

Bestimmung des Knochenalters: Vergleich einer konventionellen Röntgenmethode mit einer neuentwickelten Ultraschallmethode

Dissertation
zur Erlangung des akademischen Grades
doctor medicinae (Dr. med.)

vorgelegt dem Rat der Medizinischen Fakultät der
Friedrich-Schiller-Universität Jena

von **Marcus Eulenstein**
geboren am 27. August 1980 in Gera

Erster Gutachter: Prof. Dr. med. Dipl.-Chem. W.A. Kaiser
Zweiter Gutachter: Prof. Dr. med. E. Kauf
Dritter Gutachter: Prof. Dr. med. W. Hirsch

Tag der öffentlichen Verteidigung: 08. Januar 2007

Verzeichnis der Abkürzungen

AGS	adrenogenitales Syndrom
B-Bild	Ultraschalldarstellung im „brightness mode“ (Helligkeit, Grauwertebild)
B20-Methode	Bestimmung des Knochenalters anhand von 20 Knochen der oberen Extremität, Teil der Tanner-Whitehouse-Methode
bm	mathematisches Formelzeichen: Bestimmtheitsmaß
bspw.	beispielsweise
bzw.	beziehungsweise
ca.	circa
CED	chronisch entzündliche Darmerkrankung
cGy	Zentigray
cm	Zentimeter
g	Gramm
KEVS	konstitutioneller Entwicklungsvorsprung
KEVZ	konstitutionelle Entwicklungsverzögerung
kg	Kilogramm
kHz	Kilohertz
kV	Kilovolt
LH	Luteinisierendes Hormon
LH-RH	Luteinisierendes Hormon-Releasing Hormon
M.	Morbus
mA	Milliampere
mm	Millimeter
p	mathematisches Formelzeichen: Testwert statistischer Tests
PKU	Phenylketonurie
r	mathematisches Formelzeichen: Korrelationskoeffizient
RUS-Methode	Bestimmung des Knochenalters anhand von Radius, Ulna und der Knochen des 1., 3. und 5. Fingerstrahls
s	mathematisches Formelzeichen: Standardabweichung
SMS	skeletal maturity score

sog.	sogenannte / sogenannten
SOS	speed of sound
SQV	System Quality Verification
STH	Somatotropes Hormon
T_3	Trijodthyronin
TW1-Methode	1962 von J.M. Tanner & R.H. Whitehouse publizierte Methode zur Knochenaltersbestimmung
TW2-Methode	1975 veröffentlichter Nachfolger der TW1-Methode
U	Untersucher
u.a.	unter anderem
v.a.	vor allem
VC	Variationskoeffizient
vs.	versus
\bar{x}	mathematisches Formelzeichen: arithmetischer Mittelwert
z.B.	zum Beispiel

Tabellenverzeichnis

1	Erkrankungen mit retardierter Skelettentwicklung	14
2	Erkrankungen mit akzelerierter Skelettentwicklung	15
3	Hormone und ihr Einfluss auf die Entwicklung des Skeletts	17
4	Vorteile und Nachteile der Sonographie in der Osteologie	32
5	Ein- bzw. Ausschlusskriterien für die Teilnahme an der Studie	36
6	Häufigkeit und Verteilung der Diagnosen	37
7	Alter und anthropometrische Charakteristika	37
8	Auswertepaare für die Knochenaltersbestimmung nach der Methode von Greulich & Pyle	45
9	Auswertepaare für den Vergleich von sonographischer und radiologischer Knochenaltersbestimmung	45
10	Variationskoeffizient für das Bonage-System	51
11	Individuelle Ergebnisse der Knochenaltersbestimmung der einzelnen Untersucher und des Bonage-Systems	53
12	Korrelationskoeffizienten und Bestimmtheitsmaße der Untersucherpaare	53
13	Streuung und Lage der Differenzen aller Untersucherpaare	58
14	Korrelationskoeffizienten und Bestimmtheitsmaße der Bonage-Paare . .	59
15	Streuung und Lage der Differenzen aller Bonage-Paare	64
16	aufgetretene Tendenzen aller Untersucherpaare	65
17	Interbeurteiler-Variabilität bei Anwendung der Greulich & Pyle-Methode in verschiedenen Studien	68
18	Interbeurteiler-Variabilität bei Anwendung der Greulich & Pyle-Methode in der vorliegenden Studie	69
19	Einfluss von Erkrankungen und Therapien auf die Bonage-Messungen .	79

Abbildungsverzeichnis

1	Handradiogramme männlicher Patienten	26
2	Das Bonage-Ultraschallsystem	40
3	Messzone am Handgelenk	41
4	Darstellung des Bonage-Systems während eines Messzyklus	43
5	Schema eines Bland-Altman-Plots	48
6	Schema eines Scatterplots	49
7	Boxplots der Differenzen der verschiedenen Untersucherpaare	54
8	Bland-Altman-Plots zweier Untersucherpaare	57
9	Boxplots der Differenzen der Bonage-Paare	60
10	Bland-Altman-Plots zweier Bonage-Paare	62
11	übereinstimmende und gegensätzliche Tendenzen des Paares Bonage/ U3	66
12	übereinstimmende und gegensätzliche Tendenzen des Paares U2/ U3 . .	66

Inhaltsverzeichnis

Verzeichnis der Abkürzungen	3
Tabellenverzeichnis	5
Abbildungsverzeichnis	6
Zusammenfassung	11
1 Einleitung	13
1.1 Definition von Skelettreife bzw. Knochenalter	13
1.2 Die Entwicklung des Skeletts	13
1.3 Die Skelettentwicklung beeinflussende Faktoren	14
1.4 Notwendigkeit der Knochenaltersbestimmung	17
1.5 Geschichte der Skelettreife- bzw. Knochenaltersbestimmung	19
1.6 Methodik der Knochenaltersbestimmung	20
1.7 Die röntgenologische Untersuchung der Hand	22
1.7.1 Bestimmung beim Neugeborenen und jungen Säugling	22
1.7.2 Bestimmung des Skelettalters beim älteren Säugling, Kind und Jugendlichen bis zur Adoleszenz	23
1.7.2.1 Methode nach Greulich & Pyle	23
1.7.2.2 Oxford-Methode	25
1.7.2.3 Methode nach Tanner & Whitehouse	27
1.7.2.4 Fels-Methode	28
1.7.3 Probleme bei der Anwendung konventioneller Methoden	28
1.7.3.1 Methode nach Greulich & Pyle	29
1.7.3.2 Vergleich von Greulich & Pyle mit Tanner & Whitehouse	31

1.8	Anwendung von Ultraschall bei Untersuchungen am Skelettsystem . . .	31
1.8.1	Knochenaltersbestimmung mittels sonographischer Methoden . .	33
2	Zielstellung	34
3	Patienten und Methoden	36
3.1	Patienten	36
3.2	Methoden	38
3.2.1	Untersucher	38
3.2.2	Struktur der Studie	38
3.2.3	Bestimmung des Knochenalters anhand der Methode nach Greulich & Pyle	39
3.2.3.1	Natives Röntgen der linken Hand	39
3.2.3.2	Gerätebeschreibung und Auswahl der Geräteparameter	39
3.2.4	Bestimmung des Knochenalters mit dem Bonage-Ultraschallsystem	39
3.2.4.1	Gerätebeschreibung	40
3.2.4.2	Spezifikationen des Bonage-Systems	41
3.2.4.3	Untersuchungsablauf	42
3.2.4.4	System-Qualitätsverifikation	43
3.2.5	Statistische Auswertung	44
3.2.5.1	Messgenauigkeit des Bonage-Systems	44
3.2.5.2	Bildung von Paaren	45
3.2.5.3	Korrelation und Bestimmtheitsmaß	46
3.2.5.4	Differenzanalyse und Verteilung	46
3.2.5.5	Statistische Methodik und Tests	47
4	Ergebnisse	51

4.1	Präzision des Bonage-Systems	51
4.2	Evaluierung der Methode von Greulich & Pyle	51
4.2.1	Ergebnisse der einzelnen Untersucher	52
4.2.2	Ergebnisse des Bonage-Systems	59
4.3	Vergleich der Greulich-Pyle-Methode mit Bonage	59
4.4	Analyse der Tendenzen	64
5	Diskussion	67
5.1	Interpretation der Ergebnisse der Methode nach Greulich & Pyle	67
5.1.1	Interbeurteiler-Variabilität der Greulich & Pyle-Methode	67
5.1.2	Ursachen für die Interbeurteiler-Variabilität	69
5.1.3	Ansätze zur Verringerung der Variabilität und Verbesserung der Interbeurteiler-Variabilität	72
5.2	Interpretation der Ergebnisse der Ultraschallmethode	73
5.2.1	Präzision des Bonage-Ultraschallsystems	73
5.2.2	Bonage-Messungen	74
5.3	Beurteilung des Vergleichs der Atlasmethode und der Ultraschallmethode	76
5.3.1	Erkrankungen mit Einfluss auf die Ultraschallmethode	77
5.3.2	Einfluss verschiedener Fehlerquellen auf die Messungen mit dem Bonage-System	79
5.3.3	Anwendung anderer ultraschallbasierter Methoden zur Knochen- altersbestimmung	82
6	Schlussfolgerungen	85
	Literaturverzeichnis	87
	Lebenslauf	92

Danksagungen	93
Ehrenwörtliche Erklärung	94

Zusammenfassung

Das objektive Kriterium für die biologische Gesamtreife und Entwicklung eines Menschen ist nicht sein kalendarisches Alter, sondern sein Knochenalter. Aus diesem Grund besitzt die Knochenaltersbestimmung in der Diagnostik von Wachstums- und Entwicklungsstörungen besondere klinische Relevanz.

In der Methodik der radiologischen Skeletaltersbestimmung haben sich nur Methoden etablieren können, welche für die Altersbestimmung Radiogramme der Hand nutzen. Das Handskelett wird aufgrund seiner Vielzahl sekundärer Ossifikationszentren als repräsentativer Teil des Skelettsystems angesehen. Zu den am weitesten verbreiteten Methoden der Knochenaltersbestimmung zählt die Atlasmethode von Greulich & Pyle, die auf der subjektiven Beurteilung des Radiogramms durch den Untersucher basiert. In der Fachliteratur lassen sich aufgrund dieser Subjektivität oft widersprüchliche Angaben über die Interbeurteiler-Variabilität der Methode finden. Auch der Einfluss vom Grad der Qualifizierung eines Untersuchers sowie seiner Erfahrung im Umgang mit der Methode werden kontrovers diskutiert.

Ziele dieser Studie waren deshalb die erneute Begutachtung der widersprüchlichen Literaturangaben über die Interbeurteiler-Variabilität der Greulich & Pyle-Methode und die Evaluierung der Tauglichkeit einer neuentwickelten sonographischen Methode zur Knochenaltersbestimmung im Vergleich zur Standardmethode von Greulich & Pyle.

160 konsekutiv ausgewählte Patienten (77 Knaben, 83 Mädchen), bei denen eine Röntgenaufnahme der linken Hand indiziert war, wurden in die Studie eingeschlossen. Diese 160 Patienten, alle Kinder und Jugendliche im Alter von 5-18 Jahren, unterzogen sich am selben Tag der Röntgenuntersuchung und der neuen sonographischen Methode zur Knochenaltersbestimmung mit dem Bonage-System (Sunlight Medical Ltd., Tel Aviv, Israel). Das System evaluiert die Beziehung zwischen der Schalleitungsgeschwindigkeit im Bereich der distalen Radius- und Ulnaepiphyse und dem Wachstum unter Verwendung von im System integrierten geschlechts- und ethnizitätsspezifischen Algorithmen. Vier verschiedene Untersucher (U1 - U4) - alle geblindet für das kalendarische Alter und das Bonage-Ergebnis - analysierten die Röntgenbilder und bestimmten das Knochenalter nach der Methode von Greulich & Pyle. Miteinander verglichen wurden

sowohl die Ergebnisse der einzelnen Untersucher untereinander (U1 - U4) als auch die Ergebnisse der Untersucher 1, 2 und 3 mit den Ergebnissen von Bonage. Außerdem überprüfte die Studie, ob zwischen der Atlasmethode und der Bestimmung mit Bonage Unterschiede in der Präzision beider Verfahren bestehen. Die statistische Auswertung erfolgte mit SPSS 12.0.1.

Die Messung mit dem Bonage-System konnte bei 152 Patienten erfolgreich durchgeführt werden. Bei der Bestimmung des Knochenalters nach der Greulich & Pyle-Methode lagen die Korrelationen der Untersucher untereinander zwischen 0,977 und 0,985. Die Korrelationen zwischen Bonage und den Untersuchern 1 - 3 lagen zwischen 0,902 und 0,920. Der t-Test für gepaarte Stichproben wies für die Auswertungspaare Bonage vs. U1, U2 und U3 und für die Paare U1/ U2, U1/ U4 und U2/ U4 keine statistischen Signifikanzen auf dem 5%-Niveau nach. Im Vergleich der Untersucher U3 vs. U1, U3 vs. U2 und U3 vs. U4 ergab der t-Test signifikante Differenzen ($p < 0,05$). In der Präzision ließen sich Unterschiede zu Ungunsten der sonographischen Modalität nachweisen.

Die Ergebnisse der Untersucher zeigten (außer für die Paare, an denen U3 beteiligt war) eine hervorragende Übereinstimmung und gute Korrelationen bei allen Paaren. Die Interbeurteiler-Variabilität lag in der vorliegenden Studie zwischen 0,3 (0,29) und 2 (1,97) Monaten und ist im Hinblick auf die Subjektivität der Methode und unter Beachtung der Literaturangaben als außerordentlich gering einzuordnen.

Vorteile der sonographischen Methode sind fehlende ionisierende Strahlung, Objektivität und einfache Durchführbarkeit. Schon jetzt stellt die sonographische Knochenaltersbestimmung mit ihren vielen Vorteilen eine valide Ergänzung zu den konventionellen Verfahren dar. Im Rahmen der Erstvorstellung von Patienten ist weiterhin die Röntgenaufnahme der Hand erforderlich, um ein dissoziiertes Knochenalter, Fehlbildungen und Mineralisationsstörungen zu erkennen, die mit dem Bonage-System nicht diagnostizierbar sind.

1 Einleitung

1.1 Definition von Skelettreife bzw. Knochenalter

Das Skelettalter (= Knochenalter) ist definiert als das chronologische Alter, in dem die Population, von welcher die Standards abgeleitet wurden, einen definierten Entwicklungszustand des knöchernen Skelettsystems aufweist (Zerin und Hernandez 1991).

Obwohl Wachstum und Skelettreifung meistens einheitlich verlaufen, sind sie biologisch getrennte Prozesse. Die Reifung des Skeletts steht v.a. in Stadien fortgeschrittener Reife eher mit der sexuellen Entwicklung als mit der Entwicklung der Körperstatur in Zusammenhang. So setzt die Menarche bei - sich in der Pubertät befindlichen - Mädchen zeitnah nach dem Schluss der Epiphysenfugen der distalen Phalangen ein. Sowohl das chronologische Alter als auch die Körperstatur können häufig nicht als präzise Reifeindikatoren betrachtet werden (Zerin und Hernandez 1991).

1.2 Die Entwicklung des Skeletts

Im Prozess seiner Entwicklung und Reifung durchläuft das Skelettsystem 3 aufeinanderfolgende Stadien. Im pränatalen Stadium werden die Diaphysen (= primäre Ossifikationszentren) angelegt. Gleichfalls sind zum Zeitpunkt der Geburt bereits einzelne Knochenkerne (=Epiphysen) existent. Zu diesen gehören die Femurepiphyse und die proximale Tibiaepiphyse, die als Reifezeichen dienen. Talus, Calcaneus und das Cuboid sind ebenso angelegt. Im zweiten Stadium, der Zeit zwischen Geburt und Beginn der Pubertät, beginnen Ossifikation und Differenzierung weiterer Epiphysen (= sekundäre Ossifikationszentren). Im dritten Lebensmonat erfolgt die Anlage der Ossifikationskerne des Os capitatum und des Os hamatum. Erst im zweiten Lebensjahr tauchen die Epiphysen des Radius und der kurzen Röhrenknochen (Metacarpalia und Phalangen) auf. Im letzten Stadium, in der Zeit vom Beginn der Pubertät bis zum Ende des Wachstums, schließt sich die Ausreifung der Epiphysen und deren Fusion mit den Metaphysen an (Minas und Minas 2004). Dieser Zeitpunkt stellt das Ende des Wachstums dar. Er

beginnt bei weiblichen Jugendlichen etwa im 14. Lebensjahr, bei männlichen Jugendlichen etwa im 16. Lebensjahr.

1.3 Die Skelettentwicklung beeinflussende Faktoren

Einen Einfluss auf die Skelettentwicklung haben genetische, ethnische, klimatische und sozioökonomische Faktoren (Kemperdick 1986). Das Gewichtsverhalten übt ebenfalls einen Effekt auf die skeletale Reifung aus. Langandauernde Hungerzustände führen zu einer verminderten Entwicklung. Wird die Unterernährung behandelt, kann ein Aufholwachstum eintreten (sog. „catch-up“), bei dem bis zu 4fach beschleunigte Wachstumsraten auftreten können und auch die Reifung des Skeletts um das 2-3fache erhöht sein kann (Prader et al. 1963). Eine Vielzahl von Krankheiten können sich in unterschiedlicher Art auf die Skelettentwicklung auswirken (Tabelle 1 und Tabelle 2) (Kemperdick 1986, Heinrich 1986).

Tabelle 1: Erkrankungen mit retardierter Skelettentwicklung

Kongenitale Erkrankungen	intrauterine Wachstumsstörungen Silver-Russel-Syndrom Alkohol-Embryopathie primordiale Kleinwüchsigkeit Skelettdysplasien Malformationssyndrome
Chromosomale Erkrankungen	Trisomie 21 (Down-Syndrom) Trisomie 18 (Edwards-Syndrom) Monosomie X0 (Turner-Syndrom)
Endokrinologische Erkrankungen	hypophysär Panhypopituitarismus Wachstumshormon-Mangel Tumoren, z.B. Kraniopharyngeom thyreoidal Hypothyreose adrenal M. Addison Cushing-Syndrom gonadal Hypogonadismus
Andere Erkrankungen	chronische Erkrankungen Malnutrition Mineralisationsstörungen

Tabelle 2: Erkrankungen mit akzelerierter Skelettentwicklung

Kongenitale Erkrankungen	Cerebraler Gigantismus
	Fehlbildungssyndrome
	Beckwith-Wiedemann-Syndrom
	Akrodysostose
	Cockayne-Syndrom
	Marshall-Syndrom
	Weaver-Syndrom
	Pseudohyperparathyroidismus
	Adipositas
	Endokrinologische Erkrankungen
	Pubertas praecox vera
	Pseudopubertas praecox
	AGS
	Tumoren der Nebenniere
	Tumoren des Hodens
	Tumoren des Ovars
	gonadotropinproduzierende Tumoren iatrogen (Sexualsteroidbehandlung)
	McCune-Albright-Syndrom
	prämatüre Adrenarche
Andere Erkrankungen	Hochwuchs
	Adipositas

Hormonale Faktoren bilden ein Hauptsystem in der Beeinflussung und Steuerung des Wachstums und der Skelettentwicklung (Heinrich 1986).

Das somatotrope Hormon (= STH) nimmt in der Regulierung des Körperwachstums eine zentrale Rolle ein. Das im Hypophysenvorderlappen gebildete Proteohormon induziert in der Leber die Bildung von Somatomedinen, die sogenannten Insulin like growth factors I und II. Diese Mediatoren sind im Knorpel und in der Wachstumszone des Knochens ausschließlich für die Wirkung des STH verantwortlich (Löffler und Petrides 1998). Ein Mangel an oder eine Resistenz gegenüber STH verursachen einen Mangel an Somatomedinen. Eine Dezeleration des Wachstums und der Skelettreifung - bspw. in Form einer Nanosomie - sind die Folge. Ein Überschuss an STH führt gegensätzlich zu einer Überstimulation der Somatomedinproduktion und somit zu einer Akzeleration der Skelettentwicklung, z.B. Hochwuchs oder Akromegalie.

Die Schilddrüsenhormone beeinflussen ebenfalls die Skelettentwicklung, v.a. im Säuglings- und Kleinkindalter. Das Thyroxin und das Trijodthyronin (T_3) fördern das

Körperwachstum über die Stimulation der Biosynthese des STH im Hypophysenvorderlappen und über einen direkten Effekt am Knochen (Löffler und Petrides 1998). Auf einen Mangel an Schilddrüsenhormonen reagiert das Skelett sehr empfindlich. Ein deutlich retardiertes Wachstum sowie eine retardierte Knochenreifung sind typische Symptome der Hypothyreose (Sitzmann 2002). Die Hyperthyreose hat dagegen nur mäßige Auswirkungen auf Wachstum und Skelettreifung.

Die Wachstumshemmung ist eine bedeutsame Wirkung der Glukokortikoide. Übersteigt die Sekretionsrate das 2-3fache der Norm, kann daraus eine nahezu vollständige Stagnation des Längenwachstums resultieren. Liegen sehr hohe Dosen von Glukokortikoiden vor, tritt die gleichzeitige Hemmung von Längenwachstum und Skelettreifung auf, die zu einem irreversiblen Verlust von Wachstumspotential führen kann (Heinrich 1986). Darüber hinaus stimulieren Glukokortikoide die Osteoklasten, die durch ihre knochenresorbierende Aktivität einen direkten Effekt auf die Knochen ausüben (Steroidosteoporose).

Die Bildung und Sekretion der männlichen Sexualhormone (Testosteron, 5-alpha-Dihydrosteron, Androstendion, Androstenon) wird hauptsächlich durch das Hypophysenhormon LH (luteinisierendes Hormon) gesteuert. Hauptbildungsort der Androgene sind die Leydig-Zwischenzellen des Hodens. Kleinere Mengen von Androgenen werden aber auch in den Ovarien wie auch in der Nebennierenrinde gebildet. Androgene zählen zu den anabolen Hormonen (Löffler und Petrides 1998). Da ihre Sekretion mit Eintritt in die Pubertät stark ansteigt, liegt der größte Einfluss auf Wachstum und Skelettentwicklung im Jugendalter. Sind Androgene in geringem Überschuss vorhanden, bewirken sie eine Proliferation epiphysären Knorpels und die Förderung des Längenwachstums. In höherer Dosis fördern sie den frühzeitigen Schluss der Epiphysenfugen, die Kalzifizierung des Knochens und verursachen eine verringerte Körperendgröße (Schmidt et al. 2000). So zeigen Kinder, die an einem adrenogenitalen Syndrom (AGS) leiden, in ihrem Entwicklungsverlauf ein beschleunigtes Längenwachstum verbunden mit einem akzelebrierten Knochenalter. Um das 10. Lebensjahr lässt das Längenwachstum dann aber stark nach. Durch einen verfrühten Epiphysenschluss imponiert dann ein Wachstumsstillstand. Patienten, die an einem AGS leiden, sind also meist große Kinder, aber kleine Erwachsene (Sitzmann 2002). Die wichtigsten Vertreter aus der Gruppe der weiblichen Sexualhormone, der Östrogene, sind Östradiol, Östron und Östriol. Hauptbildungsort

der Östrogene sind v.a. die Follikel in den Ovarien. Kleinere Mengen von Östrogenen werden aber auch in den Hoden und in der Nebennierenrinde gebildet. Ihr Einfluss auf die Skelettentwicklung besteht v.a. in der Förderung des Epiphysenschlusses und in der Hemmung der osteoklastären Aktivität. Sie stimulieren die Knochenreifung bei gleichzeitiger Hemmung des Längenwachstums (Schmidt et al. 2000).

Eine zusammenfassende Darstellung von Hormoneinflüssen auf die Skelettentwicklung zeigt Tabelle 3 (Heinrich 1986, Löffler und Petrides 1998).

Tabelle 3: Hormone und ihr Einfluss auf die Entwicklung des Skeletts

Hormon		Reifung des Skeletts	Längen- wachstum	Einfluss auf Endgröße
somatotropes Hormon	Überschuss	↑	↑ ↑	↑ ↑
	Mangel	↓ ↓	↓ ↓	↓ ↓
Thyroxin / T3	Überschuss	(↑)	(↑)	(↑)
	Mangel	↓ ↓	↓ ↓	↓ ↓
Cortisol	Überschuss	↓	↓ ↓	↓ ↓
Androgene	Überschuss	↑ ↑	↑ ↑	↓ ↓
	Mangel	↓	↑	eunuchoid
Östrogene	Überschuss	↑ ↑	↑ / ↓	↓ ↓

1.4 Notwendigkeit der Knochenaltersbestimmung

Das Skelettalter spiegelt als objektives Kriterium die biologische Gesamtreife des Organismus wider und ist deshalb von großer diagnostischer Bedeutung (Bierich 1976). Als wichtiger Reifeindikator besitzt die Knochenaltersbestimmung einen hohen Stellenwert in der Aufklärung von Wachstums- und Entwicklungsstörungen (Heinrich 1986). Des Weiteren ist die Bestimmung des Skelettalters fester Bestandteil der Diagnostik bei Kindern, die für ihr chronologisches Alter zu groß bzw. zu klein sind (Sproul und Peritz 1971) und bildet die Grundlage für die gebräuchlichsten Methoden zur Abschätzung der prospektiven Endgröße (Heinrich 1986), z.B. Größenvorhersage mit der Methode von Bayley & Pinneau (Bayley und Pinneau 1952) oder der Methode von Tanner & Whitehouse (Tanner et al. 1983).

Eine weitere Anwendungsmöglichkeit findet man in der Kontrolle von Hormonbehandlungen: Erwünschte, therapeutische und unerwünschte Wirkungen lassen sich mithilfe der Skelettentwicklung verfolgen (Heinrich 1986). In der Diagnostik und Behandlung zahlreicher Endokrinopathien ist die Kenntnis des Skeletalters ebenfalls von Bedeutung (Hernandez et al. 1979).

