kaufmann, M., Bastert, G.: Die dreidimensionale Ultraschalldarstellung benigner und maligner BrustTumoren erste klinische Erfahrungen. Geburtsh. u. Frauenheilk. 52 (1992) 520-525.
- [98] Tailor, A., Jurkovic, D., Bourne, T.H., Collins, W.P., Campbell, S.:Sonographic prediction of malignancy in adnexal masses using an artificial neural network. Br. J. Obstet. Gynecol. 106 (1999) 21-30.
- [99] Tailor, A., Jurkovic, D., Bourne, T.H., Collins, W.P., Campbell, S.:Sonographic prediction of malignancy in adnexal masses using multivariate logistic regression analysis. Ultrasound Obstet. Gynecol. 10 (1997) 41-47.
- [100] Tekay, A., Jarvela, I., Jouppila, P.: Reproducibility of transvaginal Doppler velocimetry measurements in the uterine arteries of postmenopausal women. Ultrasound Obstet. Gynecol. 10 (1997) 198-204.
- [101] Tekay, A., Jouppila, P.: Blood flow in benign ovarian tumors and normal ovaries during the follicular phase. Obstet. Gynecol. 86 (1995) 55-59.
- [102] Timmermann, D., Schwarzler, P., Collins, W.P., Claerhout, F., Coenen, M., Amant, F., Vergote, I., Bourne, T.H.: Subjective assessment of adnexal masses with the use of ultrasonography: an analysis of interobserver variability and experience. Ultrasound Obstet. Gynecol. 13 (1999) 11-16.
- [103] Timmermann, D., Verrelst, H., Bourne, T.H., De Moor, B., Collins, W.P., Vergote, I., Vandewalle, J.: Artificial neural network models for the preoperative discrimination between malignant and benign adnexal masses. Ultrasound Obstet. Gynecol. 13 (1999) 17-25.
- [104] Timor-Tritsch, I., Bar-Yam, Y., Elgali, S., The technique of transvaginal sonography with the use of a 6,5 MHz probe. Am. J. Obstet. Gynecol. 158 (1988) 1019-1024.
- [105] Timor-Tritsch, I.E., Lerner, J.P., Monteagudo, A., Santos, R.: Transvaginal ultrasonographic characterization of ovarian masses by means of color flow-directed Doppler measurements and a morphologic scoring system. Am. J. Obstet. Gynecol. 168 (1993) 909-913.
- [106] Valentin, L., Sladkevicius, P., Marsal, K.: Limited contribution of Doppler velocimetry to the differential diagnosis of extrauterine pelvic tumors. Obstet. Gynecol. 83 (1994) 425-433.
- [107] Weiner, Z., Thaler, I., Beck, D., Rottem, S., Deutsch, M., Bramdes, J.M.: Differentiating malignant from benign ovarian tumors with transvaginal color flow imaging. Obstet. Gynecol. 79 (1992) 159-162.
- [108] Wells, P.N.T., Halliwell, M., Skidmore, R., Webb, A.J., Woodcock, J.P.: Tumour detection by ultrasonic Doppler blood-flow signals. Ultrasonics 15 (1977) 231-232.

- [109] Wu, C.C., Lee, C.N., Chen, T.M., Shyu, M.K., Hsieh, C.Y., Chen, H.Y., Hsieh, F.J.: Incremental angiogenesis assessed by color Doppler ultrasound in the tumorigenesis of ovarian neoplasms. Cancer 73 (1994) 1251-1256.
- [110] Yamashita, Y., Torashima, M., Hatanaka, Y., Harada, M., Higashida, Y., Takahashi, M., Mizutani, H., Tashiro, H., Iwamasa, J., Miyazaki, K., Okamura, H.: Adnexal masses: accuracy of characterization with transvaginal US and precontrast and postcontrast MR imaging. Radiology 194 (1995) 557-565.
- [111] Zell, A.: Simulation neuronaler Netze. Addison-Wesley, Bonn, 1994.
- [112] Zuna, I., Delorme, S., Huber, S., Danes, J., Teubner, J., Lorenz, A., Kaick, G.: Eine multizentrische Studie zu diagnostischen Kriterien in der Mammasonographie. Statistische Fallstricke und Wege aus dem Datendschungel. Radiologe 38 (1998) 355-363.