Verschiedene Erkrankungen können mit Reihenfolgestörungen der Handskelettentwicklung verbunden sein. Verspätetes bzw. verfrühtes Auftreten von Ossifikationszentren oder Unterschiede von karpaler und phalangealer Ossifikation können wertvolle Hinweise auf spezifische Erkrankungen geben (Heinrich 1986):

- Radiushypoplasie, Fanconi-Anämie und cerebraler Gigantismus sind oft mit verspätetem Auftreten bzw. Fehlen des Os Naviculare assoziiert,
- epiphysäre Dysplasien können sich in einem verspäteten Auftreten des Os capitatum äußern,
- Homozystinurie und das Seckel-Syndrom können mit retardiertem Erscheinen des Os lunatum vergesellschaftet sein,
- cerebraler Gigantismus und Trisomie 18 zeigen eine dissoziierte Reifung der Carpalia und der Phalangen (Akzeleration der phalangealen Entwicklung gegenüber der Entwicklung der Carpalia),
- das Beckwith-Wiedemann-Syndrom kann mit einer retardierten Entwicklung der Phalangen im Vergleich zu den Carpalia einhergehen.

In der Orthopädie ist die Knochenaltersbestimmung in der therapeutischen Intervention von Beinlängendifferenzen von großer Wichtigkeit. Für die Indikationsstellung zur operativen Korrektur ist die Kenntnis des noch zu erwartenden Längenwachstums unerlässlich. Zur Kalkulation der zu erwartenden Beinlänge bedient man sich der Methode nach Moseley, bei der die Kenntnis des Knochenalters sowohl die Basis für die Berechnung ist als auch den Zeitpunkt der Operation festlegt (Cundy et al. 1988).

Im Bereich der Kieferorthopädie ist das Skeletalter wichtig, um für kieferkorrigierende Eingriffe den Beginn des puberalen Wachstumsspurts vorausberechnen zu kön-

nen. Für den Erfolg der Therapie ist der exakte Zeitpunkt des Beginns entscheidend. Knochenaltersbestimmung und Wachstumsdiagnostik sind hier für die Erstellung eines zeitlich gut abgestimmten Behandlungsplanes essenziell (Grave 1994).

Ein zunehmender Bedarf der Knochenaltersbestimmung lässt sich im deutschen Rechtssystem bzw. in der Rechtsmedizin verzeichnen. Immer häufiger wendet sich die Justiz an medizinische Gutachter mit der Fragestellung nach dem kalendarischen Alter von jungen, ausländischen Asylsuchenden, die sich nicht identifizieren lassen oder eine falsche Identität angenommen haben. Die Gutachter sollen anhand des Knochenalters eine Aussage über das Lebensalter treffen um so bspw. über die Anwendung des Jugend- bzw. des Erwachsenenstrafrechts zu entscheiden (Kreitner et al. 1998).

1.5 Geschichte der Skelettreife- bzw. Knochenaltersbestimmung

Die Nutzung von Handradiogrammen für die Einschätzung der Reife des Skelettes zählt zu den ältesten Methoden in der Radiologie. Wilhelm Conrad Röntgen entdeckte 1895 die nach ihm benannte Strahlung und machte sie für den Menschen nutzbar. Bereits am 1. April 1896 zeigte der Münchner Pädiater von Ranke auf einer Sitzung des Münchener Ärztlichen Vereins einige Röntgenfotografien der Hand und berichtete über die Anwendbarkeit der Röntgenstrahlung zur Illustration von Ossifikationsvorgängen in der menschlichen Hand, die sich von der Geburt bis hin zur vollständigen Entwicklung vollziehen (Fendel 1976). 1898 referierte von Ranke auf der Naturforscherversammlung in Düsseldorf über seine Untersuchungen zur Ossifikation der Hand unter Röntgenbeleuchtung (Kemperdick 1981). Im selben Jahr publizierte der englische Orthopäde J. Poland sein Buch „Skiagraphic atlas showing the development of bones of the wrist and hand“, in welchem er 19 Handradiogramme britischer Kinder im Alter von 1 - 17 Jahren beschrieb (Poland 1898). In den Jahren 1907 und 1909 folgten sorgfältiger ausgearbeitete Publikationen von J.W. Pryor (Pryor 1907) und T.M. Rotch (Rotch 1909). Zahlreiche Veröffentlichungen mit Beschreibung unterschiedlicher Methoden zur Knochenaltersbestimmung folgten diesen Ausarbeitungen: 1937 T.W. Todd (Todd 1937), 1946 O. Elgenmark (Elgenmark 1946), 1952 W.W. Greulich & I. Pyle (Greulich und

Pyle 1959), 1954 R.M. Acheson (Acheson 1954), 1962 J.M. Tanner & R.H. Whitehouse (Tanner et al. 1962).

1.6 Methodik der Knochenaltersbestimmung

Zur Bestimmung des Skelettalters lassen sich drei verschiedene Ansätze unterscheiden:

- Numerische Verfahren,
- Morphische Verfahren,
- Metrische Verfahren

Für jedes Verfahren werden ein oder auch mehrere Röntgenbilder benötigt. Röntgenologische Techniken sind sehr einfach standardisierbar und lassen sich in allen Altersklassen leicht durchführen. Diese Techniken sind generell gut anzuwenden, selbst bei mäßigen bis schweren Deformationen des Handskeletts. Nur sehr unkooperative, unwillige Kinder stellen ein Problem bei dieser Art der Untersuchung dar (Zerin und Hernandez 1991).

Numerische Verfahren

Die numerischen Verfahren gehören zu den älteren Techniken. Prinzip dieser Methoden ist das Auszählen der bereits erschienenen sekundären Ossifikationszentren bzw. die Bestimmung der Reihenfolge und des Zeitpunktes, an dem die sekundären Ossifikationszentren röntgenologisch sichtbar werden.

Als Beispiele dieser numerischen Verfahren sollen die Methode nach Elgenmark (Elgenmark 1946) und die Methode nach Garn, Rohmann & Silverman (Garn et al. 1967) genannt sein. Für die Methode nach Elgenmark benötigt man entweder die Skelett-Röntgenaufnahme einer gesamten Körperhälfte oder einzelne Aufnahmen von 7 verschiedenen Körperregionen. Neben der heute nicht mehr vertretbaren Strahlenexposition - oft ist es im Verlauf einer Behandlung notwendig, Skelettaltersbestimmungen

mehrmals zu wiederholen - rechtfertigen auch der große zeitliche Aufwand sowie die hohen Kosten der Untersuchung das Abwenden von dieser Methode. Ein weiterer Nachteil dieser Methode wird offensichtlich, wenn man den zeitlichen Aspekt der Sequenz des Erscheinens der Ossifikationszentren betrachtet. 60 der 73 wichtigsten postnatalen Ossifikationszentren erscheinen bei Jungen bereits vor Erreichen des 6. Lebensjahres, bei Mädchen sogar 65 (Fendel 1976).

Die Methode nach Garn, Rohmann & Silverman dagegen nutzt sog. age-of-appearance-Listen. In diesen Listen sind fest definierte Zeitpunkte für das Erscheinen der wichtigsten sekundären Ossifikationszentren zusammengefasst, die dem Untersucher als Referenzwerte dienen. Das Knochenalter wird ermittelt, indem man die zuletzt erschienenen Ossifikationszentren bestimmt und deren „normales Erscheinungsalter“ aus der Liste ablesen kann (Garn et al. 1967). Separate Standards existieren sowohl für unterschiedliche Ethnizitätsgruppen (Schwarze und Weiße) als auch für männliches und weibliches Geschlecht (Zerin und Hernandez 1991).

Jedoch verbietet der Strahlenschutz heutzutage Methoden, bei denen die Bestimmung des Skeletalters auf Röntgenaufnahmen mehrerer Skelettabschnitte beruht (Kemperdick 1986).

Morphische Verfahren

Die morphischen Verfahren beruhen auf dem Prinzip des Vergleichens. Individuelle Röntgenaufnahmen von Teilen des Skeletts werden mit standardisierten Aufnahmen verschiedenster Stadien der Skelettentwicklung aus einer vergleichbaren Population verglichen. In die Gruppe dieser Verfahren zählen u.a. die Methode nach Greulich & Pyle (Greulich und Pyle 1959) und die Methode nach Tanner & Whitehouse (Tanner et al. 1975), die später näher beschrieben werden.

Metrische Verfahren

Den metrischen Verfahren liegen planimetrische Bestimmungen der Projektionsflächen der Knochenkerne in einer Röntgenaufnahme zu Grunde. Aus der Länge und / oder

der Breite der projizierten Knochen errechnet man die Größe der Fläche (zitiert nach: Fendel 1976). Über die Ausdehnung der Knochen in der Ebene senkrecht zum Röntgenfilm lässt sich in der zweidimensionalen Darstellung naturgemäß keine Aussage treffen. Mit diesem Verfahren lässt sich für jeden einzelnen Knochen eine Größe bestimmen. Da „Größe“ ein homogenes Merkmal ist, können mathematisch arithmetische und statistische Verfahren angewendet werden. So lässt sich aus der Einzelwert-Streuung die Varianz errechnen und aus der Varianz der Varianz-Koeffizient. Das arithmetische Mittel der Flächengrößen einzelner Knochen wird dann wiederum mit Standardgrößen verschiedener Altersgruppen verglichen (Fendel 1976).

1.7 Die röntgenologische Untersuchung der Hand

In der Routine der radiologischen Skelettaltersbestimmung haben sich nur die Verfahren durchgesetzt, welche die Hand als repräsentative Region für die Ossifikation des gesamten Skeletts nutzen. Im kindlichen Handskelett lassen sich 29 sekundäre Ossifikationszentren darstellen. Die Reifung bzw. Entwicklung dieser Ossifikationszentren gestattet zufriedenstellende Aussagen während des Heranwachsens vom älteren Säugling bis hin zum Adoleszenten (Fendel 1976). Die Tatsache, dass die linke Hand für die Untersuchungen ausgewählt wird, findet ihren Ursprung in einer internationalen Empfehlung über die Vereinheitlichung anthropologischer Messungen am Menschen. Dieses Abkommen wurde nach den Internationalen Konferenzen für Anthropologie 1906 in Monaco und 1912 in Genf getroffen (van Rijn et al. 2001).

1.7.1 Bestimmung beim Neugeborenen und jungen Säugling

Bis zu einem Alter von 3 Monaten existieren noch keine Knochenkerne im Handwurzelbereich. Deshalb ist eine Beurteilung nach Methoden, die auf einer morphischen Analyse der knöchernen Hand beruhen, in diesem Alter noch nicht möglich. Aus diesem Grund basiert die Skelettaltersbestimmung bei Neugeborenen und jungen Säuglingen auf Röntgenaufnahmen, die die Knochenkernentwicklung am Kniegelenk sowie im Bereich des Sprunggelenkes abbilden (zitiert nach: Zerín und Hernandez 1991). Die Skelettaltersbestimmung in dieser Altersgruppe wird anhand der - bei reifen Neuge-

borenen - bereits zum Zeitpunkt der Geburt vorhandenen Ossifikationszentren (siehe „Die Entwicklung des Skelett“) durchgeführt (Kemperdick 1986). Diese Knochenkerne zeigen in den ersten Lebensmonaten röntgenologisch deutlich sichtbare Veränderungen, sodass deren röntgenologische Untersuchung bei Kindern unter 2 Jahren eine hilfreiche Ergänzung zu den Untersuchungen an der Hand sein können (Zerin und Hernandez 1991). Die von den Carpalia im 1. Lebensjahr gezeigten Veränderungen sind nur geringfügig und lassen eine genaue Beurteilung des Knochenalters kaum zu (Zerin und Hernandez 1991). Für die Beurteilung der Knochenkerne von Knie und Fuß werden vorwiegend metrische Methoden eingesetzt. Trotzdem wird nach Vollendung des 3. Lebensmonates eine Röntgenaufnahme der Hand zur Skelettaltersbestimmung bevorzugt. Indikationen für die Bestimmung des Skelettalters anhand von Röntgenaufnahmen der unteren Extremität bestehen nur noch, wenn die Anlage des Os capitatum und des Os hamatum im Handradiogramm noch nicht sichtbar geworden ist (Kemperdick 1986) oder wenn sich eine hinreichende Bestimmung nicht durch eine Röntgenaufnahme der Hand erzielen lässt.

1.7.2 Bestimmung des Skelettalters beim älteren Säugling, Kind und Jugendlichen bis zur Adoleszenz

Sobald die knöchernen Kerne des Os capitatum und des Os hamatum erschienen sind, eignet sich ein Radiogramm der Hand für die Knochenaltersbestimmung. Allerdings kommen weniger metrische Methoden zur Anwendung, sondern eher qualitative Methoden mit Wichtung auf den Zeitpunkt des Erscheinens verschiedener Knochenkerne, Form- und Gestaltwandel (Reifung) von Knochen und Epiphysen sowie den Epiphysenschluss (Kemperdick 1986). Diese jeweils charakteristischen Reife- bzw. Entwicklungsstadien sind auch in der zweidimensionalen Darstellung eines Handradiogramms recht gut zu identifizieren bzw. zu beschreiben (Fendel 1976).

1.7.2.1 Methode nach Greulich & Pyle

Das von Greulich & Pyle ausgearbeitete Verfahren basiert auf den von W. Todd publizierten Daten. Dieser veröffentlichte 1937 einen Atlas, der die Knochenaltersbestim-

mung unter Anwendung einer morphischen Methode beschreibt (Todd 1937). Greulich & Pyle griffen Todds Verfahren auf, entwickelten dieses weiter und komplettierten Todds Atlas. 1950 veröffentlichten beide Autoren ihren ersten Atlas, in welchem sie eine eigene Methode mit von ihnen abgeleiteten - und in dem Atlas abgedruckten - Standards vorstellten. Im Jahr 1959 erschien dann die zweite Auflage des Atlas (Greulich und Pyle 1959). Die als Referenz dienenden Standards stammen von gesunden Kindern und Jugendlichen - alle Angehörige einer gehobeneren sozioökonomischen Gesellschaftsschicht - aus der Gegend um Cleveland (Ohio, USA) und wurden in den Jahren 1931 - 1942 zusammengetragen. Jedes Standardradiogramm, das im Atlas abgedruckt ist, wurde ausgewählt aus 100 Handröntgenbildern von Probanden gleichen Alters und Geschlechts (van Rijn et al. 2001).

Da in der vorliegenden Arbeit die Methode nach Greulich & Pyle (Auflage 1959) für die Knochenaltersbestimmung genutzt wurde, soll diese Methode im Folgenden ausführlicher beschrieben werden.

Das Prinzip dieses Verfahrens besteht darin, die zu befundenden Handradiogramme jeweils mit einem geschlechts- und alterstypischen Standard zu vergleichen. Bei der Auswertung zwingt der Atlas dem Untersucher kein festgelegtes Procedere auf, dennoch wird folgendes Vorgehen empfohlen:

Zu Beginn vergleicht der Untersucher das Handröntgenbild mit den Standards des gleichen Geschlechts. Ist ein Standard-Handradiogramm gefunden, das dem auszuwertenden Röntgenbild annähernd entspricht, so werden auch die angrenzenden Standard-Handradiogramme mit dem zu befundenden Röntgenbild verglichen. Auf diese Weise soll das Standard-Handradiogramm ermittelt werden, das die meisten Übereinstimmungen mit dem zu befundenden Röntgenbild aufweist. In den einzelnen Entwicklungsstadien, vom Kleinkind bis zum Adoleszenten, helfen spezifische Merkmale der Knochenentwicklung bei der Altersbestimmung der Knochen:

- im frühen Kindesalter: das Vorhandensein bzw. das Fehlen von carpalen und epiphysären Ossifikationszentren,
- vom Schulalter bis zum Beginn der Pubertät: Form- und Gestaltentwicklung der Knochenkerne und charakteristische Reifezeichen der einzelnen Knochen,

- während der Pubertät und der späten Adoleszenz: der Grad der Fusion der Epiphysen mit den Metaphysen

Existiert kein Standardradiogramm, das dem Handradiogramm des Patienten eindeutig zugeordnet werden kann, liegt das Knochenalter am wahrscheinlichsten zwischen zwei benachbarten Standards. In diesem Fall wird das Knochenalter interpoliert. Das angrenzende Standardpaar wird ermittelt und das Knochenalter durch das arithmetische Mittel der angrenzenden Standards bestimmt. Der resultierende Mittelwert entspricht dann am wahrscheinlichsten dem realen Knochenalter des Patienten. In der Abbildung 1 sind zur Veranschaulichung 4 Handradiogramme von männlichen Patienten unterschiedlichen Alters dargestellt.

Normalerweise zeigt sich eine ausreichend gute Übereinstimmung des Entwicklungsstadiums der einzelnen Knochen um für alle ein annähernd gleiches Knochenalter voraussetzen zu können. Dennoch existiert ein erheblicher Anteil von Patienten, deren Handradiogramm ein dissoziiertes Knochenalter aufzeigt. Bei diesen Patienten ist das Entwicklungsstadium einzelner oder auch mehrerer Knochen retardiert bzw. akzeleriert. Die alleinige Angabe des Knochenalters, welchem die Mehrzahl der im Röntgenbild sichtbaren Knochen entspricht, ist in diesen Fällen nicht ausreichend. Das Knochenalter der Knochen, die von der Entwicklungsnorm abweichen, soll zusätzlich bestimmt und dokumentiert werden. Für die Knochenaltersbestimmung in diesen Fällen existiert ein separates Kapitel im Atlas. In diesem Kapitel beschreiben W. Greulich und I. Pyle spezielle charakteristische Reife-Indikatoren, mit deren Hilfe sich das Knochenalter einzelner Knochen ermitteln lässt.

1.7.2.2 Oxford-Methode

Von J. Acheson wurde 1954 die sogenannte Oxford-Methode eingeführt. Diese bestimmt den Reifegrad einzelner Knochen - zunächst für Hand, Ellenbogen und das Kniegelenk, später auch für das Hüftgelenk - und ordnet den individuellen Reifestadien verschiedene Reifepunkte (Scores) zu. Die Summe der Punkte entspricht dem Knochenalter. Die Oxford-Methode ist der Vorgänger der Tanner & Whitehouse Methode und nach Auffassung einiger Autoren auch die komplizierteste Methode der Knochenaltersbe-

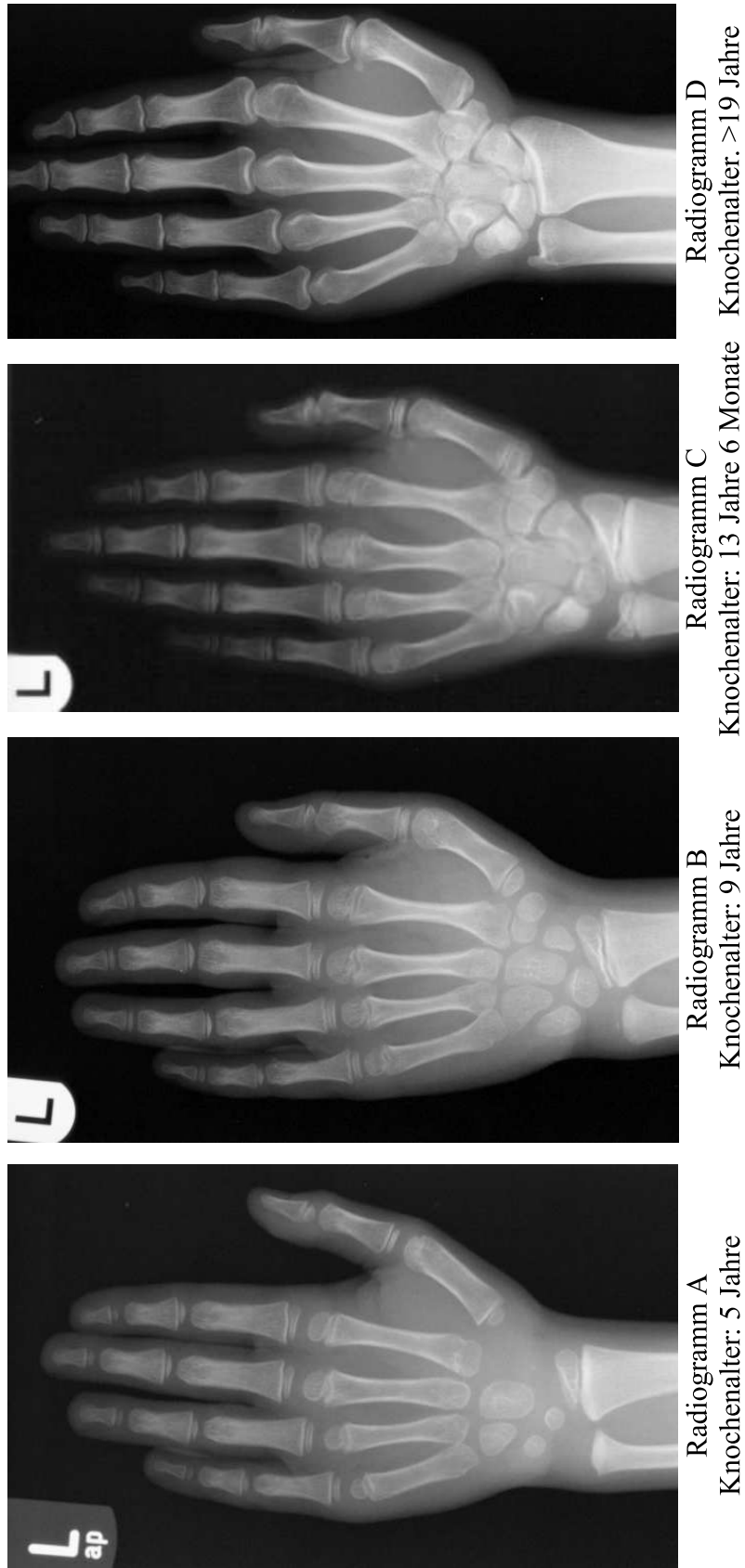


Abbildung 1: Handradiogramme männlicher Patienten. Das Knochenalter wurde nach Greulich & Pyle bestimmt und ist unter jedem Radiogramm angegeben.

stimmung (Bilgili et al. 2003). Tanner & Whitehouse und Mitarbeiter erweiterten das Score-System und veröffentlichten 1962 ihr TW1-System.

1.7.2.3 Methode nach Tanner & Whitehouse

Die heute weit verbreitete TW2-Methode (Tanner et al. 1975) ist der Nachfolger der von J.M. Tanner & R.H. Whitehouse 1962 publizierten, morphischen Methode zur Knochenaltersbestimmung (TW1-Methode). In einem eigenen Atlas werden für 20 verschiedenen Knochen der Hand und des Handgelenkes individuelle und charakteristische Entwicklungsstadien dargestellt. Den einzelnen Knochen sind entsprechend ihres Reife- bzw. Entwicklungsstadiums sogenannte gewichtete Scores (SMS = skeletal maturity score) zugeordnet, über die sich das Knochenalter errechnen lässt. Standard-Handradiogramme, skizzierte Zeichnungen und ein beschreibender Text sollen die Beurteilung des Knochenalters erleichtern. Die Gestalt und die Größe der Epiphysen bzw. der einzelnen Knochen sind für jedes Stadium genau definiert. Jedes Stadium wird anhand von 1 - 3 Kriterien beschrieben. Um die Epiphyse bzw. den Knochen einem Stadium zuordnen zu können, muss bei 1 - 2 beschriebenen Kriterien mindestens 1 Kriterium zutreffen. Bei 3 beschriebenen Kriterien müssen mindestens 2 die Beschreibung erfüllen. Eine weitere Bedingung für die Zuordnung ist die Erfüllung des ersten Kriteriums des vorhergehenden, unreiferen Stadiums.

Im Atlas werden von den Autoren drei verschiedene Möglichkeiten zur Beurteilung des Knochenalters angeboten:

- die B20-Methode (= bone twenty-Methode), bei der insgesamt 20 Knochen beurteilt werden (Phalangen und Metacarpalia des ersten, dritten und fünften Fingerstrahles, die Epiphysen von Radius und Ulna, alle Handwurzelknochen mit Ausnahme des Os pisiforme),
- die RUS-Methode (= radius, ulna, short-bones-Methode; Phalangen und Metacarpalia des ersten, dritten und fünften Fingerstrahles, die Epiphysen von Radius und Ulna),
- Score der Carpalia (mit Ausnahme des Os pisiforme)

Aus dem RUS-Score und dem Score der Carpalia lässt sich ein Gesamt-Score ermitteln, der zwischen 0 und 100 liegt. Das Knochenalter wird bestimmt, indem man die Lage des Gesamt-Scores innerhalb einer Referenzgraphik analysiert. Ihre Standards leiteten Tanner & Whitehouse aus einer Untersuchung aus den Jahren 1945 - 1958 ab. Die Handradiogramme von 2564 Kindern und Jugendlichen, die sozioökonomisch der britischen Mittelklasse zugehörig waren, wurden in dieser Zeit von Tanner & Whitehouse ausgewertet.

1.7.2.4 Fels-Methode

1988 stellten Roche und Mitarbeiter eine knochenspezifische Bestimmungsmethode vor, der Longitudinalstudien von 350 Jungen und 320 Mädchen von der Geburt bis zum Wachstumsende zu Grunde gelegt wurden. Sie ist der TW2-Methode ähnlich. Die Methode basiert auf der Beschreibung und Beurteilung von 98 Reifeindikatoren. Die Anzahl der Reifeindikatoren, die bei der Untersuchung eines Handröntgenbildes analysiert werden, ist dabei vom kalendarischen Alter des Patienten abhängig. Wie bei Tanner & Whitehouse werden je nach Reifeindikator unterschiedliche hohe Scores verteilt. Eine spezielle Computersoftware errechnet aus dem ermittelten Score das Knochenalter. Im Gegensatz zu Tanner & Whitehouse berücksichtigt die Fels-Methode aber die Tatsache, dass ein Reifestadium, das gewöhnlich über einen längeren Zeitraum besteht, weniger Information über den dynamischen Prozess der Reifeentwicklung des Skeletts aussagt als ein Stadium, das schnell durchschritten wird (Gilli 1996).

1.7.3 Probleme bei der Anwendung konventioneller Methoden

In den beiden nachfolgenden Abschnitten sollen aufgrund ihrer klinischen Relevanz nur Schwierigkeiten und Probleme in der Anwendung der Methoden nach Greulich & Pyle (Auflage 1959) und nach Tanner & Whitehouse näher erläutert werden.

Ein generelles Problem beider Methoden liegt im Prozess der skelettalen Entwicklung an sich. Die Skelettentwicklung - insbesondere Beginn und zeitlicher Ablauf - zeigt innerhalb von Populationen unterschiedlicher Ethnizitätszugehörigkeit deutliche

Variabilität. Mora et al. führten Untersuchungen an amerikanischen Kindern und Jugendlichen unterschiedlicher Abstammung durch. Legt man die von Greulich & Pyle publizierten Daten als Referenzwerte zu Grunde, befanden sich 10% der präpubertalen Kinder afrikanischer Abstammung in einem Bereich $> 2s$ und 8% der Kinder europäischer Abstammung in einem Bereich $< 2s$ (Mora et al. 2001). Diese Ergebnisse bekräftigen die Resultate von Sproul und Peritz, die 1971 publizierten, dass ein signifikanter Unterschied zwischen dem Knochenalter der schwarzen und der weißen Bevölkerung existiert. Wiederum in einem Vergleich mit den Referenzwerten von Greulich & Pyle wurde demonstriert, dass das Knochenalter der schwarzen Bevölkerung gegenüber der weißen Bevölkerung um ca. 1 Jahr akzeleriert ist (Sproul und Peritz 1971). Wiesen Mora et al. Entwicklungsunterschiede nur in der präpubertären Periode nach, beschrieben Koc et al. in ihrer Studie Unterschiede in der Skelettreifung auch während der Pubertät (Koc et al. 2001).

1.7.3.1 Methode nach Greulich & Pyle

Der Einfluss von Qualifizierung und Erfahrung des Untersuchers auf die Bestimmung des Knochenalters war Inhalt einer Reihe von Studien. Diese Studien kommen jedoch zu unterschiedlichen Ergebnissen. Groell et al. führten eine Studie durch, in welcher das Knochenalter von vier Ärzten bestimmt wurde, von zwei ständig in der Pädiatrie tätigen Radiologen und zwei radiologischen Assistenzärzten. Der direkte Vergleich verdeutlichte, dass sich die Beurteilung der Radiologen durch kleinere Differenzen ($1,0 \pm 4,8$ Monate) - entspricht einem geringeren Interbeurteiler-Fehler und einer höheren Interbeurteiler-Reliabilität - auszeichnete als die Differenzen der Beurteilungen durch die Assisstenzärzte ($1,3 \pm 7,6$ Monate) (Groell et al. 1999), wenngleich diese Differenzen auf dem 5%-Niveau statistisch nicht signifikant waren. Weitere Studien unterstützen diese Beobachtungen (Johnson et al. 1973, Mora et al. 2001).

Mainland analysierte, ob die Häufigkeit und die Regelmäßigkeit, mit der die Methode vom selben Untersucher angewendet wird, einen Einfluss auf die Knochenaltersbestimmung haben. Er beauftragte eine erfahrene Ärztin, welche Routine in der Anwendung der Methode nach Greulich & Pyle entwickelt hatte, die Knochenalter von 18 verschiedenen Patienten anhand angefertigter Handröntgenaufnahmen zu ermitteln. Danach

pausierte die Ärztin und führte keine Knochenaltersbestimmungen mehr durch. Nach einem Jahr beurteilte sie dieselben 18 Röntgenaufnahmen erneut. Der Vergleich beider Auswertungen, die über ein Jahr auseinander lagen, wies signifikante Unterschiede auf (Mainland 1954). Im Gegensatz zu diesen Erkenntnissen stehen die Untersuchungen von Roche et al.. Seinen Beobachtungen zufolge bedingen die dauerhafte Anwendung des Verfahrens und ein hoher Qualifikationsgrad des Untersuchers keine wesentlichen Verbesserungen in der Beurteilung des Knochenalters (Roche et al. 1970b).

Ein weiterer Einflussfaktor, die Kenntnis des chronologischen Alters vor der Untersuchung, wurde untersucht. Als Resultat zeigte sich, dass zwar keine Beeinflussung des Inter- und Intrabeurteilerfehlers bestehen. Dennoch wurde folgende Beobachtung beschrieben: Ein Untersucher wird das Knochenalter wahrscheinlicher als normgerecht beurteilen, wenn ihm das chronologische Alter des Patienten bekannt ist (Berst et al. 2001).

Ein zusätzliches Problem sind Kinder und Jugendliche mit einer extremen Körperstatur (die Körperhöhe betreffend). Die Validität der Methode nach Greulich & Pyle ist in solchen Fällen nicht mehr gegeben. Bei sehr großen Kindern lässt sich oft eine Akzeleration, bei sehr kleinen Kindern oft eine Retardierung diagnostizieren, die jedoch nicht auf der Grundlage von Wachstumsstörungen basieren, sondern deren Ursprung in ihrer Körperstatur zu sehen ist. Der Atlas von Greulich & Pyle ist dann nicht für die Knochenaltersbestimmung anwendbar (Sproul und Peritz 1971).

Ein ebenfalls strittiger Einflussfaktor auf die Genauigkeit der Skeletaltersbestimmung ist die Beurteilung der Carpalia. Nachdem Acheson et al. 1963 die Handwurzelknochen als eine mögliche Fehlerquelle identifizierte (Acheson et al. 1963, Acheson et al. 1966), schlussfolgerten auch Johnston und Jahina aus ihren Untersuchungen, dass der Ausschluss der Carpalia in der Beurteilung des Knochenalters nach Greulich & Pyle wünschenswert sei (Johnston und Jahina 1965). Wenn der Schwerpunkt der Analyse des Knochenalters auf der Beurteilung der Epiphysen liegen würde, so wären exaktere Bestimmungen möglich. Der Grund dafür wird in der unterschiedlichen Geschwindigkeit der Entwicklungsprozesse der Epiphysen und der Carpalia gesehen. So bedingt die Einbeziehung der Carpalia im Durchschnitt niedrigere Werte für das Knochenalter (Johnston und Jahina 1965). Auch zu dieser Feststellung lässt sich in

der Literatur Gegenteiliges finden. Roche et. al widersprachen in ihren Publikationen diesen Beobachtungen (Roche et al. 1970a, Roche et al. 1970b).

1.7.3.2 Vergleich von Greulich & Pyle mit Tanner & Whitehouse

Beide Methoden unterliegen als morphische Methoden einer starken Subjektivität des Beurteilenden (Johnson et al. 1973). Wird das Knochenalter der gleichen Population jeweils einmal mit der Methode nach Greulich & Pyle bestimmt und einmal mit der Methode nach Tanner & Whitehouse, so wird letztere immer im Durchschnitt höhere Werte für das Knochenalter liefern. Die Angaben über diese Differenz schwanken in der Literatur von 0,4 – 1 Jahr (Acheson et al. 1966, Bull et al. 1999). Reproduzierbarkeit und Genauigkeit sind bei dem von Tanner & Whitehouse entwickelten Verfahren größer und variable Fehler somit kleiner als bei der Methode nach Greulich & Pyle (Acheson et al. 1966, Bull et al. 1999). Die systematischen Fehler sind wiederum bei der Tanner-Whitehouse-Methode größer, da sich durch Beurteilung der einzelnen Knochen und Vergabe des Scores für jeden einzelnen Knochen systematisch abweichende Tendenzen aufsummieren und den Gesamtscore deutlich beeinflussen.

1.8 Anwendung von Ultraschall bei Untersuchungen am Skelettsystem

Durch ständige Weiterentwicklungen und technische Verbesserungen in den vergangenen Jahren konnten sich sonographische Methoden stets neue Einsatzmöglichkeiten erschließen. So wurde der Ultraschall auch für die Osteologie nutzbar gemacht. Seine Anwendung auf dem Gebiet der Knochendichtemessungen bereicherte die herkömmlichen radiologischen Verfahren um eine weitere Alternative. Tabelle 4 zeigt eine kurze Auflistung der Vorteile sonographischer Methoden (Fredfeldt 1986) sowie deren Nachteile (Cagriota-Scanderbeg et al. 1998).

Der Schallleitungsgeschwindigkeit (SOS = speed of sound), die bei quantitativen Ultraschalluntersuchungen am knöchernen System des Menschen bestimmt wird, kommt eine entscheidende Bedeutung zu. Dieser Wert ist ein Maß für die Geschwindigkeit,

Tabelle 4: Vorteile und Nachteile der Sonographie in der Osteologie

Vorteile	Nachteile
<ul style="list-style-type: none">• Applikation von Ultraschall ist unbedenklich und nebenwirkungsfrei• Anwendung sonographischer Methoden am Patienten bedarf nur geringen Aufwands• kostengünstig und wirtschaftlich	<ul style="list-style-type: none">• kein Übersichtsbild über untersuchte Region, sondern nur schichtartige Darstellung• keine Visualisierung von Strukturen, die sich unter stark ultraschallreflektierender oder -absorbierender Formationen befinden• kaum standardisiert, stark von Erfahrung und Geschick des Untersuchers abhängig; die Untersucher-Variabilität hat großen Einfluss auf die Genauigkeit

mit der sich der Ultraschall im Knochen ausbreitet. Mehrere Faktoren nehmen direkt Einfluss auf die Geschwindigkeit der Schallausbreitung. Sievänen et al. zeigten, dass von den physikalischen Eigenschaften der Knochen v.a. die Dichte der Kortikalis den stärksten Einfluss auf die Schalleitungsgeschwindigkeit besitzt (Sievanen et al. 2001). Die Dicke der Kortikalis und die Elastizität des Knochens beeinträchtigen die Schallausbreitung ebenfalls (Fredfeldt 1986). Im Alter nimmt die Dicke der Kortikalis ab und als Konsequenz verringert sich die Geschwindigkeit der Ultraschallausbreitung. Bei der Osteoporose, die mit einer Minderung der Knochenqualität einhergeht, wird auch eine Abnahme der SOS-Werte beobachtet (Craven et al. 1973). Eine weitere Einflussgröße ist die Stärke bzw. Dicke und Zusammensetzung des Weichteilgewebes. Bevor die Ultraschallwellen in den Knochen eintreten können, müssen sie die Gewebe durchdringen, die um den Knochen herum liegen. Um so dicker dieses Weichteilgewebe und um so größer sein Fettanteil ist, desto geringer wird die Geschwindigkeit des Ultraschalls in dem entsprechenden Gewebe sein (Barkmann et al. 2002). Die Geschwindigkeit des Ultraschalls im knöchernen Gewebe ist ca. doppelt so groß wie in Weichteilgeweben (Barkmann et al. 2002). Untersuchungen, bei denen die SOS-Werte an linkem und rechtem Radius der Unterarme bestimmt wurden, zeigten eine Diskrepanz der Werte. Die Probandinnen, bei denen die Messungen stattfanden, waren ausnahmslos Rechtshänder. Die Ergebnisse wiesen einen signifikanten Unterschied zwischen den SOS-Werten des linken und des rechten Radius auf (Vrahoriti et al. 2004). Physische Aktivität ist demnach gleichfalls ein Faktor, der einen Einfluss auf den quantitativen Ultraschall hat (Daly et al. 1997).

1.8.1 Knochenaltersbestimmung mittels sonographischer Methoden

Das Grundprinzip der sonographischen Knochenaltersbestimmung bedient sich ebenso der Analyse gemessener SOS-Werte. Die Grundidee basiert auf der unterschiedlichen Geschwindigkeit, mit der sich der Ultraschall in verschiedenen Geweben ausbreitet. Im Knorpel beträgt die Geschwindigkeit ca. $1700m/s$ und im ausgereiften Knochen $> 2200m/s$. Diesen deutlichen Unterschied nutzt die sonographische Methode aus. Röntgenaufnahmen bilden nur mineralisierten Knochen ab, sind jedoch nicht in der Lage Knorpel zur Darstellung zu bringen. Neugeborene und Kleinkinder besitzen aber gerade viel Knorpelmatrix, die sich erst im Laufe der Entwicklung in Knochensubstanz umwandelt. Die durchschnittliche Masse der Knorpelsubstanz nimmt durch die Ossifikationsprozesse ständig ab. Es lässt sich also behaupten: Je dünner der Knorpel desto fortgeschrittener die Skelettreife (Castriota-Scanderbeg et al. 1998). Heute existieren bereits Publikationen, die eine Altersabhängigkeit der Ultraschallgeschwindigkeit nachweisen (Nemet et al. 2001) und eine signifikante Korrelation zwischen SOS-Werten und Entwicklungs- bzw. Reifestadium belegen (Zadik et al. 2003). Im Prozess der Skelettentwicklung unterliegen die SOS-Meßwerte einer stetigen Zunahme bis zum Erreichen eines Peaks (Rose et al. 2001).

Ultraschallverfahren zeigen gegenüber röntgenologischen Verfahren eine höhere Sensitivität bei der Identifikation und Abschätzung der Dimension von kleineren Ossifikationszentren - z.B. den distalen femoralen Ossifikationszentren des Knies - und sind somit eine Alternative der Knochenaltersbestimmung bei Säuglingen und Kleinkindern mit stark retardiertem Knochenalter (Paesano et al. 1998).

2 Zielstellung

In der Methodik der Knochenaltersbestimmung wurden viele verschiedene Verfahren entwickelt. Von der vorhandenen Vielfalt haben sich aber nur wenige Methoden in der täglichen Routine durchgesetzt. Die am häufigsten angewandten Methoden der Knochenaltersbestimmung sind (Bull et al. 1999):

- die 1959 ausgearbeitete Methode von W.W. Greulich und I. Pyle (Greulich und Pyle 1959) und
- die 1975 von J.M. Tanner und R.H. Whitehouse aus der TW1- weiterentwickelte TW2-Methode (Tanner et al. 1975).

Beide Methoden zählen zu den morphischen Bestimmungsmethoden (Fendel 1976). Eine Röntgenuntersuchung der linken Hand ist Voraussetzung, sodass beide Methoden mit einer, wenn auch nur geringen, Exposition von Röntgenstrahlen verbunden sind (Fendel 1976). Besonders sich im Wachstum und in der Entwicklung befindliche Kinder sind nach der Röntgenverordnung bzw. der Strahlenschutzverordnung vor einer übermäßigen Exposition zu schützen. V.a. aus pädiatrischer Sicht sind aus diesem Grund Untersuchungsverfahren ohne ionisierende Strahlung von großem Interesse (Sitzmann 2002). Trotz der rasch fortschreitenden Entwicklung in der Medizintechnik existiert bis zum jetzigen Zeitpunkt keine äquivalente Untersuchungsmethode, mit der sich das Knochenalter so präzise bestimmen lässt, wie mit einer der beiden röntgenologischen Methoden. Das Ziel dieser Studie lautete deshalb, folgende Fragestellungen zu beantworten:

1. Haben Grad der Qualifizierung (Facharzt für Diagnostische Radiologie mit Zusatzbezeichnung Kinderradiologe, Facharzt für Pädiatrie, Student der Medizin) oder persönliche Erfahrung und Training bei der Anwendung der Methode nach Greulich und Pyle einen Einfluss auf die Ergebnisse der Knochenaltersbestimmung?
2. Unterscheiden sich die Ergebnisse der Bonage-Analyse von den Resultaten der Bestimmung nach der Methode von Greulich & Pyle bzw. ist das neuentwickelte

Bonage-Ultraschallsystem in der Lage, das Knochenalter von Kindern und jugendlichen Patienten im Alter von 5 - 18 Jahren genauso präzise zu bestimmen wie es mit der Atlasmethode möglich ist?

3. Ist die sonographische Knochenaltersbestimmung mit dem Bonage-System für die Anwendung in der täglichen Routine eines Krankenhauses bzw. in der Praxis eines niedergelassenen Arztes geeignet?

3 Patienten und Methoden

3.1 Patienten

Zur Bestimmung des Knochenalters wurden die Untersuchungen im Institut für Diagnostische und Interventionelle Radiologie am Klinikum der Friedrich-Schiller-Universität Jena im Zeitraum vom 30. Januar 2004 bis 14. Mai 2004 durchgeführt. Der Großteil der konsekutiv untersuchten Patienten wurde von der Klinik für Kinder- und Jugendmedizin der Friedrich-Schiller-Universität Jena an das Institut überwiesen, ein kleinerer Anteil der Patienten von niedergelassenen Pädiatern.

Der Antrag auf Zustimmung der Ethikkommission der Friedrich-Schiller-Universität Jena zur Durchführung dieser prospektiv angelegten Studie wurde eingereicht. Die Genehmigung durch die Kommission lag zu Beginn der Untersuchungen vor.

Um an der prospektiv angelegten Studie teilnehmen zu können, wurden die Patienten nach folgenden Ein- bzw. Ausschlusskriterien ausgewählt (Tabelle 5):

Tabelle 5: Ein- bzw. Ausschlusskriterien für die Teilnahme an der Studie

Kriterien:
1. von den Patienten wurde aufgrund verschiedenster Indikationen ein Röntgenbild der linken Hand angefertigt
2. eine Ruhigstellung oder eine Frakturierung des linken Armes führten zum Ausschluss aus der Studie
3. das kalendarische Alter der Patienten musste zum Zeitpunkt der Untersuchung mindestens 5 Jahre betragen, durfte aber das 18. Lebensjahr nicht überschreiten
4. eine schriftliche Einverständniserklärung musste vor der Untersuchung unterzeichnet werden

Vor der Untersuchung jedes Patienten wurde diesem bzw. seinem gesetzlichen Vertreter ein Aufklärungsbogen vorgelegt, der eine kurze Zusammenfassung über die Notwendigkeit der Knochenaltersbestimmung und die Methodik mittels Ultraschall gab. Die Teilnahme an der Studie erfolgte freiwillig. Die Diagnosen und Symptome, die die Anfertigung eines Röntgenbildes indizierten, sind in Tabelle 6 zusammengefasst. Bei einem Teil der Patienten konnten mehrere Diagnosen gestellt werden, sodass Mehrfachzuordnungen zu verschiedenen Erkrankungen bzw. Diagnosen möglich waren. Aus

Tabelle 6: Häufigkeit und Verteilung der Diagnosen

Diagnosen und Symptome	Anzahl der Patienten	% des Patientenkollektivs	
Minderwuchs, KEVZ	62	38,8	
Wachstums- und Entwicklungsstörungen	Hochwuchs, KEVS	43	26,9
	Pubertas praecox	10	6,2
	Pubertas tarda	9	5,6
Adipositas	13	8,1	
Hypo- bzw. Hyperthyreose	8	5,0	
Marfan-Syndrom	8	5,0	
AGS	4	2,5	
Asthma	3	1,9	
CED, Colitis ulcerosa, M.Crohn	3	1,9	
Neurofibromatose	2	1,3	
Leistenhoden oder Anorchie	2	1,3	
intrasellärer Tumor, Monosomie XO, ovarielle Insuffizienz, PKU, Silver-Russel-Syndrom, Somatisierungsstörung, Trisomie 21, cystische Fibrose	je 1 ×	je 0,6 %	

Tabelle 7: Alter und anthropometrische Charakteristika

	gesamtes Patientenkollektiv ($\pm s$)	männliche Patienten ($\pm s$)	weibliche Patienten ($\pm s$)
Alter	12 (± 3) Jahre	12,6 (± 3) Jahre	11,4 (± 3) Jahre
Größe	149 (± 21) cm	152 (± 22) cm	146 (± 19) cm
Gewicht	45 (± 18) kg	47 (± 19) kg	43 (± 18) kg
BMI	19,0 ($\pm 4,0$)	19,2 ($\pm 4,2$)	19,2 ($\pm 4,0$)

diesem Grund übersteigen Anzahl der Patienten und die Summe der prozentualen Anteile der Erkrankungen 160 Personen bzw. 100 Prozent.

Den Einschlusskriterien zu Folge nahmen insgesamt 160 Patienten an den Untersuchungen teil, die das Patientenkollektiv bildeten. Von diesen 160 entfallen 48,1% ($n = 77$) auf das männliche und 51,9% ($n = 83$) auf das weibliche Geschlecht. Die Altersspanne der männlichen Patienten erstreckt sich von 5,3 – 17,5 Jahren, die der weiblichen Patienten von 5,7 – 17,6 Jahren. Eine detailliertere Beschreibung des Gesamtkollektivs zeigt Tabelle 7.

3.2 Methoden

3.2.1 Untersucher

Die Röntgenbilder wurden von insgesamt vier verschiedenen Untersuchern ausgewertet.

- Untersucher 1: eine Fachärztin für Pädiatrie mit jahrzehntelanger Erfahrung in der pädiatrischen Radiologie
- Untersucher 2: ein Kinderradiologe mit mehrjähriger Erfahrung in der Methodik der pädiatrischen Knochenaltersbestimmung
- Untersucher 3: eine Fachärztin für Pädiatrie aus der Abteilung Pädiatrische Endokrinologie
- Untersucher 4: ein Medizinstudent (Doktorand, Autor der vorliegenden Arbeit)

Die Untersuchungen der Patienten mit dem Bonage-Ultraschallsystem wurden vom Doktoranden durchgeführt. Vor dem Beginn der Messungen erhielt der Doktorand eine Einweisung durch einen sachverständigen Vertreter der Firma, welche das Gerät entwickelt hat. Die Vorbereitung beinhaltete ebenso das Durchführen einiger Probenmessungen, um die Handhabung zu erlernen.

3.2.2 Struktur der Studie

Jeder der 160 Patienten unterzog sich zuerst einer Röntgenuntersuchung der linken Hand. Die Bestimmung des Knochenalters mittels Bonage-Systems erfolgte am gleichen Tag in derselben Abteilung. Das Knochenalter aller 160 Patienten wurde anhand der Handradiogramme von jedem Untersucher nach der Methode von Greulich & Pyle beurteilt. Die Untersucher waren bei der Beurteilung für das kalendarische Alter und das Ergebnis der sonographischen Knochenaltersbestimmung geblindet. Lediglich das Geschlecht der Patienten war bei der Beurteilung bekannt. Das Ergebnis jedes einzelnen Untersuchers wurde mit jedem Ergebnis der anderen Untersucher verglichen. Für den Vergleich der Methode nach Greulich & Pyle und der Ultraschall-Methode mit dem

Bonage-System wurden jeweils die Ergebnisse von Untersucher 1, 2 und 3 mit den sonographisch bestimmten Ergebnissen verglichen.

3.2.3 Bestimmung des Knochenalters anhand der Methode nach Greulich & Pyle

3.2.3.1 Natives Röntgen der linken Hand

Für die Bestimmung des Knochenalters wurde ein Radiogramm der Hand benötigt. Dafür ist die Aufnahme in nur einer Ebene, der dorso-palmaren Ebene, ausreichend. Die linke Hand des Patienten wird so auf der Röntgenkassette platziert, dass die Handfläche möglichst flach auf der Röntgenkassette aufliegt. Ein übermäßiges Abspreizen der Finger ist zu vermeiden. Bei der Einblendung ist darauf zu achten, dass die Röntgenstrahlen die distalen Abschnitte des Radius und der Ulna mit erfassen.

3.2.3.2 Gerätebeschreibung und Auswahl der Geräteparameter

Alle Röntgenbilder wurden im Institut für Diagnostische und Interventionelle Radiologie mit dem Radiographie-System MPG 80 (Filter: $1,0\text{mm}$ Aluminium und $0,1\text{mm}$ Kupfer) der Firma General Electric Medical Systems (Milwaukee, USA) angefertigt. Die an der Röntgenröhre angelegte Spannung betrug $45 - 46\text{kV}$, die Stromstärke 5mA . Der Film-Fokus-Abstand betrug 1m . Die Dosis bei einer Röntgenaufnahme der Hand beträgt ca. $0,5 - 0,75\text{cGy} \times \text{cm}^{-2}$. Diese verwendeten Geräteparameter sind Standard-einstellungen, die in der täglichen Routine im Institut zur Anwendung kommen. Die genaue Bezeichnung der für die Aufnahmen verwendeten Filme lautet Kodak Trimax Regula 400.

3.2.4 Bestimmung des Knochenalters mit dem Bonage-Ultraschallsystem

Die sonographische Bestimmung des Knochenalters erfolgte mit dem Bonage-System der Firma Sunlight Medical Ltd. (Tel Aviv, Israel), das als Prototyp für die Studie kostenfrei zur Verfügung gestellt wurde (siehe Abbildung 2). Mit dem Bonage-System



Abbildung 2: Das Bonage-Ultraschallsystem

lässt sich die Bestimmung des Knochenalters bei Kindern und Jugendlichen im Alter von 5 - 18 Jahren durchführen.

3.2.4.1 Gerätebeschreibung

Das Bonage-Ultraschallsystem besteht aus 2 unterschiedlichen Einheiten, dem Hauptmodul und dem Messelement: Das Hauptmodul enthält das Benutzer-Interface und die Software zur Kalkulation des Knochenalters. Das Messelement beinhaltet die Ultraschallsonde. Die Ultraschallsonde setzt sich aus den beiden Signalwandlern (Emitter und Receiver) zusammen, die räumlich getrennt an je einem Stativ befestigt sind. Die Stative lassen sich mechanisch aufeinander zu bewegen, bis Emitter und Receiver dem Handgelenk anliegen. Zwischen den Signalwandlern befindet sich eine Armschiene, auf welcher der Unterarm des Patienten bei der Untersuchung fixiert ist. Während der Messungen passieren die vom Emitter ausgesendeten Ultraschallwellen in Quertransmission die Messzone am Handgelenk des Patienten. Der Receiver empfängt die Ul-



Abbildung 3: Messzone am Handgelenk

traschallwellen nach der Passage der Messzone. In dieser Zone eingeschlossen sind die Ossifikationszentren (Epiphysen) von Ulna und Radius (Abbildung 3).