8 ABBILDUNGSVERZEICHNIS

Abb. 1: Fehlerfläche (Error) als Funktion der Gewichte w1 und w2 bei Backpropagation (nach [111], S. 105)
Abb. 2: Beispiel eines neuronalen Netzes mit zwei verdeckten Schichten (nach [111], Seite 73)
Abb. 3: Schematisiertes Beispiel eines einfachen neuronalen Netzes
Abb. 4: Geschwindigkeitsprofil einer laminaren Strömung. Die Länge der Pfeile entspricht der jeweiligen Geschwindigkeit (modifiziert nach [94] (S. 5) und [37] (S. 102)
Abb. 5: Diagnosebaum für die kombinierte sonographische und dopplersonographische Bewertung von Mammatumoren ($n = 89$). Die grauen Kästchen geben die Wahrscheinlichkeit für die Malignität eines Prozesses an
Abb. 6: Genauigkeit der donnlersonogranhischen Geschwindigkeitsmessung in acht
Versuchsreihen. Die durchgezogene Linie stellt die Mittelwerte, die grob gestrichelte Linie die errechneten Soll-Werte dar. Fein gestrichelt das 90-%-Konfidenzintervall. Gleiche
Werte der Einzelmessungen (Δ) sind sich überlagernd eingezeichnet
Abb. 7: Genauigkeit der dopplersonographischen Geschwindigkeitsmessung bei Winkelvariation . Durchgezogene Linie: Errechnete Soll-Geschwindigkeit bei 100 ml/h Pumpvolumen (0,07 m/s). o: Winkelkorrekturfunktion der Gerätes (Cosinus-Funktion). x: Dopplersonographische Einzelmessungen. Ergebnisse aus zehn Versuchsreihen
Abb. 8: Prozentualer Messfehler bei der dopplersonographischen Geschwindigkeitsmessung mit verschiedenen Winkeln, Mittelwerte aus zehn Messreihen. Negative Prozentzahlen: Abweichung der gemessenen Geschwindigkeit nach unten 42
Abb. 9: Dopplersonographische Geschwindigkeitsmessung bei unterschiedlich hoher Empfangsverstärkung . Messergebnisse aus je acht Versuchsreihen. Auf der x-Achse sind die Positionen des Empfangverstärkungs-Drehreglers aufgetragen. Gleiche Werte werden nur einfach aufgeführt. Durchgezogene Linie: Errechnete Sollgeschwindigkeit
Abb. 10: Mittlere prozentuale Abweichung der dopplersonographischen Geschwindigkeitsmessung bei unterschiedlich hoher Empfangsverstärkung (x-Achse: Position des Drehreglers). Negative Prozentzahlen: Abweichung der Geschwindigkeit nach unten
Abb. 11: Dopplersonographische Geschwindigkeitsmessung bei unterschiedlich gesetztem Dopplerfenster. Messergebnisse aus je acht Versuchsreihen bei Ableitung der Dopplerflusskurve mit je nur zentralem, nur randständigem und den ganzen Schlauch umfassendem Cursor. Gleiche Werte sind nicht doppelt aufgeführt. Durchgezogene Linie: Errechnete Sollgeschwindigkeit

Abb. 12: Tiefenabhängigkeit der dopplersonographischen Geschwindigkeitsmessung. Messergebnisse aus acht Versuchsreihen bei Ableitung der Dopplerflusskurve in 1 cm und in 4 cm Tiefe, gleiche Werte sind nicht doppelt aufgeführt. Durchgezogene Linie: Errechneter Sollwert
Abb. 13: Schlauchdurchmesser bei Farbduplex-Darstellung untereinander liegender Schläuche. In sechs Versuchsreihen wurde der Durchmesser des Farbstrahles gemessen, gleiche Werte sind nicht doppelt aufgeführt. Durchgezogene Linie: Tatsächlicher Innendurchmesser des Silikonschlauches
Abb. 14: Darstellung eines Blutflusses bei unterschiedlichen maximalen Geschwindigkeitsskalen (CDI-Modus). Schraffierte Fläche: Pumpvolumina, bei denen kein Blutfluss dargestellt werden konnte. Karierte Fläche: Punktuelles Aufleuchten des farbkodierten Dopplersignals. Ausgemalte Fläche: Rekonstruktion eines Gefäßverlaufes möglich. Darüber: Deutlicher Fluss im ganzen Schlauchlumen
Abb. 15: Grenze der Farbdarstellung bei flachen Winkeln und variierter maximaler Geschwindigkeitsskala (CDI-Modus). Karierte Fläche: Winkel, bei denen kein Blutfluss darstellbar war. Ausgemalte Fläche: Winkel, bei dem Blutfluss zur Darstellung kam49
Abb. 16: Grenze der Darstellung eines Blutflusses bei unterschiedlichen Dynamikbereichen der Energie (CDE-Modus). Schraffierte Fläche: Pumpvolumina, bei denen kein Blutfluss dargestellt werden konnte. Karierte Fläche: Punktuelles Aufleuchten des farbkodierten Dopplersignals. Ausgemalte Fläche: Rekonstruktion eines Gefäßverlaufes möglich. Darüber: Deutlicher Fluss im ganzen Schlauchlumen
Abb. 17: Relation der Breite des Farbspektrums im CDI-Modus zum tatsächlichen Innendurchmesser des Schlauches (durchgezogene Linie) in Abhängigkeit von der eingestellten Farbintensität und der maximalen Geschwindigkeitsskala. Auf der x- Achse sind die Positionen des Farbintensitäts-Drehreglers aufgetragen. Je vier Versuchsreihen bei einer Geschwindigkeitsskala von 0,09 m/s ($o = Einzelwerte, x =$ Mittelwert) und bei 0,023 m/s ($\Delta = Einzelwerte, + =$ Mittelwert)
Abb. 18: Relation der Breite des Farbspektrums im CDE-Modus zum tatsächlichen Innendurchmesser des Schlauches (durchgezogene Linie) in Abhängigkeit von der eingestellten Farbintensität. Werte und Mittelwerte aus je vier Versuchsreihen. Auf der x- Achse sind die Positionen des Farbintensitäts-Drehreglers aufgetragen
Abb. 19: ROC-Kurve für die Bestimmung der optimalen Grenzwerte des Volumens für die Dignitätsbeurteilung von Adnextumoren, $n = 61$. $o =$ Markierung der drei besten Grenzwerte.57
Abb. 20: Entscheidungsmatrix für die Anwendung der multiplen logistischen Regression zur sonographischen Dignitätsbeurteilung von Adnextumoren (n = 63). Die Prozentzahlen geben die Wahrscheinlichkeit für das Vorliegen von Malignität bei Erfüllung des vorgeschalteten Kriteriums an ($p < 0.001$)
Abb. 21: Nachweis eines Notch (spätsystolische Einkerbung) bei Gefäßen benigner und maligner Ovarialtumoren. 60