Anhand der gemessenen Schallleitungsgeschwindigkeit (speed of sound-Werte), der Distanz zwischen Sender und Empfänger, sowie einer geschlechts- und ethnizitätsspezifischen Datenbank - als Teil der Software - kalkuliert das Bonage-Ultraschallsystem das Knochenalter. Da das Bonage-System während der Messungen kein B-Bild erzeugt, kann der direkte Weg der Schallausbreitung im Knochen nicht nachvollzogen werden.

3.2.4.2 Spezifikationen des Bonage-Systems

Das System arbeitet mit einer Ultraschallfrequenz von 750 kHz. Die Abmessungen und das Gewicht des Systems betragen:

- $39\text{cm} \times 12\text{cm} \times 33\text{cm}$ für das 7kg schwere Hauptmodul
- $50\text{cm} \times 44\text{cm} \times 23\text{cm}$ für das 10kg schwere Messelement

Die vom Hersteller angegebene Präzision des Systems beträgt: $0,23 (\pm 0,15)$ Jahre für männliche und $0,24 (\pm 0,19)$ Jahre für weibliche Patienten.

3.2.4.3 Untersuchungsablauf

Vorbereitung Jeder neue Patient wird in eine neue Datei im Bonage-System aufgenommen, in welcher seine persönlichen (Name, Vorname, Geburtstag) und anthropometrischen (Körpergröße, Körpergewicht) Daten registriert werden. Das System ist vor und nach jeder Benutzung zu reinigen und zu desinfizieren. Die Messungen erfolgen am linken Arm des vor dem Bonage-System sitzenden Patienten. Der zu Untersuchende platziert seinen linken Arm in Pronationsstellung auf der Armschiene des Messelements. Der Untersucher palpiert dann den Processus styloideus ulnae am Handgelenk des Patienten und zeichnet mit einem Fettstift und einer speziell dafür vorgesehenen Schablone eine horizontale Markierung am Processus styloideus ulnae an. Ist der Processus markiert, umschließt der Patient mit der Faust einen Handgriff. Dieser Griff ist verstellbar an der Armschiene montiert und so auf die individuelle Armlänge einstellbar. Die Position des Armes muss so gewählt werden, dass sich das Handgelenk in der Nähe der Ultraschallsignalwandler befindet. Der Winkel zwischen Ober- und Unterarm sollte ca. 45° betragen. Im Anschluss wird Ultraschallgel auf beide Signalwandler (Emitter und Receiver) aufgetragen, um die Ankopplung zu ermöglichen. Die Signalwandler lassen sich anschließend über ein Bedienfeld am Messelement manuell exakt positionieren. Der Emitter, der sich auf der ulnaren Seite des Unterarmes befindet, wird so am Handgelenk des Patienten eingestellt, dass die Markierung am Emitter mit der Markierung über dem Processus styloideus ulnae genau übereinstimmt.

Messung Durch Betätigen des Start-Buttons bewegen sich Emitter und Receiver auf das Handgelenk zu, bis sie mit einem Anpressdruck von ca. 500g an beiden Seiten dem Handgelenk anliegen. An diesem Punkt findet die erste Messung statt. Ist diese beendet, weichen Emitter und Receiver automatisch auseinander, bewegen sich auf der palmar-dorsalen Ebene 2mm nach cranial und begeben sich in Warteposition. Solch ein Messzyklus dauert 20 Sekunden. Der nächste Messzyklus wird durch erneutes Betätigen des Start-Buttons initiiert. Die Position der Ultraschallsignalwandler während einer Messung wird in Abbildung 4 veranschaulicht. Um ein Ergebnis kalkulieren zu können sind mindestens 5-6 dieser Messzyklen erforderlich. Maximal 11 Messzyklen werden vom Bonage-System durchgeführt. Je größer die Anzahl der Messzyklen, desto höher ist die Präzision. Während der gesamten Messprozedur sollte sich der Patient möglichst



Abbildung 4: Darstellung des Bonage-Systems während eines Messzyklus

nicht bewegen, um die Messungen nicht durch eventuelle Positionsveränderung des Armes zu verfälschen. Nach dem letzten Zyklus bewegen sich die Signalwandler in ihre Ausgangsposition. Die Messungen sind beendet.

Ergebnis Durch Betätigen des Result-Buttons ermittelt die Software anhand der Messungen das Knochenalter des Patienten. Das Ergebnis wird in 2 Formen dargestellt:

- Angabe des Knochenalters in numerischer Form: in Jahr und Monat
- Angabe des Knochenalters als Messpunkt innerhalb einer Perzentilengrafik.

Die erhobenen Daten können gespeichert werden und ermöglichen so eine Verlaufskontrolle.

3.2.4.4 System-Qualitätsverifikation

Um sicherzustellen, dass das Bonage-System konstant exakte und fehlerfreie Messungen durchführt, soll regelmäßig eine System-Qualitätsprüfung erfolgen. Diese Überprüfung

ist vom Anwender bzw. dem Untersucher selbst durchzuführen und sollte einmal wöchentlich stattfinden. Der Systemcheck kontrolliert, ob die Messkomponenten (v.a. die Ultraschallsonden) fehlerfrei funktionieren. Für die System-Qualitätsprüfung hat die Firma Sunlight Medical Ltd. (Tel Aviv, Israel) ein Phantom entwickelt. Das Phantom ist aus einem Material gefertigt, von dem bekannt ist, mit welcher Geschwindigkeit sich der Ultraschall in diesem Material ausbreitet. Da die Geschwindigkeit des Ultraschalls in dem Material ebenso von dessen Temperatur abhängt, wurden 18 Referenzwerte für den Temperaturbereich von 14°C bis 31°C bestimmt. Anstelle des Armes wird das Phantom zwischen die Ultraschallsonden platziert und eine spezielle Messprozedur (system quality verification = SQV) am Bonage-System gestartet. Der Untersucher muss dem System lediglich die Temperatur des Phantoms vorgeben, die an einer Temperaturskala direkt am Phantom ablesbar ist. Bonage bestimmt dann den aktuellen SOS-Parameter des Phantoms und vergleicht ihn mit dem zu erwartenden Referenzwert. Weicht der gemessene Wert vom Referenzwert ab, so erscheint im Display des Systems eine Warnung. Bonage erzeugt ebenfalls eine Warnung, wenn die letzte Qualitätsprüfung zu lange zurückliegt.

3.2.5 Statistische Auswertung

Die Auswertung der erhobenen Daten erfolgte nach intensiver Beratung durch das Institut für Medizinische Dokumentation und Statistik der Friedrich-Schiller-Universität Jena. Für die statistische Analyse wurde das PC-Programm Statistical Package for the Social Sciences (SPSS, Version 12.0.1 für Windows) genutzt.

3.2.5.1 Messgenauigkeit des Bonage-Systems

Ein Maß für die Präzision = Validität der Messungen, die mit dem Bonage-System durchgeführt werden, stellt der Variationskoeffizient (VC) dar. Dieser Koeffizient zählt zu den Streuungsmaßen und beschreibt die Standardabweichung als prozentualen Anteil des Mittelwertes. Dabei geht man von der Annahme aus, dass die Variation von Daten bzw. Werten proportional zum Mittelwert ansteigt, dass also Standardabweichungen/ Varianzen im Prinzip automatisch um so größer werden können, je höher

die durchschnittlichen Messwerte liegen. In der Literatur wird der Variationskoeffizient auch als relative Standardabweichung bezeichnet. In der Regel wird der Variationskoeffizient in Prozent angegeben. Er berechnet sich, indem die Standardabweichung durch das arithmetische Mittel dividiert und der Quotient mit 100% multipliziert wird.

3.2.5.2 Bildung von Paaren

Für die Auswertung der Ergebnisse wurden Auswertungspaare gebildet. Um Unterschiede bzw. Übereinstimmungen in den Ergebnissen - nur für die Methode nach Greulich und Pyle - zu verifizieren, wurden folgende Paare gebildet, die in Tabelle 8 zusammengefasst sind.

Tabelle 8: Auswertepaare für die Knochenaltersbestimmung nach der Methode von Greulich & Pyle

Paar	Untersucher	Paar	Untersucher
1	U1 - U2	4	U2 - U3
2	U1 - U3	5	U2 - U4
3	U1 - U4	6	U3 - U4

Weitere Paare wurden für den Vergleich der sonographischen Knochenaltersbestimmung mit der Methode nach Greulich & Pyle gebildet. Diese sind in Tabelle 9 dargestellt. Die Ergebnisse von Untersucher 4 wurden für den Vergleich dieser beiden Methoden nicht berücksichtigt, da dieser Untersucher nicht über eine entsprechende Qualifikation auf dem Gebiet der Knochenaltersbestimmung verfügte.

Tabelle 9: Auswertepaare für den Vergleich von sonographischer und radiologischer Knochenaltersbestimmung

Paar	Untersucher
1	Bonage - U1
2	Bonage - U2
3	Bonage - U3

3.2.5.3 Korrelation und Bestimmtheitsmaß

Zunächst wurde für jedes Auswertungspaar der Korrelationskoeffizient nach Pearson ermittelt, um zu prüfen, ob zwischen den Ergebnissen eine signifikante Korrelation besteht. Dieser Koeffizient beschreibt die Enge des linearen Zusammenhangs zweier Merkmale. Der Korrelationskoeffizient wird angegeben in Form einer Maßzahl r , welche die Richtung und die Stärke des linearen Zusammenhangs beschreibt und Werte zwischen $+1$ und -1 annimmt. Nähert sich $r + 1$ oder -1 an, wird die lineare Abhängigkeit immer wahrscheinlicher. Erreicht $r = +1$ oder -1 liegt ein funktionaler linearer Zusammenhang vor. Man spricht dann von einem perfekt positiven bzw. einem perfekt negativen Zusammenhang. Nähert sich r der 0 , so wird ein linearer Zusammenhang der untersuchten Merkmale immer unwahrscheinlicher. Aus dem Korrelationskoeffizienten wurde durch Quadrieren von r das Bestimmtheitsmaß bm abgeleitet ($bm = r^2$). Das Bestimmtheitsmaß lässt sich interpretieren als Anteil der Streuung der Werte eines Merkmals Y , der durch die lineare Abhängigkeit von X erklärt werden kann. Das heisst: Ein $r = 0,9$ ergibt ein $r^2 = bm$ von $0,81$. Daraus resultiert, dass sich 81% der Y -Werte durch ihren linearen Zusammenhang zu X erklären lassen. Die restlichen 19% werden durch andere zufällige Einflüsse erklärt. Das Bestimmtheitsmaß nimmt Werte zwischen 0 und $+1$ an. Auch hier gilt, je näher das Bestimmtheitsmaß bm an 1 liegt, desto höher ist die Wahrscheinlichkeit des linearen Zusammenhangs. Allerdings ist ein bm nahe bei 1 kein Beweis für einen linearen Zusammenhang. Es wird nur ausgesagt, dass ein angenommener linearer Zusammenhang einen großen Anteil der Streuung der Y -Werte erklärt.

3.2.5.4 Differenzanalyse und Verteilung

Im Anschluss wurden die Differenzen zwischen den Ergebnissen der Auswertungspaare analysiert. Streubereich, Mittelwert und Standardabweichung wurden für jedes Paar bestimmt. Die Mittelwerte der Differenzen stellen ein Maß für die Reliabilität dar. Dieses Maß beurteilt die Konsistenz der Ergebnisse zwischen verschiedenen Untersuchern (Interbeurteiler-Reliabilität). Werden die Beträge der mittleren Differenzen verglichen, resultiert daraus eine Information über die Unterschiede zwischen den Ergebnissen der einzelnen Untersucher. Um Übereinstimmungen bzw. Unterschiede genauer zu evaluie-

ren, wurden die Differenzen jedes Auswertungs-paares mit dem t-Test für gepaarte bzw. abhängige Stichproben untersucht. Eine Voraussetzung für die Anwendung des t-Tests ist die Normalverteilung des untersuchten Merkmals. Diese Voraussetzung ist für die Untersuchungen in der vorliegenden Studie aufgrund des Grenzwerttheorems erfüllt. Das Grenzwerttheorem besagt, dass die Verteilung der Differenzmittelwerte bei zunehmendem Stichprobenumfang in eine Normalverteilung übergeht, deren Varianz proportional zum Stichprobenumfang klein wird (Bortz 1999, Nachtigal und Wirtz 1998). Bei Stichprobenumfängen von ≥ 30 kann bereits von einer Normalverteilung ausgegangen werden. Des Weiteren stellt der t-Test ein statistisches Verfahren dar, welches robust auf Voraussetzungsverletzungen reagiert (Bortz 1999). Das Signifikanzniveau α (entspricht dem Testwert p) wurde auf $\alpha \leq 0,05$ festgelegt. Das bedeutet, dass bei einhundert Tests die Nullhypothese, welche die Ungleichheit der zu vergleichenden Methoden ablehnt, fünfmal fälschlicherweise widerlegt werden kann. Bei $\alpha \leq 0,05$ wurde Ungleichheit angenommen und die Nullhypothese verworfen.

3.2.5.5 Statistische Methodik und Tests

Eine andere Methode, um eventuelle Übereinstimmungen bzw. Unterschiede auch graphisch darstellen zu können, wurde 1986 von Bland und Altman (Bland und Altman 1986) vorgestellt. Mit ihrem Verfahren lässt sich u.a. auch eine mögliche Falschinterpretation des Pearson'schen Korrelationskoeffizienten aufdecken. Um einen Bland-Altman-Plot zu erstellen, werden die Differenzen der Ergebnisse der angewendeten Methoden auf der Ordinate gegen das arithmetische Mittel der Differenzen auf der Abszisse abgetragen. In den Plot wird die mittlere Lage der Differenzen (mittlere Differenz = Mittelwert = \bar{x}) auf der Ordinate abgetragen. Ausgehend von dieser mittleren Differenz werden zwei weitere Grenzen eingezeichnet, die als limits of agreement bezeichnet werden. Die Lage der beiden Grenzen wird bestimmt, indem von der mittleren Differenz sowohl die zweifach positive als auch die zweifach negative Standardabweichung der inter-individuellen Unterschiede abgetragen wird ($\bar{x} \pm 2 \times s$). Des Weiteren wird das 95%-Konfidenzintervall des zu erwartenden intra-individuellen Unterschiedes im Plot gekennzeichnet. Die Grenzen des 95%-Konfidenzintervalls markieren graphisch jenes Intervall, in dem mit 95%-iger Wahrscheinlichkeit der Mittelwert der Differenzen liegt. Um so größer die Stichprobe ist, desto kleiner wird das Konfidenzintervall, da

sich die Lage des Mittelwertes mit größerer Sicherheit bestimmen lässt. Abbildung 5 demonstriert schematisch die Charakteristika eines Bland-Altman-Plots.

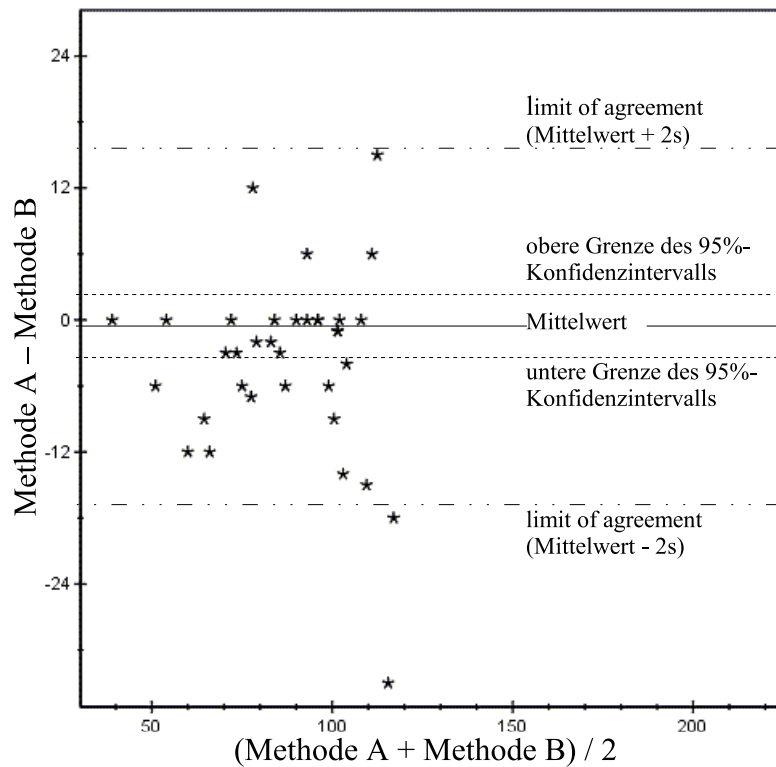


Abbildung 5: Schema eines Bland-Altman-Plots. Mit diesem Plot lassen sich Differenzen zwischen verschiedenen Methoden (bzw. Untersuchern) analysieren.

Der Bland-Altman-Plot bietet so verschiedene Informationen über den Grad der Übereinstimmung (Krummenauer und Doll 2000): Liegt die Nulllinie innerhalb des 95%-Konfidenzintervalls, ist auf dem Signifikanzniveau von 5% kein Unterschied zwischen den Methoden zu erwarten. Liegt die Nulllinie jedoch außerhalb des Konfidenzintervalls ist auf dem 5%-Niveau ein signifikanter Unterschied anzunehmen. Der Plot liefert ebenso eine Information über die Streuung der Differenzen und somit über das Ausmaß der Übereinstimmung der Methoden. Je enger der Bereich ist, in dem sich die Differenzen um ihre mittlere Lage (Mittelwert \bar{x}) streuen, desto größer ist die Übereinstimmung. Der Theorie der Normalverteilung zufolge sollten sich zudem 95,4% der Ergebnisse innerhalb des Bereiches befinden, der von den limits of agreement eingeschlossen wird.

Anhand des Bland-Altman-Plots lässt sich also erkennen, inwiefern Methode A der Methode B gleicht bzw. sich von ihr unterscheidet. Aber der Plot zeigt nicht, ob die

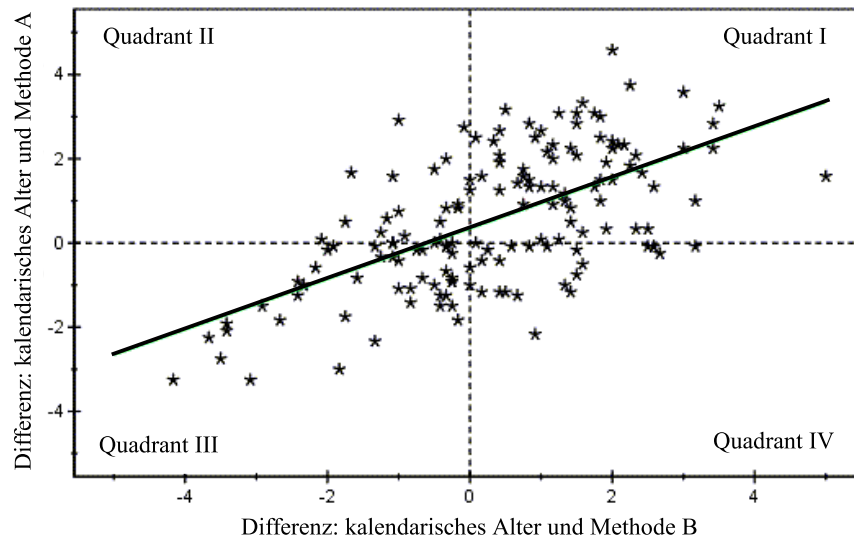


Abbildung 6: Schema eines Scatterplots. Mithilfe dieser Art des Plots werden die Tendenzen der zu vergleichenden Methoden beurteilbar.

Tendenz beider Methoden im Falle eines Irrtums dieselbe Richtung aufweist. Das heisst z.B. ob bei einem Patienten, bei dem mit Methode A ein akzeleriertes Knochenalter festgestellt wird, auch mit Methode B ein akzeleriertes Ergebnis diagnostiziert wird. In diesem Fall weisen beide Methoden eine übereinstimmende Tendenz auf. Gegensätzliche Tendenzen der Methoden bedeutet, dass z.B. eine Methode ein akzeleriertes die andere Methode hingegen ein retardiertes Knochenalter ergibt. Aus diesem Grund wurde für jedes Paar ein weiterer Plot erstellt, bei dem jeweils die Differenz aus kalendarischem Alter und Methode A gegen die Differenz aus kalendarischen Alter und Methode B abgetragen wurden. Die Abbildung 6 stellt allgemein diese Art eines Scatterplots dar. Durch die Nullpunkte beider Achsen werden 2 Geraden gelegt, die den Plot in 4 Quadranten unterteilen. Bei übereinstimmender Tendenz weisen die Differenzen das gleiche Vorzeichen auf, liegen also in den Quadranten I (positiv-positiv) oder III (negativ-negativ). Widersprechen sich die Tendenzen der Methoden liegen die Differenzen in den Quadranten II (negativ-positiv) oder IV (positiv-negativ). Im Idealfall einer exakten Übereinstimmung der Methoden verläuft die Regressionsgerade durch den Schnittpunkt der beiden Ursprungsgeraden und teilt die Quadranten I und III in 2 kongruente Dreiecke. Auf diese Weise wurden die Tendenzen der zu vergleichenden Methoden zur Knochenaltersbestimmung analysiert. Die Anzahl der Fälle, bei denen gegensätzlichen Tendenzen auftraten, wurde ermittelt.

Sofern ein Unterschied zwischen den Methoden ausgeschlossen werden konnte, wurde zusätzlich die Testung auf einen Präzisionsunterschied der beiden Methoden durchgeführt. Dafür wurde der Test nach Maloney und Rastogi genutzt. Analog zum t-Test wurde bei $\alpha \leq 0,05$ (entspricht Testwert $p \leq 0,05$) die Nullhypothese verworfen, Ungleichheit und ein daraus resultierender Präzisionsunterschied angenommen.

4 Ergebnisse

4.1 Präzision des Bonage-Systems

Um den Variationskoeffizienten für das Bonage-System zu ermitteln, wurde folgendes Procedere durchgeführt: An 5 Probanden wurde das Knochenalter an 3 unterschiedlichen Tagen mehrmals mit Bonage gemessen. Auf der Grundlage dieser Messungen wurden für jeden Probanden das arithmetische Mittel und die Standardabweichung ihrer Ergebnisse bestimmt. Aus diesen Werten wurde schließlich der Variationskoeffizient für jeden Probanden ermittelt. Eine Auflistung dieser Werte zeigt Tabelle 10.

Tabelle 10: Variationskoeffizient für das Bonage-System (s = Standardabweichung, VC = Variationskoeffizient)

Proband	Mittelwert des Knochenalters (in Monaten)	s (in Monaten)	VC (in %)
Proband 1	158,67	2,29	1,44
Proband 2	207,78	2,54	1,22
Proband 3	64,89	5,78	8,90
Proband 4	204,44	5,36	2,62
Proband 5	171,67	4,14	2,41
mittlerer VC			3,32

Der niedrigste Wert des Variationskoeffizienten beträgt 1,22% bei Proband 2, der höchste Wert 8,90% bei Proband 3. Alle Koeffizienten befinden sich in einem relativ engen Bereich von 1,22% – 2,62%, mit Ausnahme des Variationskoeffizienten, der sich für den Probanden 3 ergibt. Dieser übersteigt das Dreifache des höchsten Wertes der anderen 4 Probanden. Aus den verschiedenen Variationskoeffizienten der einzelnen Probanden ließ sich ein mittlerer Variationskoeffizient von 3,32% für das Bonage-System errechnen.

4.2 Evaluierung der Methode von Greulich & Pyle

Für 8 der 160 Patienten (entspricht 5% der Studienpopulation), die an der Studie teilnahmen, konnte mit dem Bonage-System kein Knochenalter bestimmt werden. Diese 8

Patienten wurden daher weder für den inter-individuellen Vergleich der einzelnen Untersucher noch für den Vergleich beider Methoden berücksichtigt. Aus diesem Grund umfasst die tatsächlich ausgewertete Studienpopulation 152 statt der ursprünglich geplanten 160 Patienten.

Die Untersucher 1, 2 und 4 ermittelten von jedem Patienten ($n = 152$) das Knochenalter nach der Methode von Greulich & Pyle. Untersucher 3 bestimmte das Knochenalter, ebenfalls nach der Methode von Greulich & Pyle, von 127 Patienten. Mit dem Bonage-System wurde das Knochenalter aller 152 Patienten bestimmt.

4.2.1 Ergebnisse der einzelnen Untersucher

In Tabelle 11 sind charakteristische Kenngrößen der Ergebnisse der Untersucher U1, U2, U3 und U4 zusammengestellt. Die Ergebnisse der Untersucher U1 bis U4 liegen in einem Altersbereich von 36 – 228 Monaten. Das kleinste von Untersucher 2 bestimmte Knochenalter beträgt 36 Monate, das größte 228 Monate. Im Vergleich mit den anderen Untersuchern findet sich bei U2 mit 192 Monaten der größte Streubereich. Die Ergebnisse der Knochenaltersbestimmungen von Untersucher 3 reichen von min. 39 bis max. 207 Monate. Der daraus resultierende Streubereich umfasst 168 Monate und ist der kleinste in dieser Gruppe. Die Mittelwerte der einzelnen Untersucher liegen in einem relativ engen Intervall von 0,7 Monaten (137,2 bis 137,9 Monate). Die kleinste Standardabweichung zeigt sich in den Ergebnissen von Untersucher 3 ($s = 37,7$ Monate), die größte in den Ergebnissen von Untersucher 1 ($s = 39,9$ Monate). Eine Gemeinsamkeit bei der Beurteilung des Knochenalters ist die Lage des Medians. Bei allen 4 Untersuchern liegt dieser bei 144 Monaten.

Um eine Aussage über die Genauigkeit der Standardmethode nach Greulich & Pyle zur Bestimmung des Knochenalters zu erhalten, wurden die Einzelergebnisse der Untersucher 1 bis 4 miteinander verglichen. Als Goldstandard bzw. als Maß für Richtigkeit dienten in dieser Studie die Ergebnisse von Untersucher 1, da dieser eine langjährige klinische Erfahrung auf dem Gebiet der radiologischen Knochenaltersbestimmung und speziell im Umgang mit der Methode von Greulich & Pyle besitzt. Um festzustellen, ob zwischen den Ergebnissen der Knochenaltersbestimmungen der verschiedenen

Tabelle 11: Individuelle Ergebnisse der Knochenaltersbestimmung der einzelnen Untersucher und des Bonage-Systems (alle Angaben in Monaten)

		U1	U2	U3	U4	Bonage
		(Monate)	(Monate)	(Monate)	(Monate)	(Monate)
n	(gesamt)	152	152	127	152	152
	(männlich)	73	73	60	73	73
	(weiblich)	79	79	67	79	79
Minimum		39	36	39	42	58
Maximum		228	228	207	222	215
Streubereich		189	192	168	180	157
Mittelwert	(gesamt)	137,2	137,6	137,6	137,9	139,9
	(männlich)	141,7	141,7	140,0	142,6	148,4
	(weiblich)	133,3	133,8	135,3	133,6	132,1
Standard- abweichung	(gesamt)	39,9	38,7	37,7	37,9	36,1
	(männlich)	45,6	45,2	43,3	42,3	39,9
	(weiblich)	33,5	31,2	32,0	33,1	30,4
Median		144	144	144	144	142

Untersucher überhaupt ein Zusammenhang besteht, wurde der Pearson'sche Korrelationskoeffizient für jedes Untersucherpaar ermittelt. Tabelle 12 zeigt die errechneten Korrelationskoeffizienten r und das daraus resultierende Bestimmtheitsmaß bm .

Tabelle 12: Korrelationskoeffizienten und Bestimmtheitsmaße der Untersucherpaare

Untersucherpaar	Korrelationskoeffizient r	Bestimmtheitsmaß bm
U1 - U2	0,979	0,958
U1 - U3	0,977	0,955
U1 - U4	0,985	0,970
U2 - U3	0,980	0,960
U2 - U4	0,973	0,947
U3 - U4	0,977	0,955

Der größte Wert für den Korrelationskoeffizienten $r = 0,985$ zeigte sich für das Untersucherpaar U1 und U4. Ausnahmslos alle Korrelationskoeffizienten nahmen Werte an, die $> 0,970$ sind. Den kleinsten Korrelationskoeffizienten wies das Untersucherpaar U2/ U4 auf, bei welchem er $0,973$ betrug. Das Bestimmtheitsmaß verhält sich analog zum Korrelationskoeffizienten, da es sich aus diesem ableitet. Aus dieser Kausalität folgt, dass der größte Wert ebenfalls bei Untersucherpaar U1/ U4 ($bm = 0,970$) und der kleinste Wert bei dem Paar U2/ U4 ($bm = 0,947$) auftrat.

Um die Ergebnisse der Knochenaltersbestimmung der verschiedenen Untersucher genauer analysieren zu können, wurden die Differenzen zwischen den Einzelergebnissen der Untersucher gebildet. Es sollte ein möglichst präziser Vergleich erzielt werden. Aus diesem Grund wurden die Ergebnisse jedes einzelnen Untersuchers den Ergebnissen der anderen Untersucher gegenübergestellt. In Abbildung 7 sind diese Differenzen in Form von Boxplots dargestellt. Des Weiteren sind in der Abbildung die mittleren Differenzen, die Standardabweichungen und die p-Werte des t-Tests zusammengestellt.

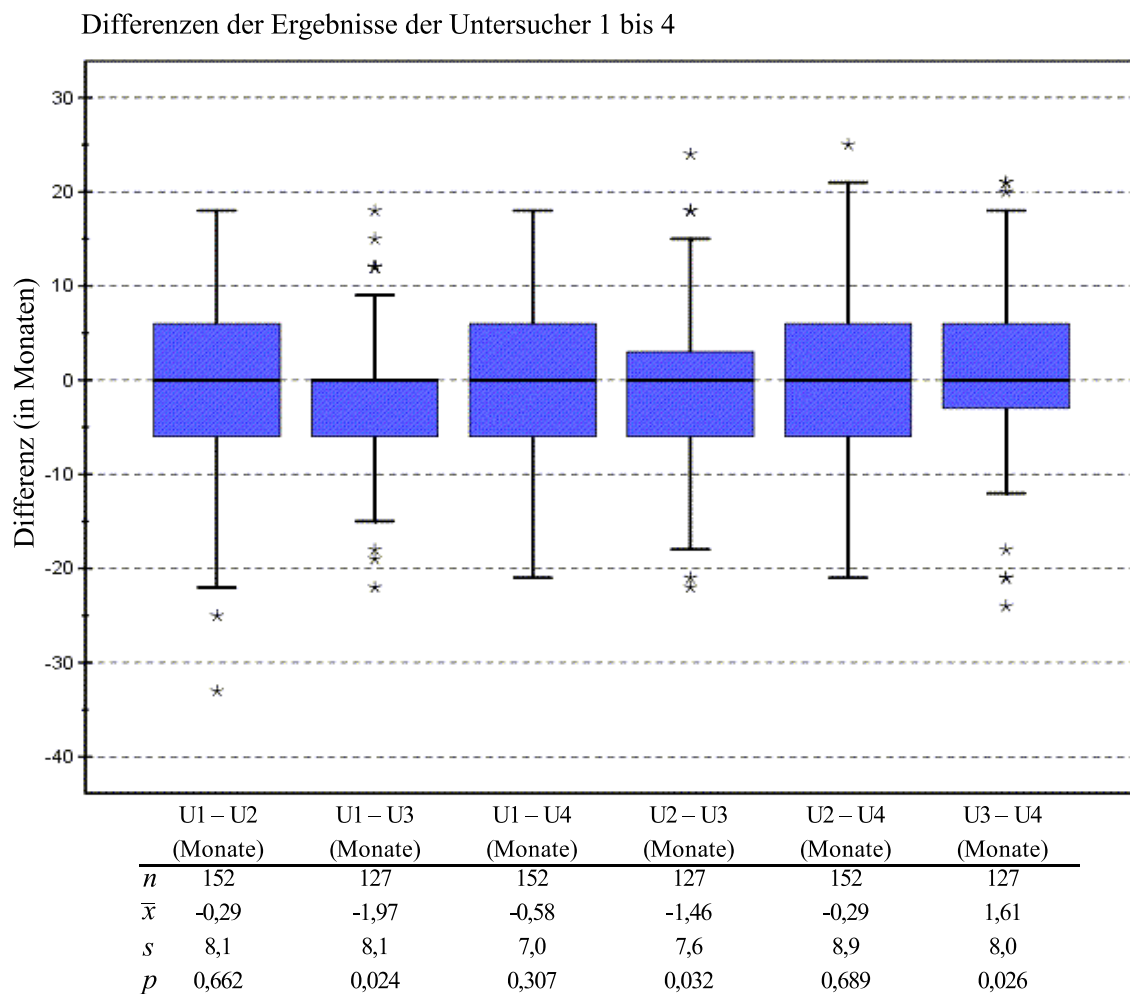


Abbildung 7: Boxplots der Differenzen der verschiedenen Untersucherpaare. Angegeben sind die Mittelwerte \bar{x} , Standardabweichungen s und Testwerte p . Die Mittelwerte der Differenzen \bar{x} entsprechen dabei der Interbeurteiler-Variabilität mit der zugehörigen Standardabweichung s . Ausreißer wurden mit Sternchen markiert.

Alle Differenzen, die bei dem Vergleich der 6 Auswertungspaare errechnet wurden, lagen innerhalb eines Bereiches von -22 bis 21 Monaten. Einzelne Differenzen erreichten Werte außerhalb dieser Grenzen, wurden aber als Ausreißer identifiziert. So resultierte

die größte negative Differenz ($\Delta = -33$ Monate) aus dem Vergleich der Ergebnisse von Untersucher 1 mit den Ergebnissen von Untersucher 2. Die größte positive Differenz ($\Delta = 25$ Monate) erschien bei Auswertungspaar U3/ U4. Bei Betrachtung der einzelnen Boxplots fällt auf, dass die Auswertungspaare an denen Untersucher 3 beteiligt war, deutlich geringere Streubereiche aufwiesen als die anderen Paarungen. So betrug der Streubereich des Untersucherpaares U1/ U3 nur 24 Monate; bei dem Paar U3/ U4 betrug der Streubereich 30 Monate und 33 Monate bei dem Untersucherpaar U2/ U3. Den größten Streubereich aller Vergleichsgruppen zeigte das Paar U2/ U4 mit 42 Monaten.

Eine weitere Beobachtung lässt sich bei den Paarungen, an denen U3 beteiligt war, konstatieren. Obwohl die Streubereiche sichtlich kleiner waren, nahmen die Standardabweichungen dieser Paare ähnliche Werte wie die der Vergleichsgruppen mit größerem Streubereich an. Die Standardabweichung des Untersucherpaares U1/ U3 betrug 8,1 Monate bei einem Streubereich von 24 Monaten und war somit genauso groß wie die Standardabweichung des Untersucherpaares U1/ U2 bei einem Streubereich von 40 Monaten. Die größte Streuung der Differenzen trat zwischen Untersucher 2 und Untersucher 4 auf. Sie betrug für dieses Paar 8,9 Monate. Die geringste Streuung ergab sich für das Paar U1/ U4 ($s = 7$ Monate).

Ein Vergleich der Mittelwerte als Maß für die Interbeurteiler-Variabilität veranschaulichte, dass sich die mittleren Differenzen in einem Bereich zwischen $-1,97$ und $1,61$ Monaten befanden. Die mittlere Differenz für die Paare U1/ U2 und U2/ U4 betrug $-0,29$ Monate und stellte gleichzeitig die kleinste im Vergleich der verschiedenen Untersucher dar. Der t-Test für gepaarte Stichproben wies für beide Auswertungspaare keine Signifikanzen nach und bestätigte die Nullhypothese: Für das Paar U1/ U2 ergab sich als Ergebniss des t-Tests ein p -Wert von $0,662$ bei 151 Freiheitsgraden. Für das Paar U2/ U4 wurde gleichfalls bei einem Freiheitsgrad von 151 ein p von $0,689$ errechnet. Im Vergleich von U1 und U4 betrug die mittlere Differenz $-0,58$ Monate. Auch für dieses Paar wurde wiederum bei 151 Freiheitsgraden mit einem p -Wert von $0,307$ die Nullhypothese bestätigt und das Bestehen von Signifikanzen abgelehnt. An dieser Stelle wurde ebenfalls deutlich, dass sich die drei Auswertungspaare, die U3 beinhalteten, klar von den anderen unterschieden. Der t-Test lieferte zu all jenen Paaren

zugehörige p -Werte, die eine deutliche Signifikanz der Differenz in den Ergebnissen der Untersucher belegten (siehe Abbildung 7).

Anhand eines Bland-Altman-Plots lassen sich die Unterschiede sehr gut graphisch veranschaulichen. Die zwei in Abbildung 8 dargestellten Plots für die Auswertungspaare U1/ U3 und U1/ U4 sollen stellvertretend für die Plots aller anderen Paare stehen, bei denen die graphische Auswertung mittels eines Bland-Altman-Plots analog erfolgte.

In der Gegenüberstellung beider Plots lässt sich der Unterschied in den Ergebnissen beider Untersucherpaare erkennen. Im Plot von U1/ U3 liegt die Obergrenze des 95%-Konfidenzintervalls der mittleren Differenz bei $-0,6$ Monaten, die Untergrenze bei $-3,4$ Monaten. Die Nulllinie liegt demnach nicht innerhalb dieses 95%-Konfidenzintervalls und somit besteht eine Signifikanz auf dem 5%-Niveau. Die Nullhypothese muss abgelehnt werden. Dieses Ereignis trat immer in gleicher Weise bei den Paaren auf, in die U3 miteinbezogen war. Bei allen Paaren von U3 mit anderen Untersuchern lag die Nulllinie außerhalb des 95%-Konfidenzintervalls und die Bedingungen für die Beibehaltung der Nullhypothese konnten nicht erfüllt werden. Im Plot von U1/ U4 hingegen liegt die Obergrenze des Intervalls bei $1,7$ Monaten und die Untergrenze bei $-0,5$ Monaten. Die Nulllinie liegt folglich mitten im 95%-Konfidenzintervall. Im Vergleich dieser beiden Untersucher sind keine Signifikanzen auf dem 5%-Niveau zu erwarten. Die Plots der Paare U1/ U2 und U2/ U4 zeigten ebenfalls Konfidenzintervalle, die die Nulllinie miteinschlossen und so die Nullhypothese bestätigten.

Die Bland-Altman-Plots beider Untersucherpaare zeigen auch Differenzen, die außerhalb der limits of agreement liegen. Solche außerhalb gelegenen Differenzen kamen bei jedem Untersucherpaar vor. Bei allen Paarungen von U3 mit anderen Untersuchern befanden sich jeweils $6,3\%$ aller Differenzen jenseits dieser Grenzen. Im Plot der Paare U1/ U2 und U1/ U4 lagen $5,3\%$ der Differenzen außerhalb der Grenzen. Nur $3,3\%$, der prozentual geringste Anteil, aller Differenzen von Paar U2/ U4 lagen nicht innerhalb der Grenzen. Alle diese Werte befinden sich nahe dem statistisch zu erwartenden Wert, der $4,6\%$ beträgt.

Durch die graphische Darstellung im Bland-Altman-Plot konnte eine weitere Analyse der Differenzen zwischen den Untersuchern erfolgen. Jede einzelne Differenz ist im

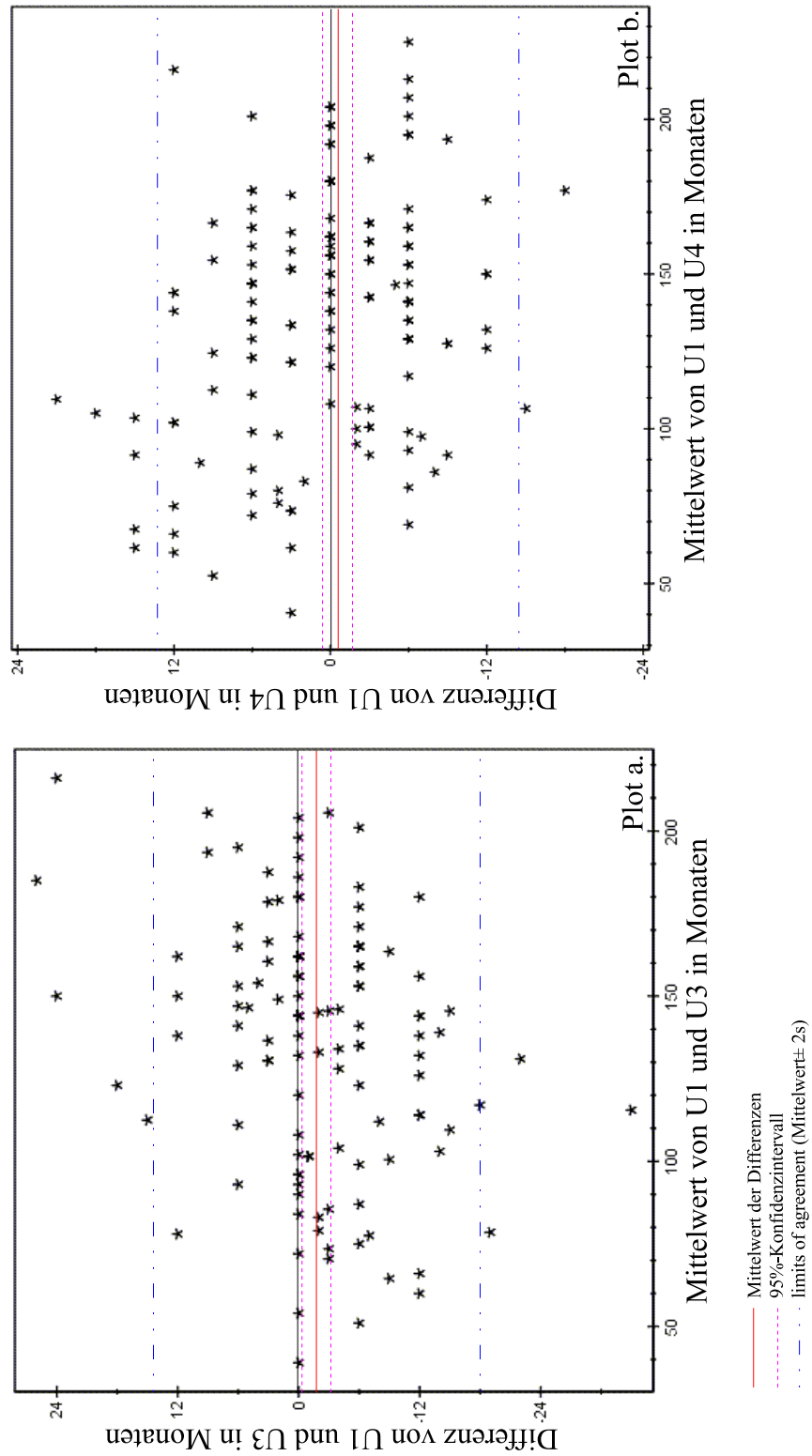


Abbildung 8: Bland-Altman-Plots zweier Untersucherpaare

Plot eingezeichnet. So lässt sich anhand ihrer Lage im Plot die Größe der Differenz direkt aus der Graphik ablesen (Tabelle 13). Aufgrund dieser Informationen ließ sich die Streuung der Differenzen um die Nulllinie, als Maß für die Übereinstimmung der einzelnen Untersucher, ebenfalls gut aus dieser Darstellung heraus beurteilen. Bei allen Paaren, die in Tabelle 13 beschrieben sind, entfielen 57,9% (U2/ U4)- 77,6% (U1/ U4) der Differenzen auf ein Intervall, das ± 6 Monate beträgt. Über 89% der Differenzen, die sich zwischen den Ergebnissen der Untersucher ergaben, lagen in einem Intervall von ± 12 Monaten. Bei Auswertungspaar U1/ U4 lagen bereits 94,7% in dem Bereich von ± 12 Monaten. In das Intervall von ± 18 Monaten fielen stets über 95% der Differenzen. Bei 3 von 6 Paarungen traten auch Differenzen auf, die größer als 24 Monate waren.

Tabelle 13: Streuung und Lage der Differenzen aller Untersucherpaare

Intervalle (in Monaten)	U1 - U2		U1 - U3		U1 - U4	
	Anzahl	%	Anzahl	%	Anzahl	%
± 6	109	71,7	93	73,2	118	77,6
$ 7 - 12 $	33	21,7	22	17,3	26	17,1
$ 13 - 18 $	7	4,6	7	5,5	7	4,6
$ 19 - 24 $	1	0,7	4	3,2	1	0,7
$> 24 $	2	1,3	1	0,8	0	0,0
Σ	152	100,0	127	100,0	152	100,0

Intervalle (in Monaten)	U2 - U3		U2 - U4		U3 - U4	
	Anzahl	%	Anzahl	%	Anzahl	%
± 6	97	76,4	88	57,9	86	67,7
$ 7 - 12 $	20	15,7	48	31,6	31	24,4
$ 13 - 18 $	7	5,5	13	8,5	4	3,2
$ 19 - 24 $	3	2,4	2	1,3	6	4,7
$> 24 $	0	0,0	1	0,7	0	0,0
Σ	127	100,0	152	100,0	127	100,0

Konnten für die einzelnen Untersucher keine signifikanten Unterschiede ihrer Ergebnisse belegt werden, ließ sich zusätzlich mit dem Test nach Maloney & Rastogi die Präzision der Greulich & Pyle-Methode beurteilen. Bei den Untersucherpaaren U1/ U2, U1/ U4 und U2/ U4 ließ sich der Test sinnvoll anwenden, da zwischen diesen Untersuchern keine Signifikanzen nachzuweisen waren. Ein Unterschied in der Präzision der Untersucher konnte nur für U1/ U4 nachgewiesen werden ($p = 0,001$). Für das Paar U1/ U2 betrug $p = 0,07$ und für das Paar U2/ U4 war $p = 0,314$. Diese beiden p -Werte ließen für diese Paarungen nicht auf einen Präzisionsunterschied der Methode

von Greulich & Pyle schließen. Da sich die Ergebnisse von Untersucher 3 signifikant von den Ergebnissen aller anderen Untersucher unterschied, konnte der Test in diesen Fällen nicht sinnvoll angewendet bzw. interpretiert werden.

4.2.2 Ergebnisse des Bonage-Systems

Die Ergebnisse des Bonage-Systems (siehe 11) lagen innerhalb eines Streubereiches, der von 58 bis 215 Monaten reicht. Folglich betrug die Spannweite des Streubereiches 157 Monate. Im Vergleich mit den Untersuchern U1 bis U4 zeigte das Bonage-System damit den kleinsten Streubereich. Der Mittelwert der Ergebnisse von Bonage ($\bar{x} = 139,9$ Monate) lag über den Mittelwerten der anderen Untersucher. Betrachtet man den Mittelwert bezogen auf das Geschlecht, so fällt auf, dass bei der Beurteilung männlicher Probanden der Mittelwert von Bonage ($\bar{x} = 148,4$ Monate) deutlich größer ist, als die Mittelwerte von U1, U2, U3 und U4. Bei der Knochenaltersbestimmung weiblicher Probanden lag das ermittelte Knochenalter im Mittel bei $\bar{x} = 132,1$ Monaten und war somit kleiner als die Ergebnisse der anderen Untersucher. Die Standardabweichung für die Ergebnisse der Ultraschallmethode war mit 36,1 Monaten die kleinste Standardabweichung der Einzelergebnisse. Der Median lag bei 142 Monaten und somit 2 Monate unter dem Median der Untersucher U1 bis U4.

4.3 Vergleich der Greulich-Pyle-Methode mit Bonage

Der Pearson'sche Korrelationskoeffizient wurde erneut bestimmt, um zu eruieren, ob ein Zusammenhang zwischen den Ergebnissen der Untersucher besteht. Das Bestimmtheitsmaß wurde gleichfalls bestimmt. Die ermittelten Koeffizienten sind aus Tabelle 14 abzulesen.

Tabelle 14: Korrelationskoeffizienten und Bestimmtheitsmaße der Bonage-Paare

Untersucherpaar	Korrelationskoeffizient r	Bestimmtheitsmaß bm
Bonage - U1	0,902	0,814
Bonage - U2	0,920	0,846
Bonage - U3	0,912	0,832

Aus der Tabelle geht hervor, dass die Korrelationen zwischen Bonage und den Untersuchern wesentlich kleiner war als die Korrelationen zwischen den einzelnen Untersuchern untereinander (siehe Tabelle 12). Die Werte der Korrelationskoeffizienten überstiegen nicht 0,920 und befanden sich deutlich unter 0,970. Noch nachdrücklicher zeigte sich dies am Bestimmtheitsmaß. Im Gegensatz zu den Werten, die sich aus den Paarungen von U1, U2 und U3 ergaben und die alle über 0,950 lagen (siehe Tabelle 12), erreichten die Bonage-Paare nur Werte bis zu 0,846.

Um die Methode von Greulich & Pyle direkt mit der neuentwickelten Ultraschallmethode vergleichen zu können, wurden wiederum Paare gebildet und die Differenzen der Ergebnisse analysiert. Die Ergebnisdifferenzen der 3 Paare sind in den Boxplots in Abbildung 9 dargestellt. Die Abbildung zeigt ebenfalls die mittleren Differenzen, die Standardabweichungen und die p -Werte des zugehörigen t-Tests.

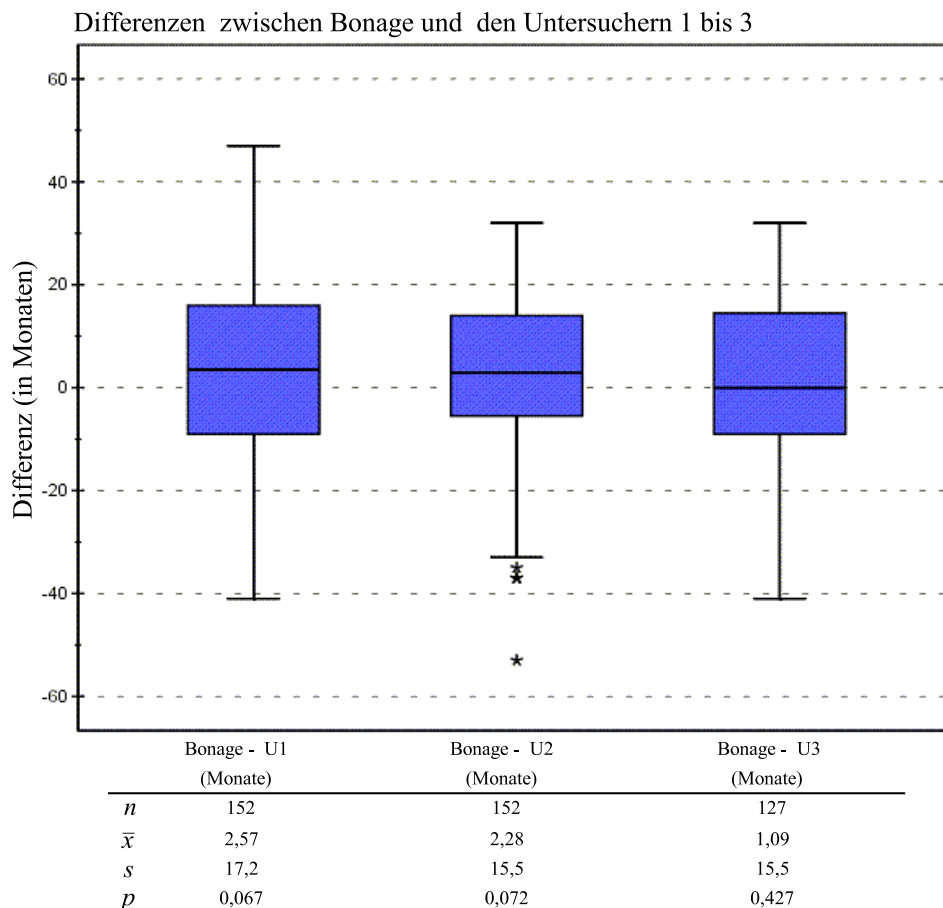


Abbildung 9: Boxplots der Differenzen der Bonage-Paare. Angegeben sind die Mittelwerte \bar{x} , Standardabweichungen s und Testwerte p . Die Mittelwerte der Differenzen \bar{x} entsprechen dabei der Interteilervariabilität mit der zugehörigen Standardabweichung s . Ausreißer wurden mit Sternchen markiert.