Abb. 22: ROC-Diagramm für den Vergleich der Sensitivitäten und Spezifitäten von V_{max}, V_{min}, V_{mean}, PI, RI und AB bei der Dignitätsbeurteilung von Adnextumoren. Mittelwerte Abb. 23: ROC-Diagramm für den Vergleich der Sensitivitäten und Spezifitäten verschiedener Werte des Pulsatilitäts-Index (PI) bei der Dignitätsbeurteilung von Adnextumoren. MW = Mittelwert aller Gefäße einer Patientin; Max = höchster Wert, Min = niedrigster Wert für den Pulsatilitäts-Index aller (zentraler und peripherer) Gefäße Abb. 24: ROC-Diagramm für den Vergleich der Sensitivitäten und Spezifitäten des Mittelwertes für den Resistance-Index (RI) zentraler, peripherer und aller Gefäße (gesamt) bei der Dignitätsbeurteilung von Adnextumoren. MW = Mittelwert verschiedener Abb. 25: Verteilung der systolischen Maximalgeschwindigkeit (V_{max}) bei benignen und malignen Adnextumoren. Δ = einzelne Mittelwerte der V_{max} aller Gefäße einer Abb. 26: Verteilung der diastolischen Maximalgeschwindigkeit (Vmin) bei benignen und malignen Adnextumoren. Δ = einzelne Mittelwerte der V_{min} aller Gefäße einer Abb. 27: Verteilung des Pulsatilitäts-Index (PI gesamt) bei benignen und malignen Adnextumoren. Δ = einzelne Mittelwerte des PI aller Gefäße einer Raumforderung. 67 Abb. 28: Verteilung des Resistance-Index (RI gesamt) bei benignen und malignen Adnextumoren. Δ = einzelne Mittelwerte des RI aller Gefäße einer Raumforderung. 67 Abb. 29: Diagnosematrix zur dopplersonographischen Dignitätsbeurteilung von Adnextumoren. Die Prozentzahlen geben die Wahrscheinlichkeit für das Vorliegen von Malignität eines Prozesses bei Erfüllung des jeweils vorgeschalteten Kriteriums an. 68 Abb. 30: Entscheidungsmatrix für eine kombinierte Anwendung der in der multiplen logistischen Regression ermittelten signifikanten sonographischen und qualitativen dopplersonographischen Parameter bei der Differenzierung unklarer Adnextumoren (Modell I). Die Prozentzahlen geben die Wahrscheinlichkeit für das Vorliegen von Malignität eines Prozesses bei Erfüllung des jeweils vorgeschalteten Kriteriums an. 70 **Abb. 31:** Entscheidungsmatrix für eine kombinierte Anwendung der in der multiplen logistischen Regression ermittelten signifikanten sonographischen sowie qualitativen und quantitativen dopplersonographischen Parameter bei der Differenzierung unklarer Adnextumoren (Modell II). Die Prozentzahlen geben die Wahrscheinlichkeit für das Vorliegen von Malignität eines Prozesses bei Erfüllung des jeweils vorgeschalteten

9 TABELLENVERZEICHNIS

Tab. 1 : Stadieneinteilung von Ovarialkarzinomen nach der Fédération Internationale deGynécologie et d'Obstétrique (FIGO), vereinfacht nach [93]	:6
Tab. 2: Histologie der Ovarialtumoren (n=63)	6
Tab. 3: Klassifikation der malignen Ovarialtumoren (n=21) 2	6
Tab. 4: Menopausenstatus der Patientinnen mit Ovarialtumor (n=63). 2	:7
Tab. 5 : Stadieneinteilung von Mammakarzinomen nach der TNM-Klassifikation,vereinfacht nach [93].3	2
Tab. 6: Histologie der Mammatumoren (n = 219)	2
Tab. 7: Klassifikation der malignen Mammatumoren (n = 153)	2
Tab. 8: Menopausenstatus der Patientinnen mit Mammatumor (n=217). Retrosp. = retrospektiv. 3	3
Tab. 9: Spezifitäten, Sensitivitäten und Treffsicherheiten der sonomorphologischenParameter, die in die Untersuchung eingingen. Neben der Fallzahl n (maximal 63) Angabedes jeweiligen Malignitätskriteriums (Malignitätskrit.) bzw. des entsprechendenGrenzwertes, n.s. = nicht signifikant.5	e 66
Tab. 10: Spezifitäten, Sensitivitäten und Treffsicherheiten der dopplersonographischenParameter (qualitative Kriterien und Mittelwerte der quantitativen Kriterien) bei derDignitätsbeurteilung von Adnex-Tumoren. Neben der Fallzahl n (maximal 63) Angabe derjeweiligen Malignitätskriteriums (Malignitätskrit.) bzw. des entsprechenden Grenzwertes.n.s. = nicht signifikant.	s
Tab. 11: Vergleich der diagnostischen Wertigkeit der Entscheidungsmatrix für dieSonographie, die Dopplersonographie und die Kombination aus Sonographie mit derDopplersonographie bei der Dignitätsbeurteilung von Adnex-Tumoren (n = 63). Modell I:nur qualitative dopplersonographische Kriterien, Modell II: qualitative und quantitativedopplersonographische Kriterien. V _{mean} = mittlere Maximalgeschwindigkeit, PI =Pulsatilitätsindex	'2
Tab. 12: Vergleich der mit der Entscheidungsmatrix erreichten Sensitivitäten, Spezifitäterund Validitäten bei retro- und prospektiven Datensätzen der Patientinnen mitMammatumor (n retrospektiv = 89, n prospektiv = 130).7	n '3
Tab. 13: Mittels neuronaler Netze erzielte Sensitivitäten, Spezifitäten und Validitäten bei der Auswertung der sonographischen und dopplersonographischen Dignitätsbeurteilung von Mammatumoren. VP = qualitative Vaskularisationsparameter. Retro = retrospektive Patientengruppe (n = 89), pro = prospektive Patientengruppe (n = 130)	'6