Die Differenzen, die sich zwischen den Ergebnissen von Bonage und den Resultaten der Untersucher ergaben, lagen alle in einem Intervall zwischen -41 und 47 Monaten. Bei dem Paar Bonage/ U2 trat eine Differenz von -53 Monaten auf, die aber aus statistischer Sicht ebenfalls einen Ausreißer darstellt. Vergleicht man den Streubereich der Einzelergebnisse der Untersucher U1 bis U4 (-21 bis 22 Monate) mit dem Streubereich der Bonage-Paare (-41 bis 47 Monate), so lässt sich erkennen, dass letzterer mehr als doppelt so groß ist.

Das Auswertungspaar mit der größten Streuung ist Bonage/ U1, für welches sich eine Standardabweichung von $17,2$ Monaten errechnen ließ. Sowohl für das Paar Bonage/ U2 als auch für Bonage/ U3 resultierte eine Standardabweichung von $15,5$ Monaten.

Die arithmetischen Mittel der Differenzen betragen $2,57$ Monate für das Paar Bonage/ U1, $2,28$ Monate für das Paar Bonage/ U2 und $1,09$ Monate für das Paar Bonage/ U3. Damit erreichten die mittleren Differenzen von Bonage/ U1 und Bonage/ U2 das Zweifache des Mittelwertes, der sich für das Paar Bonage/ U3 ergab. Vergleicht man die Mittel der Differenzen der Bonage-Paare mit den Differenzen der Paare aus Abbildung 7, lässt sich feststellen, dass die Paare Bonage/ U1 bzw. Bonage/ U2 die größten Differenzen aufzeigten. Dennoch konnte der t-Test die Nullhypothese, welche einen Unterschied der Methoden ablehnt, auf dem 5%-Niveau für alle 3 Bonage-Paare bestätigen. Die p -Werte für die Paare Bonage/ U1 und Bonage/ U2 betragen bei 151 Freiheitsgraden im t-Test $0,067$ und $0,072$. Bei 126 Freiheitsgraden für das Paar Bonage/ U3 war $p = 0,427$. Für kein Auswertungspaar ließen sich signifikante Unterschiede in den Ergebnissen identifizieren. Um zu testen, ob die Methode von Greulich & Pyle und die Ultraschallmethode Unterschiede in ihrer Präzision aufwiesen, kam noch einmal der Maloney-Rastogi-Test zur Anwendung. Dieser stellte eindeutig Signifikanzen fest. Die p -Werte für Bonage/ U1 bzw. Bonage/ U2 betragen $0,006$ bzw. $0,038$. Nur für das Paar Bonage/ U3 konnte kein Präzisionsunterschied beobachtet werden ($p = 0,153$).

Wiederum wurde für jedes Bonage-Paar ein Bland-Altman-Plot erstellt. In Abbildung 10 sind beispielhaft 2 Bland-Altman-Plots dargestellt, welche den Vergleich dieser beiden Paare graphisch veranschaulichen.

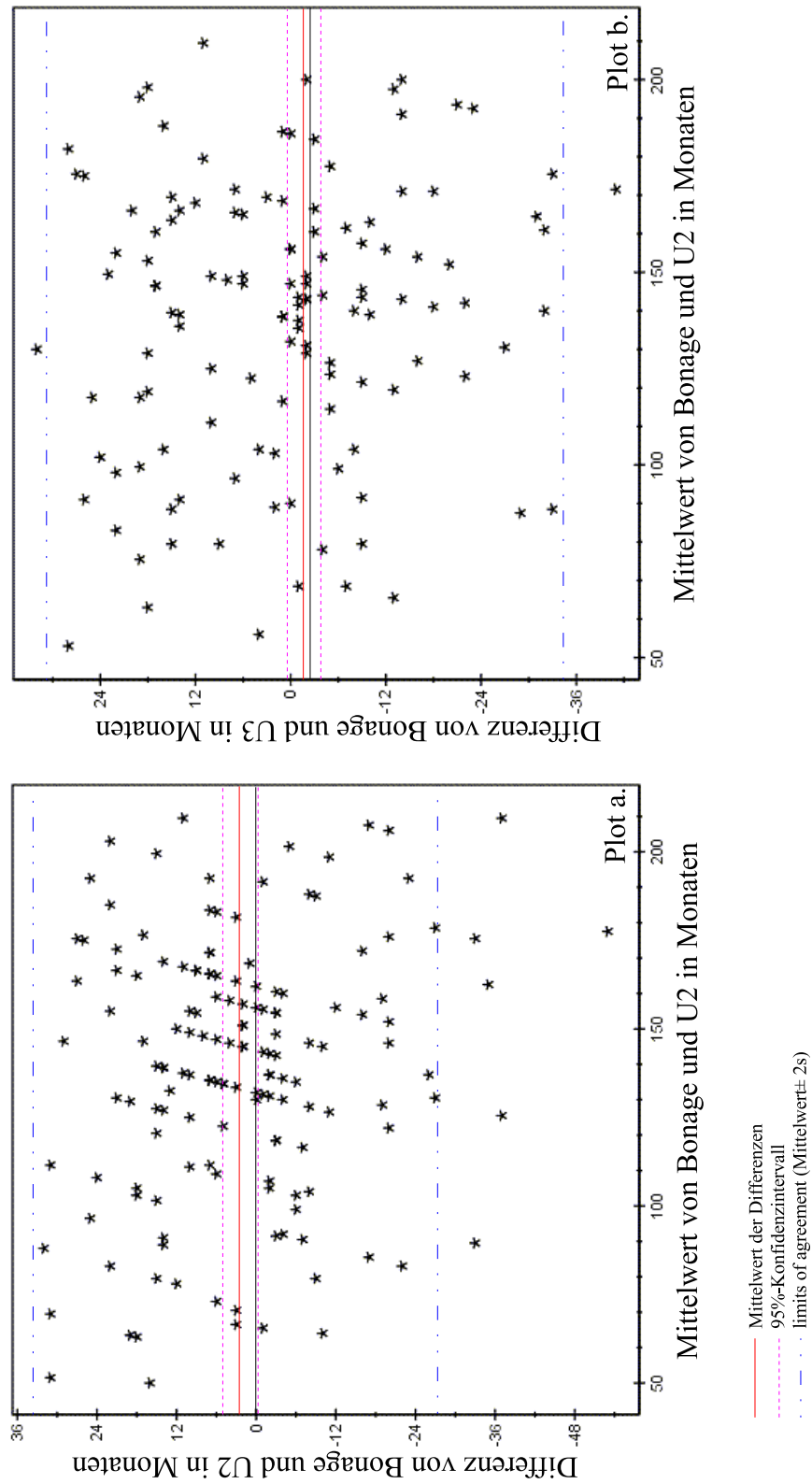


Abbildung 10: Bland-Altman-Plots zweier Bonage-Paare

In allen 3 erstellten Plots sind jeweils der Mittelwert der Differenzen, das 95%-Konfidenzintervall dieser mittleren Differenz und die limits of agreement eingezeichnet. In allen 3 Plots liegt die Linie, die den Nullpunkt auf der Y-Achse schneidet, innerhalb des Konfidenzintervalls. Es lässt sich erkennen, dass die Lage des Konfidenzintervalls mit den p -Werten des t-Tests korreliert und diese bestätigt. Je näher sich der p -Wert dem festgesetzten Signifikanzniveau von $\alpha = 0,05$ nähert, desto näher liegt die Nulllinie an einer der Grenzen des Konfidenzintervalls. In dem Plot, der den Vergleich des Auswertungspaares Bonage/ U3 beschreibt, liegt die Nulllinie am weitesten im Konfidenzintervall. Der zugehörige p -Wert ($p = 0,153$) war auch der größte Wert der 3 Paare dieser Gruppe. Das entgegengesetzte Bild veranschaulicht der Plot des Paares Bonage/ U1. Dieses Paar zeigte den niedrigsten p -Wert ($p = 0,067$) in dieser Gruppe und dementsprechend besitzt die Nulllinie den geringsten Abstand zur Grenze des Konfidenzintervalls. Der prozentuale Anteil der Einzelmessungen, der sich außerhalb der limits of agreement befindetet, ist mit 3,9% bei dem Paar Bonage/ U2 der geringste. Mit 5,5% bei Bonage/ U3 hat dieses Paar die größte Anzahl von Messergebnissen, die außerhalb der limits of agreement liegen. Damit entspricht der prozentuale Anteil der Differenzen, die außerhalb der limits of agreement liegen, dem statistischen Erwartungswert.

Wiederum konnten die Differenzen durch ihre Lage im Bland-Altman-Plot genauer analysiert und ihre Streuung um die Nulllinie, als Maß für die Übereinstimmung der beiden zu vergleichenden Methoden, beurteilt werden. Tabelle 15 gibt eine Zusammenfassung über die Lage der Differenzen, die aus den erstellten Bland-Altman-Plots zu entnehmen war. Im Intervall ± 6 Monate befanden sich 22,4% (Bonage/ U1) - 35,5% (Bonage/ U2) der Differenzen, im Intervall ± 12 Monate waren es 48,7% (Bonage/ U1) - 59,2% (Bonage/ U2). 83,6% (Bonage/ U1) - 88,2% (Bonage/ U3) aller Differenzen lagen in dem Bereich ± 24 Monate. Überdies traten bei allen 3 Paaren Differenzen auf, die größer als 24 Monate waren (Bonage/ U1: 16,4%; Bonage/ U2: 12,5%; Bonage/ U3: 11,8%). Zu einem geringen Prozentsatz (Bonage/ U1: 3,2%; Bonage/ U2: 2,0%; Bonage/ U3: 0,8%) errechneten sich Differenzen, die überdies außerhalb des Intervalls ± 36 Monate lagen.

Tabelle 15: Streuung und Lage der Differenzen aller Bonage-Paare

Intervalle (in Monaten)	Bonage - U1		Bonage - U2		Bonage - U3	
	Anzahl	%	Anzahl	%	Anzahl	%
± 6	34	22,4	54	35,5	43	33,9
$ 7 - 12 $	40	26,3	36	23,7	24	18,9
$ 13 - 18 $	34	22,4	24	15,8	30	23,6
$ 19 - 24 $	19	12,5	19	12,5	15	11,8
$ 25 - 30 $	12	7,9	9	5,9	8	6,3
$ 31 - 36 $	8	5,3	7	4,6	6	4,7
$> 36 $	5	3,2	3	2,0	1	0,8
Σ	152	100,0	152	100,0	127	100,0

4.4 Analyse der Tendenzen

Die Abbildungen 11 und 12 zeigen beispielhaft die Scatterplots der Paare U2/ U3 und Bonage/ U3. Für jedes Auswertungspaar wurde solch ein Plot erstellt. Anhand der abgebildeten Plots kann graphisch veranschaulicht werden, wie sich die Tendenzen der Knochenaltersbestimmungen mit der Atlasmethode mit den Tendenzen der Knochenaltersbestimmungen mit dem Bonage-System vergleichen ließen. In Abbildung 11 liegen 97 der 127 (entspricht 76,4%) Knochenaltersbestimmungen in den Quadranten I und III. In Bezug auf das kalendarische Alter des Patienten haben in diesen 97 Fällen beide Methoden übereinstimmend entweder ein akzeleriertes oder ein retardiertes Knochenalter bestimmt und wiesen somit die gleiche Tendenz auf. In 24 von 127 Fällen (entspricht 18,9%), diese liegen in den Quadranten II und IV, haben Atlasmethode und Bonage widersprüchliche Ergebnisse erbracht. Hat bei einem Patienten die Atlasmethode ein tendenziell akzeleriertes (bzw. retardiertes) Knochenalter bestimmt, so ergab die Knochenaltersbestimmung mit Bonage gegensätzlich ein retardiertes (bzw. akzeleriertes) Knochenalter. In diesen 24 Fällen zeigten beide Methoden unterschiedliche Tendenzen. Der Scatterplot erlaubt allerdings keine Aussage darüber, welche der beiden Methoden die genauere, respektive die ungenauere, Knochenaltersbestimmung erbringt. Die übrigen 6 Fälle (entspricht 4,7%) ließen sich keinem Quadranten exakt zuordnen, da ihre Messpunkte auf einer der beiden Ursprungsgeraden liegen, die die Quadranten gegeneinander abgrenzen. Der Plot aus der Abbildung 12, sowie alle anderen erstellten Scatterplots, wurden analog ausgewertet und sind der Tabelle 16 zu entnehmen.

Tabelle 16: aufgetretene Tendenzen aller Untersucherpaare

	Quadrant I+III (übereinstimmende Tendenzen)		Quadrant II+IV (gegensätzliche Tendenzen)	
	Anzahl	%	Anzahl	%
Bonage - U1	106	69,8	37	24,3
Bonage - U2	117	77,0	29	19,1
Bonage - U3	97	76,4	24	18,9
U1 - U2	132	86,8	13	8,6
U1 - U3	106	83,5	14	11,0
U1 - U4	129	84,9	16	10,5
U2 - U3	111	87,4	12	9,4
U2 - U4	132	86,8	15	9,9
U3 - U4	105	82,7	16	12,6

Das Paar, bei dem es die meisten Übereinstimmungen gab, wird von U2 und U3 gebildet. Von 127 Ergebnissen der Knochenaltersbestimmungen wiesen 111 (87,4%) dieselbe Tendenz auf. Das Paar, dessen Ergebnisse den geringsten Anteil gleicher Tendenzen zeigten, ist das Paar Bonage/ U1 (69,8% der Ergebnisse). Vergleicht man die Paare, die sich aus den Untersuchern 1 bis 4 ergeben mit den Bonage-Paarungen ist ein eindeutiger Trend erkennbar. Während die Bonage-Paare nur Übereinstimmungen von 69,8 – 77,0% erreichten, lagen die Übereinstimmungen der Paare von U1 bis U4 zwischen 82,7% und 87,4%. Der Anteil der gegensätzlichen Tendenzen, die in den Ergebnissen aller Paare identifiziert wurden, schwankte ebenfalls deutlich. Der geringste Anteil fand sich bei U1/ U2 (8,6%), der höchste Anteil gegensätzlicher Tendenzen bei Bonage/ U1. Aus der Tabelle 16 geht wiederum ein Trend hervor, der erkennen lässt, dass die Bonage-Paare mit 18,9%, 19,1% und 24,3% die meisten Fälle zeigten, in denen die Tendenzen der Knochenaltersbestimmung unterschiedliche Richtungen aufwiesen. Im Gegensatz dazu existierten 3 Paare (U1/ U2, U2/ U3 und U2/ U4) bei denen dieser Anteil unter 10,0% lag. Die Abbildungen 11 und 12 sollen beispielhaft die graphische Veranschaulichung der Scatterplots der Paare U2/ U3 und Bonage/ U3 darstellen, anhand derer sich die Tendenzen der Knochenaltersbestimmungen erkennen ließen.

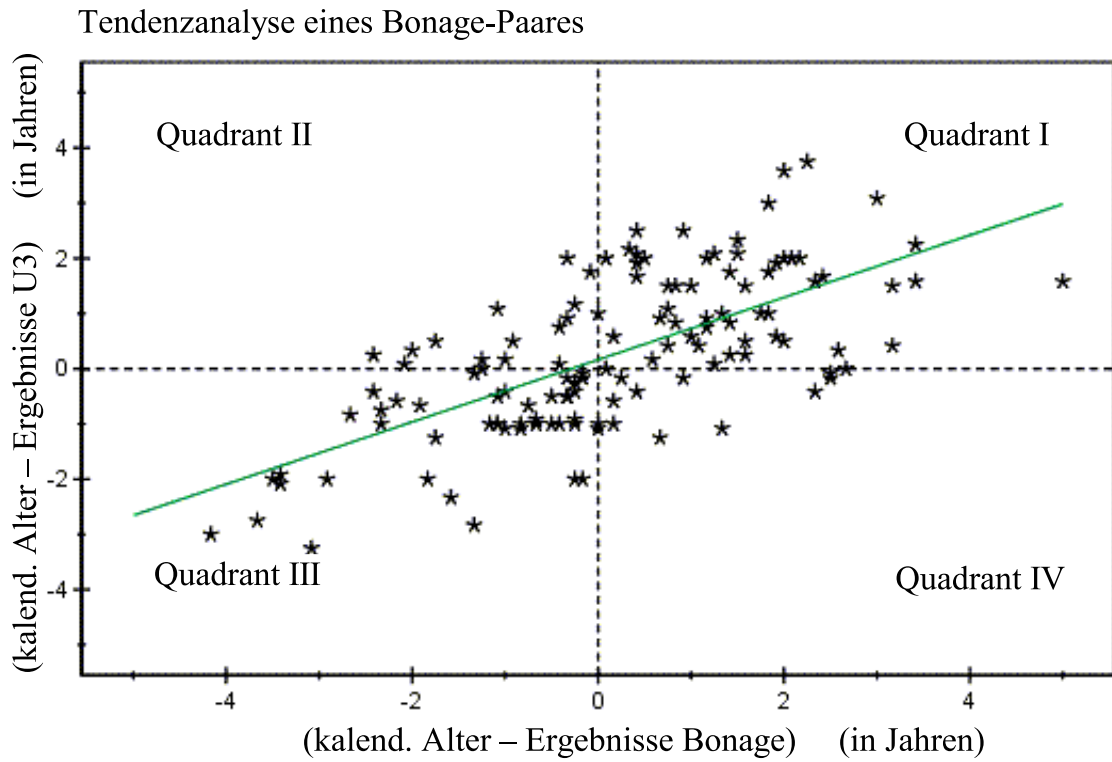


Abbildung 11: übereinstimmende und gegensätzliche Tendenzen des Paares Bonage/ U3

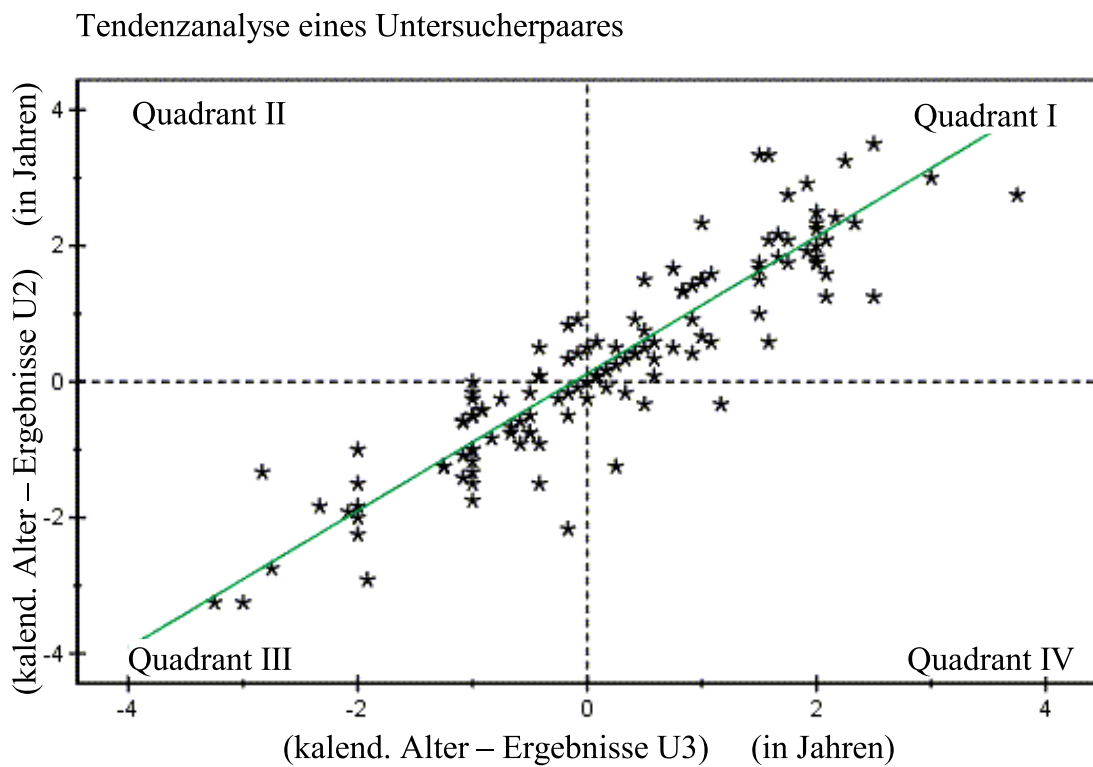


Abbildung 12: übereinstimmende und gegensätzliche Tendenzen des Paares U2/ U3

5 Diskussion

5.1 Interpretation der Ergebnisse der Methode nach Greulich & Pyle

5.1.1 Interbeurteiler-Variabilität der Greulich & Pyle-Methode

Die Methoden nach Greulich & Pyle (Auflage 1959) und nach Tanner & Whitehouse sind die beiden Verfahren der Knochenaltersbestimmung, die in der klinischen Praxis am weitesten verbreitet sind und am häufigsten angewendet werden (Berst et al. 2001, Bull et al. 1999). Deshalb sind die Interbeurteiler-Fehler beider Verfahren bereits Gegenstand mehrerer Studien gewesen, die teilweise zu sich widersprechenden Ergebnissen gelangten. So zeigten Groell et al., dass bei Anwendung der Methode von Greulich & Pyle erfahrene gegenüber unerfahreneren Untersuchern einen geringeren Interbeurteiler-Fehler im Vergleich ihrer Ergebnisse aufwiesen (siehe Kapitel 1.7.3.1)(Groell et al. 1999). Auch Johnson et al. analysierten die Interbeurteiler-Variabilität zwischen verschiedenen qualifizierten Untersuchern. Die Differenzen der Kinderradiologen lagen im Mittel bei $0,58 \pm 0,53$ Jahren (entspricht $7 \pm 6,4$ Monate). Die Differenz der weniger erfahrenen Beurteiler betrug dagegen nur $0,38 \pm 0,43$ Jahre (entspricht $4,6 \pm 5,2$ Monate) (Johnson et al. 1973). Die Variabilität der Ergebnisse der weniger Erfahrenen war somit geringer als die der Kinderradiologen. Dieses Ergebnis steht den Resultaten von Groell et al. entgegen. Eine Studie, die behauptet, dass die Interbeurteiler-Variabilität der Methode sehr hoch und folglich die Methode an sich sehr ungenau sei, ist die Studie von Cundy et al. (Cundy et al. 1988). In ihrer Studie ließen die Autoren dasselbe Set Röntgenbilder (60 Bilder) von vier verschiedenen Radiologen auswerten. Diese bestimmten das Knochenalter aller 60 Bilder nach der Greulich & Pyle-Methode. Ein Vergleich ihrer Ergebnisse führte zu folgendem Ergebniss: 50% der Differenzen waren > 1 Jahr und 10% der Differenzen sogar > 2 Jahre. In Tabelle 17 sind die ermittelten Interbeurteiler-Fehler aus einigen Studien zusammengestellt.

In der vorliegenden Arbeit unterschieden sich die Untersucher im Grad ihrer Qualifizierung. Ebenso variierte die Erfahrung der einzelnen Untersucher im Umgang mit der

Tabelle 17: Interbeurteiler-Variabilität bei Anwendung der Greulich & Pyle-Methode in verschiedenen Studien

Autor / Erscheinungsjahr	Statistische Maßzahl	Interbeurteiler-Variabilität (in Jahren)
Acheson et al. / 1966	Standardfehler	0,55 - 0,57
Gilli / 1996	mittlere Differenz	0,10 - 0,34
Groell et al. / 1999	mittlere Differenz	0,01 - 0,19
Johnson et al. / 1973	mittlere Differenz	0,39 - 0,84
Roche et al. / 1970 (März)	mittlere Differenz	0,37 - 0,60
Roche et al. / 1970 (Juni)	mittlere Differenz	0,43
die hier vorliegende Studie	mittlere Differenz	0,02 - 0,16

Methode nach Greulich & Pyle. Sowohl die Pädiaterin mit Spezialisierung auf dem Gebiet der pädiatrischen Radiologie (U1) als auch die Fachärztin für Pädiatrie mit Spezialisierung für den Fachbereich Endokrinologie (U3) hatten durch die tägliche Anwendung der Methode fundierte Kenntnisse und ausreichende Erfahrung erlangt. Der Radiologe (U2) führte die Knochenaltersbestimmung seltener durch als die beiden Pädiater (U1 und U3). Der Doktorand (U4), der vor Beginn dieser Studie diese Methode noch nicht angewendet hatte, konnte nicht auf Erfahrungen in der Thematik der Knochenaltersbestimmung zurückgreifen. Die Resultate der Interbeurteiler-Variabilität für die vorliegende Studie sind in Tabelle 18 zusammengefasst. Der geringste Interbeurteiler-Fehler wurde zwischen den Ergebnissen der Untersucher 1, 2 und 4 beobachtet (mittlere Differenz = 0,29 Monate = 0,02 Jahre). Die höchste Variabilität ergab sich zwischen den Ergebnissen der Untersucher 1 und 3 (mittlere Differenz = 1,97 Monate = 0,16 Jahre). Im Vergleich mit den Interbeurteiler-Fehlern der Studien aus Tabelle 17 zeigen die Ergebnisse der hier vorliegenden Arbeit eine ausgezeichnete Übereinstimmung der Ergebnisse aller Untersucher. Im Gegensatz dazu steht die statistische Bewertung der Paare U1/ U3, U2/ U3 und U3/ U4. Obwohl die Variabilität dieser Paare im Vergleich mit den Ergebnissen der Studien anderer Autoren (Tabelle 17) eine Interbeurteiler-Variabilität zeigt, die auf eine hohe Reproduzierbarkeit der vorliegenden Studienergebnisse schließen lässt, zeigte der t-Test für alle drei Paare signifikante Abweichungen auf dem 5%-Niveau. Der Grund für diese statistische Signifikanz der Paare ist am wahrscheinlichsten auf die Größe des Stichprobenumfangs ($n = 127$ bzw. 152) zurückzuführen. Dieses Phänomen ist eine in der Statistik bekannte Problematik, da die klassischen Signifikanztests insoweit irreführend sind, dass ein Untersuchungser-

gebnis selbst bei kleinsten Effekten signifikant wird, wenn der Stichprobenumfang nur groß genug ist (Bortz 1999). An dieser Stelle muss aber kritisch hinterfragt werden, ob die entdeckten Signifikanzen klinisch relevant sind. Denn selbst der höchste ermittelte Interbeurteiler-Fehler in der Größenordnung von 0,16 Jahren ist für ein subjektives Auswertungsverfahren ein hervorragendes Ergebnis.