Tab. 14: Vergleich der Treffsicherheiten bei der Auswertung sonographischer und dopplersonographischer Untersuchungen von Mammatumoren ($n = 219$) mittels multiple	er
logistischer Regression (MLR) und neuronaler Netze (NN).	.79
Tab. 15: Prozentuale Abweichung dopplersonographischer Messungen der	
Blutflussgeschwindigkeit bei einem Winkelfehler von 5 Grad (nach [94], S. 23).	.84
Tab. 16: Vergleich verschiedener in der Literatur angegebener Grenzwerte für	
Pulsatilitäts- und Resistance-Index bei der dopplersonographischen Untersuchung von	
Adnex-Tumoren mit Sensitivitäten (Sens.) und Spezifitäten (Spez.). Zahl mäligner	0.4
1 umoren zur gesamten 1 umorzan1 = 2 mai. / n.	.94
Tab. 17: Vergleich der Sensitivitäten und Spezifitäten verschiedener Score-Systeme und	t
Modelle multipler logistischer Regression zur Beurteilung von Adnex-Tumoren.	
US = Ultraschall, FD = Farbdoppler, MLR = multiple logistische Regression. Zahl	00
mangner 1 umoren zur gesamten 1 umorzani = $2 \text{ mal.} / \text{n.}$.99

DANKSAGUNG

Herrn Professor Dr. KTM Schneider danke ich für die Bereitstellung des Themas sowie für die fruchtbaren Ideen und motivierenden Anregungen, mit denen er zum Entstehen der Arbeit maßgeblich beigetragen hat.

Mein besonderer Dank gilt Herrn Privatdozent Dr. Marcus Schelling für die fachkundige und unermüdliche Betreuung, die Beratung beim Durchführen der Versuche und seine Geduld nicht nur beim Einlernen in die klinische Diagnostik.

Bei Herrn Dipl.-Stat. M. Scholz und bei Frau Dipl.-Math. R. Busch möchte ich mich für die konstruktive Zusammenarbeit in allen statistischen Fragestellungen bedanken, durch die auch kompliziertere Problemstellungen elegant gelöst werden konnten.

Das Team des Instituts für Statik und Dynamik der Luft- und Raumfahrtkonstruktionen der Universität Stuttgart und insbesondere mein Mann Klaus haben mich bei der Anwendung der neuronalen Netze tatkräftig unterstützt, wofür ich mich ganz speziell bedanken möchte.

Danken möchte ich auch allen Mitarbeitern der Frauenklinik und des hämatologischen Labors des Klinikums rechts der Isar, die mir beim Durchführen dieser Arbeit behilflich waren. Außerdem sage ich all jenen Dank, die mich auf dem Weg bis zu ihrer Fertigstellung begleitet haben, besonders meiner Familie.

LEBENSLAUF

Name:	Brigitta Reglinde Reimann
Geburtsdatum:	5. Juni 1970
Geburtsort:	Ingolstadt

Ausbildung:

Schule:	1976-1980	Grundschule an der Prinzregentenstraße, Rosenheim
	1980-1989	Theresien-Gymnasium München
Studium:	1989-1991	Romanistik und Slawistik, Ludwig-Maximilian-
		Universität München
	1991-1993	Humanmedizin, vorklinischer Teil, Ludwig-
		Maximilian-Universität München
	1993-1997	Humanmedizin, klinischer Teil, Technische
		Universität München

Ärztliche Tätigkeit:

Arzt im Praktikum:	1997-1998	Innere Medizin, Kreiskrankenhaus Erding
	1999	Anästhesie, Olga-Hospital, Stuttgart
Assistenzärztin:	seit 2001	Anästhesie, Intensiv- und Notfallmedizin,
		Kreiskrankenhaus Sigmaringen