Tabelle 18: Interbeurteiler-Variabilität bei Anwendung der Greulich & Pyle-Methode in der vorliegenden Studie (s = Standardabweichung)

Untersucherpaar	Statistische Maßzahl	Interbeurteiler-Variabilität ($\pm s$) (in Jahren)
U1 - U2	mittlere Differenz	0,02 ($\pm 0,69$)
U1 - U3	mittlere Differenz	0,16 ($\pm 0,68$)
U1 - U4	mittlere Differenz	0,05 ($\pm 0,58$)
U2 - U3	mittlere Differenz	0,12 ($\pm 0,63$)
U2 - U4	mittlere Differenz	0,02 ($\pm 0,74$)
U3 - U4	mittlere Differenz	0,13 ($\pm 0,67$)

5.1.2 Ursachen für die Interbeurteiler-Variabilität

Variable Faktoren, die bereits bei der Anfertigung eines Röntgenbildes in unterschiedlichster Weise auftreten können (z.B. relative Über- und Unterbelichtung, Positionierung der Hand, etc.), haben keine Auswirkung auf die Reproduzierbarkeit der Knochenaltersbestimmung (Roche et al. 1970a).

Eine mögliche Ursache für die Variabilität der Ergebnisse der Knochenaltersbestimmung nach der Greulich & Pyle-Methode in der durchgeführten Untersuchung stellt die Auswahl der Standardbilder im Atlas dar. In einigen Standard-Handradiogrammen, die im Atlas direkt aufeinander folgen, sind die Form- bzw. Gestaltänderungen einzelner Knochen so minimal, dass eine korrekte Altersschätzung erschwert wird. Einige Knochen in verschiedenen Standard-Handradiogrammen scheinen sich laut Gilli sogar in einem jüngeren Reifestadium zu befinden als in einer vorangehenden Standardabbildung niedrigeren Knochenalters. In dieser Hinsicht ist die Auswahl der Standard-Handradiogramme im Atlas teilweise unzureichend. Diese Problematik bei der Anwendung der Greulich & Pyle-Methode wirkt sich nachteilig auf die Reproduzierbarkeit der Methode aus (Gilli 1996). Gilli führt noch einen weiteren Fakt an, der die Schätzung

des Knochenalters erschwert und damit wiederum die Reproduzierbarkeit der Methode verringert. In einigen Handradiogrammen divergiert das Alter einiger abgebildeter Knochen sehr stark. Die Unterschiede des Knochenalters einzelner Knochen im selben Radiogramm beträgt in den männlichen Standards bis zu 20 Monaten und bis zu 10 Monaten in den weiblichen Standards (Gilli 1996).

Einen weiteren Einfluss auf die Reproduzierbarkeit sieht Gilli in der Methodik selbst. Die Beurteilung des Knochenalters basiert auf dem Prinzip, dass die Gestalt der Knochen und deren Gestaltsänderungen in beschreibenden Texten verbalisiert wurden, um Standards zu veranschaulichen. Im Verlauf der Untersuchung müssen die geschriebenen Kriterien wiederum visualisiert werden. Schwierigkeiten in der präzisen Beschreibung von Reifeindikatoren und deren korrekte Interpretation könnten eine Quelle für die Interbeurteiler-Variabilität sein (Gilli 1996).

Die Analyse der Differenzen demonstrierte eine weitere Schwierigkeit, die bei der Anwendung der Atlasmethode auftreten kann: Alle Differenzen, die außerhalb der limits of agreement lagen, wurden näher betrachtet. Ca. 80% dieser Differenzen begründen sich auf Handradiogramme von Patienten, die eine dissoziierte Entwicklung der Knochen ihres Handskeletts aufwiesen. Das heisst, die größten Abweichungen in den Beurteilungen der Untersucher traten in dieser Studie bei den Patienten auf, die eine dissoziierte Knochenentwicklung zeigten (n = 27 von 152). Bei diesen Patienten ist es besonders kompliziert, manchmal sogar unmöglich, ein pauschales Knochenalter mit der Greulich & Pyle-Methode zu bestimmen. Auch die von Greulich & Pyle angeratene Interpolation des Knochenalters kann in Fällen großer Differenzen nicht als befriedigende Beurteilung, sondern vielmehr als erforderlicher Kompromiss angesehen werden. Die Methode von Greulich & Pyle zählt zu den subjektiven Verfahren und dem Untersucher wird keine konkrete Vorgehensweise auferlegt. Diese Tatsache kommt besonders in solchen Problemfällen zum Tragen. Die Wichtung von speziellen Reifefaktoren und somit die Zuordnung des am ehesten zutreffenden Knochenalters ist allein von der Einschätzung und der Erfahrung des Untersuchers abhängig. Basierend auf dieser Vorgehensweise werden die Grenzen der Methode deutlich. Umso mehr die Bewertung des Knochenalters auf der subjektiven Einschätzung des Untersuchers beruht, desto schlechter wird die Reproduzierbarkeit. Dieses Problem wurde bereits von anderen Autoren geschildert (Roche et al. 1970a).

Die Alternative zur Greulich & Pyle-Methode ist die Knochenaltersbestimmung nach der von Tanner & Whitehouse vorgestellten Methode (Bull et al. 1999). Da dieses Verfahren allerdings auch zu den morphischen Bestimmungsmethoden zählt, existieren zwischen beiden Verfahren Parallelen hinsichtlich ihrer Anfälligkeit gegenüber zufälligen und systematischen Fehlern. Die Studie von Acheson et al. untersuchte bei beiden Verfahren, in welcher Art und Weise solche Fehler die Ergebnisse der Knochenaltersbestimmung beeinflussen (Acheson et al. 1966). Die Autoren fanden heraus, dass zufällige Fehler bei der Greulich & Pyle-Methode deutlich stärkere Verfälschungen der Ergebnisse bewirken als bei der Tanner-Whitehouse-Methode. Als Maß für die Auswirkung zufälliger Fehler wurden die Intrabeurteiler-Variabilitäten mehrerer Untersucher miteinander verglichen. Die 95%-Konfidenzintervalle waren bei der Methode von Tanner & Whitehouse signifikant kleiner. Folglich wies diese Methode eine geringere Intrabeurteiler-Variabilität auf als die Greulich & Pyle-Methode.

Neben den zufälligen Fehlern existieren zudem systematische Fehler. Cundy et al. konnten zwar in ihrer Studie das Vorhandensein systematischer Fehler bei der Anwendung der Methode von Greulich & Pyle belegen (zwei der vier Radiologen unterschätzten konsequent das Knochenalter), jedoch gaben sie keine genaueren Ursachen für diese Art der Fehlerquellen an (Cundy et al. 1988). Einer der systematischen Fehler äußert sich darin, dass die Tanner-Whitehouse-Methode im direkten Vergleich mit der Greulich & Pyle-Methode konsequent Knochenalter bestimmt, die ca. 1 Jahr höher sind (Acheson et al. 1966). Acheson et al. sehen diese Diskrepanz in den unterschiedlichen Populationen begründet, aus welchen die Referenzstandards zusammengestellt wurden (Standards der Greulich & Pyle-Methode aus der Population amerikanischer Jugendlicher vs. Standards der Tanner-Whitehouse-Methode aus der Population britischer Jugendlicher). Weiterhin zeigte sich, dass die Methode von Tanner & Whitehouse anfälliger für systematische Fehler ist, als die Methode von Greulich und Pyle (Acheson et al. 1966). Eine mögliche Ursache könnte in der Bildung des skeletal maturity scores liegen, der eigentlich einen Mittelwert mehrerer Scores verschiedener Knochen darstellt. Neigt ein Untersucher dazu, das Knochenalter systematisch zu unter- bzw. überschätzen, so summiert sich dieser Fehler bei der Bildung eines Gesamtscores auf. Auch für die Greulich & Pyle-Methode ließen sich in unterschiedlichen Studien mögliche Ursa-

chen systematischer Fehlerquellen identifizieren (siehe Kapitel 1.7.3.1, Einflussfaktor der Carpalia) (Acheson et al. 1966, Johnston und Jahina 1965).

Einigen Studien zufolge erbringt die Tanner-Whitehouse-Methode deutlich genauere Ergebnisse der Knochenaltersbestimmungen als die von Greulich & Pyle entwickelte Methode; so z.B. die Studie von Bull (Bull et al. 1999). Der hohe Zeitaufwand spricht, trotz der besseren Ergebnisse, jedoch oft gegen den Einsatz der erstgenannten Methode im klinischen Alltag. Eine deutliche Verbesserung dieses Problems bietet eine neue Computersoftware an, welche verspricht, den Zeitaufwand erheblich zu minimieren und die Methode somit effizienter zu gestalten. Ein Beispiel eines solchen Programmes ist Bone-O-Matic (<http://bone-o-matic.com>). Diese neue Software ermöglicht laut Minas ein komfortableres Arbeiten (Minas und Minas 2004).

5.1.3 Ansätze zur Verringerung der Variabilität und Verbesserung der Interbeurteiler-Variabilität

Einen möglichen Ansatz für die Verringerung der Variabilität und eine Erhöhung der Reproduzierbarkeit geben Roche et al. in ihrer Veröffentlichung an (Roche et al. 1970a). Sie schlagen vor, die Knochenaltersbestimmung von 2 Untersuchern durchführen zu lassen und deren Ergebnisse zu interpolieren. Die Autoren beschrieben in ihrer Studie, dass sich die mittlere Interbeurteiler-Variabilität einzelner Untersucher von 0,43 Jahren um 0,20 Jahre reduzierte. Der mittlere Interbeurteiler-Fehler der Untersucherpaare betrug dann nur noch 0,23 Jahre (Roche et al. 1970a). Allerdings muss die Frage gestellt werden, ob diese Lösung in der täglichen Routine des Klinikalltages praktikabel ist. Nicht alle radiologischen Abteilungen, v.a. die kleinerer Krankenhäuser, sind mit 2 Radiologen besetzt, die überdies Erfahrung im Umgang mit der Methode von Greulich & Pyle aufweisen müssten.

Maßnahmen im Sinn einer Qualitätssicherung führen ebenfalls zu einer Verringerung des Interbeurteiler-Fehlers. Roche et al. untersuchten, ob ein Trainingseffekt Einfluss auf die Interbeurteiler-Variabilität besitzen würde. Zwar konnten die Ergebnisse durch häufiges Anwenden der Methode nicht verbessert werden, gleichwohl zeigte sich aber, dass im Rahmen einer Zwischenauswertung die Diskussion über aufgetretene Probleme

und Schwierigkeiten im Umgang mit der Methode nachfolgend zu einem geringeren Interbeurteiler-Fehler führte. Vor der Diskussion betrug die Interbeurteiler-Variabilität 0,32–0,56 Jahre, nach der Diskussion hingegen 0,25–0,53 Jahre (Roche et al. 1970b). Dieser Ansatz wäre in der alltäglichen Praxis nicht nur gut vorstellbar, sondern auch gut zu integrieren. Des Weiteren würde diese Möglichkeit auch die Standardisierbarkeit der doch in Teilen recht subjektiven Methode unterstützen.

5.2 Interpretation der Ergebnisse der Ultraschallmethode

5.2.1 Präzision des Bonage-Ultraschallsystems

Der für das System ermittelte Variationskoeffizient von 3,32% bedeutet, dass der Messfehler von dem Ergebnis abhängt, das mit Bonage ermittelt wurde (siehe Kapitel 4.1). Wird z.B. ein Knochenalter von 8 Jahren (96 Monate) bestimmt, so beträgt der Messfehler $\pm 3,2$ Monate. Mit steigendem Knochenalter eines Probanden bzw. Patienten wird dieser Messfehler zunehmen. Bei einem Knochenalter von 17 Jahren (204 Monate) muss ein Messfehler von $\pm 6,8$ Monaten bei der Bewertung des Knochenalters einkalkuliert werden.

Der errechnete Variationskoeffizient begründet sich auf wiederholten Messungen an fünf Probanden. Bei 4 der 5 Testpersonen waren keine Erkrankungen bekannt, die einen Einfluss auf das Knochenalter ausüben könnten. Lediglich bei einer Testperson (Proband 5) war ein konstitutioneller Entwicklungsvorsprung diagnostiziert worden. Bis zu diesem Zeitpunkt befand sich der Proband jedoch nicht in einer medikamentösen Therapie, sondern in einem Zyklus regelmäßiger, fachärztlicher Kontrollen zur Beurteilung der weiteren Entwicklung.

Auffällig ist der Variationskoeffizient von 8,9% bei Proband 3. Zum Zeitpunkt der Messungen des Probanden 3, war dieser 5 Jahre und 7 Monate alt und damit der jüngste der fünf Probanden. In dem jungen Alter des Probanden, das relativ nah an der zugelassenen Altersgrenze für das Bonage-System liegt, ist eventuell auch der hohe Variationskoeffizient begründet (siehe Kapitel 5.2.2). Betrachtet man den Variationskoeffizienten von Proband 3 als Ausreißer und schliesst ihn aus der statistischen

Auswertung aus, so reduziert sich der Variationskoeffizient um 1,4 Prozentpunkte auf 1,92%.

Auf der Grundlage eines angenommenen Variationskoeffizienten von 3,32% beträgt der größte zu erwartende Messfehler, der bei der Knochenaltersbestimmung eines 18-jährigen Patienten auftreten könnte, demzufolge 7,2 Monate. Diese Größenordnung entspricht einer Interbeurteiler-Variabilität von 0,6 Jahren und liegt in einem Wertebereich, der auch in anderen Studien über die Greulich & Pyle-Methode beschrieben wurde (siehe Tabelle 17). Legt man den Variationskoeffizienten von 1,92% zu Grunde, der sich durch den Ausschluss von Proband 3 (Ausreißer) errechnet, so beträgt der größte zu erwartende Messfehler 4,15 Monate.

Nach Angabe des Herstellers (Sunlight Medical Ltd., Tel Aviv, Israel) beträgt der Messfehler für weibliche Patienten $0,25 \pm 0,19$ Jahre (entspricht $3 \pm 2,3$ Monate), für männliche Patienten $0,23 \pm 0,15$ Jahre (entspricht $2,8 \pm 1,8$ Monate). Im Mittel ist laut Hersteller also eine Abweichung des Ergebnisses um $0,24 \pm 0,16$ Jahren (entspricht $2,9 \pm 1,9$ Monate) einzukalkulieren.

5.2.2 Bonage-Messungen

Dass die Präzision des Bonage-Systems zunehmend ungenauer wird, wenn sich das Alter des zu messenden Patienten dem minimalen und maximalen Grenzwert der Altersbeschränkung von 5 bzw. 18 Jahren nähert, lässt sich neben dem Variationskoeffizienten auch anhand der einzelnen Messergebnisse erkennen.

Die Bestimmung des Variationskoeffizienten zeigt einen Trend: Mit Ausnahme von Proband 2 steigt der Koeffizient mit zunehmendem Alter der Probanden. Aus dieser Beobachtung und dem großen Wert von Proband 3 ist anzunehmen, dass die Präzision bei Annäherung an ein Alter nahe der Altersbeschränkung abnimmt. Natürlich würde man einen größeren Stichprobenumfang benötigen, um diese These zu bestätigen. Der vermutete Trend lässt sich jedoch durch eine Betrachtung der Messwerte einzelner Patienten erhärten. Das geringste Knochenalter, welches jeweils von den Untersuchern bestimmt wurde, lag in einem Bereich von 36 Monate (Untersucher 2) bis 42 Monate

(Untersucher 4). Der kleinste Wert, der von Bonage für dieselbe Studienpopulation ermittelt wurde, beträgt 58 Monate. In Analogie dazu lässt sich eine ähnliche Beobachtung für die obere Altersgrenze konstatieren. Alle Untersucher mit Ausnahme von U3 bestimmten den höchsten Wert für ein Knochenalter aus der Studienpopulation auf ≥ 222 Monate (U3: 207 Monate). Der höchste Wert, den Bonage dagegen ermittelte, betrug 215 Monate. Folgende Schlussfolgerungen werden daher wahrscheinlich: Patienten, die das 5. Lebensjahr vollendet haben, aber in ihrer biologischen Reife noch nicht 5 Jahre erreicht haben (retardiertes Knochenalter), werden von Bonage nicht korrekt diagnostiziert. Auch der Umkehrschluss trifft zu: Patienten, die das 18. Lebensjahr noch nicht erreicht haben, aber in der biologischen Reife über 18 Jahren liegen, werden ebenfalls vom Bonage-System nicht korrekt erfasst.

Diese Ergebnisse müssen bei der Frage der Relevanz in der Klinik differenziert gesehen werden. Die Tatsache, dass die Messgenauigkeit von Bonage an der oberen Grenze von 18 Jahren abnimmt, hat klinisch nur geringe Bedeutung. Zum einen besitzen Patienten weiblichen Geschlechts bereits mit 16 Jahren ein insoweit ausgereiftes Skelett, dass das Wachstum in diesem Alter fast beendet ist. So ergibt sich daraus keine klinische Bedeutung, da eine wachstumsmodulierende Intervention keine Wirkung mehr zeigen und somit nicht mehr sinnvoll erscheinen würde. Patienten männlichen Geschlechts dagegen erreichen den Abschluss des Längenwachstums zwar erst in einem Alter von 17 bis 18 Jahren, aber der Erfolg einer wachstumsmodulierenden Therapie würde zu diesem Zeitpunkt ebenso eher fraglich erscheinen. Wird hingegen ein retardiertes Knochenalter eines 5-jährigen Kindes nicht erkannt, so resultiert daraus eine klinische Relevanz. Die Plastizität des Skelettsystems und die Wachstumsreserve eines jüngeren Kindes sind größer als bei älteren Kindern und nehmen im Verlauf der Skelettreifung immer mehr ab (Sitzmann 2002). Am häufigsten werden die Untersuchungen zur Knochenaltersbestimmung jedoch bei Kindern im Schulalter durchgeführt. Dass die Software und die Referenzdatenbanken des Bonage-Systems die Knochenaltersbestimmung nur von Kindern zulässt, die >5 Jahre sind, ist eindeutig als Nachteil der sonographischen Bestimmungsmethode zu werten. Theoretisch ist die sonographische Knochenaltersbestimmung auch bei Kindern möglich, die jünger als 5 Jahre sind. Dafür müssten allerdings noch weitere Studien erarbeitet bzw. durchgeführt werden, welche zu einer

Erweiterung der Datenlage genutzt werden könnten, um eine suffiziente Knochenaltersbestimmung mit Bonage auch in jüngeren Altersgruppen zu ermöglichen.

5.3 Beurteilung des Vergleichs der Atlasmethode und der Ultraschallmethode

Der statistische Vergleich der Methoden zeigte, dass keine signifikanten Unterschiede der Methoden auf dem 5%-Niveau bestehen. In diesem Fall bedeutet das, dass der t-Test für abhängige Stichproben keinen signifikanten Unterschied der Stichprobenmittelwerte der beiden Methoden feststellen konnte. In Bezug auf die mittlere Lage der Differenzen existiert also kein Unterschied zwischen der Methode nach Greulich & Pyle und der Ultraschallmethodik, die Bonage nutzt. Sehr auffällig ist jedoch die unterschiedliche Präzision beider Verfahren. Die Testwerte für den Maloney-Rastogi-Test belegen für die Untersucherpaare Bonage/ U1 und Bonage/ U2 eindeutig Unterschiede in der Präzision (p -Wert $\leq 0,05$). Nur für das Paar Bonage/ U3 ließ sich kein Präzisionsunterschied zwischen beiden Methoden finden. Dies beruht auf den Ergebnissen von Untersucher 3, die besser mit dem Bonage-System übereinstimmen als mit den Ergebnissen der anderen Untersucher. So weist der t-Test im Vergleich von U3 mit allen anderen Untersuchern signifikante Unterschiede auf. Im Vergleich von U3 mit Bonage existiert jedoch eine gute Übereinstimmung der Ergebnisse, die durch den Testwert p ($p = 0,153$) für das 5%-Niveau bestätigt wird.

Die großen Präzisionsunterschiede der Methoden, die sich v.a. im Vergleich der Untersucher 1 und 2 mit Bonage manifestieren, sind bereits den Bland-Altman-Plots dieser Paare zu entnehmen. Der Streubereich, in dem sich die Differenzen befinden, ist bei den Bonage-Paaren erheblich größer als bei den Untersucherpaaren. Verdeutlicht wird dies auch durch eine weitere statistische Kenngröße, die Standardabweichung. Diese sind im Vergleich bei den Bonage-Paaren ($s = 15,5$ bis $17,2$ Monate) doppelt so groß wie die Standardabweichung der Untersucherpaarungen ($s = 7,0$ bis $8,9$ Monate).

5.3.1 Erkrankungen mit Einfluss auf die Ultraschallmethode

Die große Streuung der Differenzen um den Mittelwert wirft die Frage nach möglichen Ursachen auf. Bei 60 Patienten ergaben sich Differenzen (zwischen den Ergebnissen von Bonage und den Ergebnissen der Untersucher 1, 2 und 3), die mehr als 18 Monaten betragen. In diesen Fällen wurde nach Gründen gesucht, welche die großen Abweichungen (Differenzen >18 Monate) erklären könnten. Für 51,7% (n = 31) ließen sich keine Ursachen eruieren. Für die anderen 48,3% (n = 29) konnten die klinischen Diagnosen der Patienten die Diskrepanzen teilweise erklären. Da das Bonage-System die Altersbestimmung anhand des Ultraschallparameters speed of sound ermittelt, nehmen alle Komorbiditäten, die in irgendeiner Art eine Veränderung dieses Parameters bewirken, direkt Einfluss auf die Validität des Ergebnisses. Da der Parameter SOS stark von physikalischen Eigenschaften und der Struktur des Knochens abhängig ist (Fredfeldt 1986, Sievanen et al. 2001), wird der SOS-Wert besonders durch Erkrankungen verzerrt, die sich am Organsystem Skelett/ Knochen manifestieren.

9 der 29 Patienten befanden sich zum Zeitpunkt der Ultraschallmessung in einer wachstumsmodulierenden Hormontherapie bzw. hatten die Hormontherapie innerhalb des Zeitraumes von 6 zurückliegenden Monaten abgeschlossen. Die Therapie bestand entweder in der Verabreichung des Hormonpräparates Enantone, das als LH-RH-Agonist die Bildung von Sexualhormonen supprimiert, oder in der Verabreichung von Testosteron. Da die Sexualhormone als wichtiger Modulator der Skelettreifung und des skelettalen Wachstums agieren (Schmidt et al. 2000), ist als Konsequenz mit einer Veränderung des SOS-Parameters zu rechnen.

Bei 7 der 29 Patienten war eine Osteopenie bzw. Osteoporose bekannt. Da die Osteoporose mit einer verminderten Mineraldichte des Knochens vergesellschaftet ist und zu einer Veränderung der Knochenstruktur - sowohl der trabekulären als auch der kortikalen Struktur - führt (Jergas und Schmid 2005), wird auch der SOS-Wert beeinflusst (Craven et al. 1973), der in besonderem Maße von der Beschaffenheit der Kortikalis abhängt (Sievanen et al. 2001). Bei 4 der 7 Patienten mit einer Osteoporose war diese wahrscheinlich steroidbedingt (regelmäßige Einnahme von Prednisolon oder anderen Steroiden im Rahmen der Therapie chronisch-entzündlicher Darmerkrankungen), die anderen 3 waren auf renale Ursachen (z.B. Nierenhypoplasie, Störung des Vitamin-

D-Stoffwechsels und der Calciumhomöostase) zurückzuführen. Bei 5 der 7 Patienten war die verminderte Mineralisation der Knochen in den Handradiogrammen bereits röntgenologisch nachweisbar. Die Demineralisation führt zu einer Verminderung der Schalleitungsgeschwindigkeit (Mohr et al. 2004) und demnach zu einem niedrigeren sonographisch bestimmten Knochenalter.

Bei 5 Patienten war eine Hypothyreose diagnostiziert worden, die bei allen 5 Patienten eine Substitutionsbehandlung mit L-Thyroxin notwendig machte. Auch Erkrankungen der Schilddrüse können sich auf die Knochendichte auswirken. So wurde z.B. bei sehr stark TSH-supprimierender Hormonsubstitution ein signifikanter Knochendichteverlust nachgewiesen (Schindler 2003).

Weitere 5 Personen zeigten eine Adipositas. Wie sich die Adipositas genau am knöchernen System manifestiert, ist zur Zeit noch umstritten. Sicher ist aber, dass das größere Körpergewicht zu einer mechanischen Mehrbelastung der Knochen führt und dass die Adipositas den Stoffwechsel deutlich verändert. Zusätzlich resultiert als Folge von Adipositas oft ein relativer Bewegungsmangel, der sich über Veränderungen des Knochens auch auf die SOS-Werte auswirkt und somit die sonographische Knochenaltersbestimmung beeinflusst. Eine weitere Beeinflussung der Knochenaltersbestimmung mit dem Bonage-System lässt sich auf den Habitus der Adipositaspatienten zurückführen. Durch die oftmals dickeren Unterarme dieser Patienten wird die Vorlaufstrecke des Ultraschalls vergrößert. Dies bedingt wiederum veränderte SOS-Werte und in der Folge eine Verfälschung der Knochenaltersbestimmung.

Weitere 2 Patienten wurden mit einem Enzymdefekt geboren, der zu der Entwicklung eines adrenogenitalen Syndroms bei beiden Patienten führte. Durch diese Enzymdefekte entsteht eine reaktive Mehrsekretion von ACTH und in deren Folge die Überstimulierung der Sekretion von Androgenen (Sitzmann 2002). Die Androgene modulieren dann ihrerseits das Wachstum und die skelettale Entwicklung. Durch die beschleunigte Reifung des Skeletts haben diese Patienten eine höhere Knochendichte durch eine verstärkte Mineralisation. Die Messungen der Geschwindigkeit des Ultraschalls durch den Knochen (SOS) ergeben dann höhere Transmissionsgeschwindigkeiten und führen folglich zu einem zu hoch bestimmten Knochenalter.

Ein Patient hatte sich ca. ein Jahr vor der Messung einer Chemotherapie unterziehen müssen, nachdem bei ihm ein Rhabdomyosarkom diagnostiziert worden war. Unter der Chemotherapie kam es kurzzeitig zu einer Verminderung des Wachstums.

Tabelle 19 fasst noch einmal Art und Anzahl der Erkrankungen zusammen, bei denen ein Einfluss auf das Ergebnis der quantitativen Ultraschallmessung wahrscheinlich ist.

Tabelle 19: Einfluss von Erkrankungen und Therapien auf die Bonage-Messungen. Bei 29 Patienten ließen sich Einflussfaktoren eruieren, die sich auf die Ergebnisse der Ultraschallmethode auswirkten.

Einfluss bzw. Erkrankung	Anzahl	
Hormontherapie	9	
Osteopenie bzw. -porose	7	(davon 4 steroidbedingt und 3 renal bedingt)
Hypothyreose	5	
Adipositas	5	
AGS	2	
Chemotherapie	1	
Σ	29	

5.3.2 Einfluss verschiedener Fehlerquellen auf die Messungen mit dem Bonage-System

Von den 8 Patienten, bei denen mit dem Bonage-System kein Knochenalter bestimmt werden konnte, waren 2 Patienten zu unruhig, sodass die Messungen selbständig vom Bonage-System abgebrochen wurden. Bei 5 Patienten bestand eine erhebliche somatische Entwicklungsverzögerung. Der biologische Reifegrad entsprach bei diesen Patienten nicht annähernd dem kalendarischen Alter. Die Unterarme dieser 5 Patienten waren zu dünn, um erfolgreiche Messungen durchführen zu können. Selbst mehrmals wiederholte Messungen führten nicht zu einem Ergebnis. Eine Patientin war an Cystischer Fibrose erkrankt und entwickelte im Verlauf der Krankheit Deformierungen des Handskeletts. Die Bestimmung des Knochenalters mit der Ultraschallmethode war bei ihr nicht möglich. Leiden Patienten an pathologisch-progressiven Prozessen des Handskeletts, die Deformierungen insbesondere im Bereich der Handwurzel verursachen, wird die Ultraschallmessung mit Bonage ebenfalls erschwert.

Natürlich unterliegt die Ultraschallmessung systematischen und zufälligen Fehlern. Um aber systematische Fehler identifizieren zu können, werden weitere experimentelle Studien nötig sein.

Ein denkbarer systematischer Fehler der Ultraschallmethode könnte in der Ausbreitung der Schallwellen begründet sein. Die Methode basiert auf einem Messprinzip, das auf der Quertransmission des Ultraschalls durch den Knochen beruht. Die komplexe Ausbreitung von Ultraschall in einem inhomogenen Gewebe wie z.B. Knochen ist bisher noch nicht vollständig nachvollziehbar (Barkmann et al. 1999). Trifft der Ultraschall nicht exakt auf die Epiphysenfuge bzw. deren angrenzende Zonen, die er möglichst direkt durchqueren sollte, verändert sich neben der Ausbreitungsrichtung des Ultraschalls auch das zeitliche Intervall zwischen Aussendung und Empfang des Ultraschallsignals. Diese Veränderungen können zu verfälschten Messparametern (SOS) und folglich zu einem falschen Messergebnis führen.

Während die Messung erfolgt, sollte sich der Patient nach Möglichkeit nicht bewegen, damit die Schallsonden die Messung nicht außerhalb des festgelegten Bereiches durchführen bzw. die Ausbreitung des Ultraschalls in der Messzone nicht gestört wird. Hier wird eine Limitation der Ultraschallmethode offensichtlich. Nicht allen Kindern fällt es leicht, während der Messungen möglichst bewegungslos zu verharren. Dies gilt besonders für jüngere Kinder. Sind sie zu unruhig oder zu aufgeregt, sind Messungen mit dem Bonage-System nicht möglich bzw. werden sehr ungenau. In dieser Hinsicht sind die Verfahren, welche auf der Beurteilung von Handradiogrammen beruhen, dem Bonage-System überlegen. Zwar gibt es auch Kinder, die aus den verschiedensten Gründen (sei es z.B. die Trennungsangst von den Eltern) auch bei der Röntgenaufnahme der Hand nicht ruhig genug sitzen bleiben, aber die Anfertigung eines Handradiogramms dauert nur einen wenige Augenblicke und ist wesentlich kürzer als die Messung mit der Ultraschallmethode.

Vor der Messung muss der Untersucher den Processus styloideus ulnae mittels einer Schablone anzeichnen. In diesem Schritt des Untersuchungsablaufes ist eine weitere Quelle für zufällige Fehler erkennbar. Wird der Processus nicht genau markiert, können die Schallwandler nicht exakt auf den vorgesehenen Messbereich eingestellt werden. In der Studie trat dieses Problem besonders bei adipösen Kindern auf. Die Markierung

mittels der zugehörigen Schablone setzt voraus, dass der Processus ausreichend prominent ist, um die Schablone exakt am Handgelenk anlegen zu können. Nur dann kann der Untersucher die Markierung an der richtigen Stelle anzeichnen. Das ist ein Grund dafür, dass Messungen bei Kleinkindern (<5 Jahre) nicht möglich sind. Bei Kindern in diesem Alter ist die distale Ulnaepiphyse noch nicht knöchern angelegt. Bei den adipösen Kindern ließ sich der Processus zwar palpieren, aber mit der Schablone nur schlecht markieren.

Bei adipösen Kindern tauchte zusätzlich ein weiteres Problem auf, da der Ultraschallparameter SOS direkt von dem Gewebe abhängt, das der Schall durchdringen muss. Da sich der Ultraschall im Weichteilgewebe viel langsamer ausbreitet als im Knochen, werden die Messungen des SOS immer ungenauer, je mehr Weichteilgewebe vom Schall durchdrungen werden muss (Barkmann et al. 2002, Baroncelli et al. 2001). So ist die Vorlaufstrecke des Ultraschalls durch das vermehrte Weichteilgewebe bei Adipositas relativ vergrößert und führt zu einer Verringerung des SOS-Wertes. Dadurch wird ein zu langes Zeitintervall vorgetäuscht und ein falsch-niedriges Knochenalter bestimmt.

Zur Studienpopulation gehörten auch Patienten, bei denen der Processus styloideus ulnae nur minimal ausgebildet war. Auch bei diesen Patienten konnte der Processus mit der Schablone nicht optimal markiert werden.

Ein zufälliger Fehler könnte in der Ankopplung der Ultraschallsonden während des Messvorgangs liegen. Ist die Ankopplung der Sonden nicht optimal, kann das Messergebnis verfälscht werden. Besonders Einschlüsse von Luftblasen im Gel, welches zur Schallankopplung genutzt wird, können erhebliche Veränderungen des Messergebnisses bei quantitativem Ultraschall verursachen (Barkmann et al. 1999). Die Ankopplung wird zudem gestört, wenn nicht ausreichend Gel auf die Sonden bzw. am Handgelenk des Patienten aufgetragen wird. Wie in Kapitel 3.2.4.3 beschrieben, verändern die Ultraschallsonden während des Messvorganges ihre Position. Nach jedem Messzyklus werden die Sonden aus der Ankopplung gelöst und positionieren sich neu, um dann erneut angekoppelt zu werden. Dieser Vorgang wird vom Bonage-System selbst gesteuert und ist nicht durch den Untersucher beeinflussbar. Es ist aber durchaus vorstellbar, dass nicht bei jedem Messzyklus die Schallankopplung optimal verläuft.

Bei der Datenauswertung der vorliegenden Studie fielen im Vergleich der Bonage-Methode mit der Greulich & Pyle-Methode bei einigen Patienten sehr diskrepante Werte auf. Die Daten dieser Patienten wurden deshalb ein weiteres Mal überprüft. Als Ursache für die Diskrepanz zwischen den beiden Methoden war die dissoziierte Knochenentwicklung anzusehen. Die Daten wurden dann unter dieser speziellen Fragestellung noch einmal analysiert. Es zeigte sich, dass bei 27 der 152 Patienten (entspricht 17,8% der Studienpopulation) eine dissoziierte Entwicklung des Handskeletts röntgenologisch nachzuweisen war.

Divergiert das Alter von Handwurzel und Mittelhand bzw. von Handwurzel und Phalangen, so lässt sich das nicht mit der neuentwickelten Ultraschallmethode feststellen. Das von Bonage ermittelte Knochenalter entsteht aus der Summe von Einzelmesswerten, die sich aus einer lokal begrenzten Messzone ableiten. Diese Messzone ist streng auf die distalen Epiphysenzonen von Radius und Ulna beschränkt. Zwar beschreiben schon die Autoren Wagner et al. gerade die distale Radiusepiphyse als sehr genauen Indikator des Knochenalters, der deshalb besonders für die Knochenaltersbestimmung als geeignet erscheint (Wagner et al. 1995). Trotzdem muss die Tatsache, dass nur 2 Ossifikationszentren die Basis der Knochenaltersbestimmung darstellen, kritisch betrachtet werden. Im Gegensatz dazu steht die Methode von Greulich & Pyle, die auf der Beurteilung sämtlicher Ossifikationszentren des Handskeletts beruht. Zudem visualisiert Bonage die zu analysierenden Ossifikationszentren nicht. Ein B-Bild wird nicht erstellt. Die Altersbestimmung kann daher nur auf Basis der gemessenen Ultraschallparameter und nicht auf der Basis von Form- und Gestaltwandel der Ossifikationszentren erfolgen. Unterschiede in der Variabilität beider Verfahren lassen sich also u.a. auf die verschiedenen Prinzipien der Methoden zurückführen.

5.3.3 Anwendung anderer ultraschallbasierter Methoden zur Knochenaltersbestimmung

Dass sonographische Verfahren prinzipiell zur Knochenaltersbestimmung geeignet sind, zeigen bereits die von Castriota-Scanderbeg et al. durchgeführten Studien. Das Konzept der Methode ist die sonographische Darstellung und Ausmessung der Knorpeldicke der proximalen Femurkopfeiphyse. Eine Verringerung der Knorpeldicke um 0,1mm kor-

respondiert mit einer Veränderung des Knochenalters um etwa 3 Monate. In der 1996 veröffentlichten Studie gelang es, die Kausalität zwischen Knorpeldicke der proximalen Femurepiphyse und des Knochenalters nachzuweisen (Castriota-Scanderbeg et al. 1996). Aufgrund dieser Erkenntnis erhofften sich Castriota-Scanderbeg et al., dass Reproduzierbarkeit und Präzision des Ultraschallverfahrens eine neue und suffiziente Methode der Knochenaltersbestimmung ermöglichen würden. Die Folgestudie, die 1998 publiziert wurde, sollte die Tauglichkeit der neuen Methode untersuchen. Zu große Variation in den Ergebnissen der sonographisch bestimmten Knochenalter und eine zu geringe Sensitivität konnten die Anwendung dieser neuen Methode in der klinischen Praxis zum damaligen Zeitpunkt nicht rechtfertigen (Castriota-Scanderbeg et al. 1998). Die Ergebnisse zeigten andererseits aber auch, dass sich die sonographische Methode durch eine deutlich höhere Spezifität gegenüber den konventionellen röntgenologischen Verfahren auszeichnete (Castriota-Scanderbeg et al. 1998).

Die vorliegende Studie beweist, dass sich Methodik und Technik der Ultraschallverfahren seitdem verbessert haben. So konnte für den Vergleich der auf sonographiebasierenden Bonage-Methode mit der röntgenologischen Methode von Greulich & Pyle auf dem 5%-Niveau kein signifikanter Unterschied der beiden Methoden festgestellt werden. Eine Parallele zu der Studie von Castriota-Scanderbeg et al. ist darin zu sehen, dass auch in der vorliegenden Studie die Variationen der Bonage-Ergebnisse größer als die der Greulich & Pyle-Methode sind, wenngleich auch nicht mehr signifikant.

Neben den Studien von Castriota-Scanderbeg et al. war die Anwendung sonographischer Methoden zur Knochenaltersbestimmung Gegenstand weiterer wissenschaftlicher Untersuchungen. Auf der Grundlage ihrer Ergebnisse schlussfolgerten Nesi et al., dass das von ihnen entwickelte Ultraschallverfahren eine einfache, wertvolle und strahlungsfreie Alternative zu den röntgenologischen Techniken darstellen würde (Nesi et al. 1997). In ihren Untersuchungen analysierten sie drei unterschiedliche Ossifikationszentren der Hand mittels Ultraschall und wiesen diesen einen Score zu: Os pisiforme, Ossa sesamoidea und der Knorpel der distalen Phalanx des dritten Fingers. Der ermittelte Score und das röntgenologisch bestimmte Knochenalter wurden anschließend mithilfe statistischer Tests miteinander verglichen. Bei der Beurteilung des Os pisiforme und der Ossa sesamoidea ließen sich zwischen der sonographischen und der röntgenologischen Methode keine statistisch signifikanten Unterschiede nachweisen. Bei der Beurteilung

des Knorpels der Phalanx des dritten Fingers unterschieden sich beide Techniken. Diese Unterschiede führten die Autoren auf die schlechtere Darstellbarkeit dieses Knorpels im B-Bild während der Ultraschalluntersuchung zurück (Nessi et al. 1997).

Eine andere Möglichkeit, ultraschallbasierte Techniken für die Knochenaltersbestimmung einzusetzen, wurde von Sherif et al. vorgestellt (Sherif et al. 2003). In dieser Studie wurde untersucht, ob ein Zusammenhang zwischen dem Knochenalter und dem Volumen des ossifizierten Anteils des Os calcaneus besteht. Bei 400 Kindern wurden dafür sonographische Volumenbestimmungen des Os calcaneus durchgeführt, um Referenzstandards zu erstellen. Weitere 100 Kinder unterzogen sich dann der sonographischen Knochenaltersbestimmung, um die Standards auf ihre Validität zu prüfen (Sherif et al. 2003). Es zeigte sich, dass das sonographisch bestimmte Volumen des Fersenbeins gut mit dem kalendarischen Alter übereinstimmte und somit als valide Methode für die Knochenaltersbestimmung erscheint.

Bilgili et al. arbeiteten ein neues Ultraschallverfahren aus, welches auf der Greulich & Pyle-Methode basiert (Bilgili et al. 2003). Um die Validität des neuen sonographischen Verfahrens zu untersuchen, wurden die Handradiogramme von 97 Kindern mit den sonographisch erstellten „hand and wrist ultrasound charts“ dieser 97 Kinder verglichen. Die „hand and wrist ultrasound charts“ stellen Schemata dar, in welchen die Ossifikationszentren der Carpalia, die Radius- und die Ulnaepiphyse sowie die proximale Phalanx des Daumens gekennzeichnet und näher charakterisiert wurden. Diese „hand and wrist ultrasound charts“ wurden anschließend unter Nutzung des Atlases von Greulich & Pyle interpretiert und ein Knochenalter bestimmt. Das Knochenalter des Radiogramms wurde mit dem Knochenalter der „hand and wrist ultrasound chart“ verglichen. Statistisch zeigten beide Bestimmungsverfahren gute Übereinstimmungen. Bei den männlichen Patienten fand sich bei 71,1% eine völlige Übereinstimmung beider Methoden; bei 84,4% war der Unterschied < 6 Monate. Bei den weiblichen Patienten fand sich bei 65,5% eine völlige Übereinstimmung; der Unterschied war bei 88,5% < 6 Monate (Bilgili et al. 2003). Auch die Ergebnisse dieser Studie belegen, dass Ultraschallverfahren eine valide Alternative für die Knochenaltersbestimmung darstellen.

6 Schlussfolgerungen

Es ist bekannt, dass die Geschwindigkeit, mit der sich Ultraschall in Knochen ausbreitet, vom Knochenalter abhängig ist (Barkmann et al. 2002, Rose et al. 2001). Macht man sich dieses Wissen zu Nutze und entwickelt man Methoden, um diese Veränderungen exakt quantifizieren zu können, so stellt die Sonographie eine attraktive Alternative zu den etablierten radiographischen Verfahren zur Knochenaltersbestimmung dar.

Die Vorteile des Ultraschalls gegenüber den konventionellen Methoden liegen in der einfachen und schnellen Applikation, für die es praktisch keine Kontraindikationen gibt, und in seiner wirtschaftlichen Effizienz. Des Weiteren ist Ultraschall ein Verfahren, das völlig frei von ionisierender Strahlung ist. Im Vergleich mit anderen sonographischen Verfahren zur Knochenaltersbestimmung (Castriota-Scanderbeg et al. 1996, Castriota-Scanderbeg et al. 1998) dürften insbesondere die Einfachheit, die Objektivität der Durchführung und der geringere Zeitbedarf für das Bonage-System vorteilhaft sein.

Ein weiterer prinzipieller Vorteil des Ultraschalls ist die Fähigkeit, Knorpelgewebe und somit auch kleine Ossifikationszentren zu identifizieren bzw. darzustellen. Gerade in der Beurteilung von Entwicklung und Wachstum ist dies ein eindeutiger Vorteil gegenüber den röntgenologischen Methoden, die nur mineralisierten Knochen zur Darstellung bringen (Paesano et al. 1998). Für die konventionellen Methoden bedarf es außerdem immer der Anwesenheit eines Facharztes für Radiologie, der die Untersuchung durchführt, sowie dem notwendigen technischen Equipment. Das Bonage-System dagegen kann von jedem genutzt werden, der eine Unterweisung im Umgang mit dem Gerät erhalten hat, unabhängig von der persönlichen medizinischen Qualifikation.

In der vorliegenden Studie konnte für die Analyse der Stichprobenmittelwerte kein signifikanter Unterschied zwischen den Methoden nach Greulich & Pyle und der Bonage-Ultraschallmethode nachgewiesen werden. Die schlechtere Präzision der neuentwickelten Ultraschallmethode verhindert jedoch, dass sich die konventionellen, etablierten Verfahren zum jetzigen Zeitpunkt durch das Ultraschallverfahren generell ersetzen lassen.

Vorteile der sonographischen Methode sind fehlende ionisierende Strahlung, Objektivität und einfache Durchführbarkeit. Schon jetzt stellt die sonographische Knochenaltersbestimmung mit ihren vielen Vorteilen eine valide Ergänzung zu den konventionellen Verfahren dar. So eignet sich das Bonage-System bspw. zur Verlaufsbeurteilung bereits diagnostizierter Wachstums- und Entwicklungsstörungen oder zur Abschätzung des Knochenalters. Im Rahmen der Erstvorstellung von Patienten ist weiterhin die Röntgenaufnahme der Hand erforderlich, um ein dissoziiertes Knochenalter, Fehlbildungen und Mineralisationsstörungen zu erkennen, die mit dem Bonage-System nicht diagnostizierbar sind.

Setzen sich technische Verbesserungen und Weiterentwicklungen auf dem Gebiet der sonographischen Verfahren in den kommenden Jahren fort und werden diese weiterhin gefördert, so werden die Ultraschallverfahren in naher Zukunft mit Sicherheit eine valide, äquivalente Alternative zu den heute angewendeten, konventionellen Verfahren darstellen. Um aber die konventionellen Verfahren verdrängen zu können, müsste die Präzision der sonographischen Methoden deutlich verbessert werden. Zudem müssten die Ultraschallverfahren dahingehend adaptiert werden, dass sie mehrere Ossifikationszentren aus anatomisch verschiedenen Knochengruppen der Hand (Radius- und Ulnaepiphyse, Carpalia und Metacarpalia) analysieren. Dies ist notwendig um z.B. dissoziierte Knochenentwicklungen bei der Bestimmung des Knochenalters zu berücksichtigen. Auch der Einbezug weiterer Parameter - z.B. Knochendichte - würde die Ergebnisse der sonographischen Verfahren, die auf der Beurteilung der Ultraschallgeschwindigkeit in Knorpel und Knochen beruhen, verbessern.

Literaturverzeichnis

- Acheson RM. 1954. A method of assessing skeletal maturity from radiographs; a report from the Oxford child health survey. *J Anat*, 88(4):498–508.
- Acheson RM, Fowler G, Fry EI, Janes M, Koski K, Urbano P, Werfftenboschjj VA. 1963. Studies in the reliability of assessing skeletal maturity from x-rays. I. Greulich-Pyle Atlas. *Hum Biol*, 35:317–49.
- Acheson RM, Vicinus JH, Fowler GB. 1966. Studies in the reliability of assessing skeletal maturity from x-rays. III. Greulich-Pyle Atlas and Tanner-Whitehouse method contrasted. *Hum Biol*, 38(3):204–18.
- Barkmann R, Glüer CC, Heller M. 1999. Methoden der in-vivo Ultraschallmeßtechnik am Skelett: Grundlagen und technische Realisierung. *Journal für Mineralstoffwechsel*, 6(5):22–27.
- Barkmann R, Rohrschneider W, Vierling M, Troger J, de TF, Cadossi R, Heller M, Gluer CC. 2002. German pediatric reference data for quantitative transverse transmission ultrasound of finger phalanges. *Osteoporos Int*, 13(1):55–61.
- Baroncelli GI, Federico G, Bertelloni S, de Terlizzi F, Cadossi R, Saggese G. 2001. Bone quality assessment by quantitative ultrasound of proximal phalanxes of the hand in healthy subjects aged 3-21 years. *Pediatr Res*, 49(5):713–8.
- Bayley N, Pinneau SR. 1952. Tables for predicting adult height from skeletal age: revised for use with the Greulich-Pyle hand standards. *J Pediatr*, 40(4):423–41.
- Berst MJ, Dolan L, Bogdanowicz MM, Stevens MA, Chow S, Brandser EA. 2001. Effect of knowledge of chronologic age on the variability of pediatric bone age determined using the Greulich and Pyle standards. *AJR Am J Roentgenol*, 176(2):507–10.
- Bierich JR. 1976. Die Bedeutung der radiologischen Skeletalterbestimmung für die Klinik. *Radiologe*, 16(9):381–90.
- Bilgili Y, Hizel S, Kara SA, Sanli C, Erdal HH, Altinok D. 2003. Accuracy of skeletal age assessment in children from birth to 6 years of age with the ultrasonographic version of the Greulich-Pyle atlas. *J Ultrasound Med*, 22(7):683–90.
- Bland JM, Altman DG. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, i:307–10.
- Bortz J. 1999. Statistik für Sozialwissenschaftler, Kap 3.2.3, 5.1.2. Fünfte Aufl. Springer-Verlag, 93–4, 138.
- Bull RK, Edwards PD, Kemp PM, Fry S, Hughes IA. 1999. Bone age assessment: a large scale comparison of the Greulich and Pyle, and Tanner and Whitehouse (TW2) methods. *Arch Dis Child*, 81(2):172–3.
- Castriota-Scanderbeg A, De Micheli V, Scarale MG, Bonetti MG, Cammisa M. 1996. Precision of sonographic measurement of articular cartilage: inter- and intraobserver analysis. *Skeletal Radiol*, 25(6):545–9.

- Castriota-Scanderbeg A, Sacco MC, Emberti-Gialloreti L, Fraracci L. 1998. Skeletal age assessment in children and young adults: comparison between a newly developed sonographic method and conventional methods. *Skeletal Radiol*, 27(5):271–7.
- Craven JD, Costantini MA, Greenfield MA, Stern R. 1973. Measurement of the velocity of ultrasound in human cortical bone and its potential clinical importance. An in vivo preliminary study. *Invest Radiol*, 8(2):72–7.
- Cundy P, Paterson D, Morris L, Foster B. 1988. Skeletal age estimation in leg length discrepancy. *J Pediatr Orthop*, 8(5):513–5.
- Daly RM, Rich PA, Klein R. 1997. Influence of high impact loading on ultrasound bone measurements in children: a cross-sectional report. *Calcif Tissue Int*, 60(5):401–04.
- Elgenmark O. 1946. The normal development of the ossification centers during infancy and childhood: clinical, roentgenologic, and statistical study. *Acta Paediatr Scand*, 33(Suppl 1):1–79.
- Fendel H. 1976. Die Methodik der radiologischen Skeletalterbestimmung. *Radiologe*, 16(9):370–80.
- Fredfeldt KE. 1986. Sound velocity in the middle phalanges of the human hand. *Acta Radiol Diagn (Stockh)*, 27(1):95–6.
- Garn SM, Rohman CG, Silverman FN. 1967. Radiographic standards for postnatal ossification and tooth calcification. *Med Radiogr Photogr*, 43(2):45–66.
- Gilli G. 1996. The assessment of skeletal maturation. *Horm Res*, 45 Suppl 2:49–52.
- Grave K. 1994. The use of the hand and wrist radiograph in skeletal age assessment; and why skeletal age assessment is important. *Aust Orthod J*, 13(3):196.
- Greulich WW, Pyle SI. 1959. Radiographic atlas of skeletal development of the hand and wrist. Stanford University Press, Stanford, California.
- Groell R, Lindbichler F, Riepl T, Gherra L, Roposch A, Fotter R. 1999. The reliability of bone age determination in central European children using the Greulich and Pyle method. *Br J Radiol*, 72(857):461–4.
- Heinrich UE. 1986. Die Bedeutung der radiologischen Skeletalterbestimmung für die Klinik. *Radiologe*, 26(5):212–5.
- Hernandez RJ, Poznanski AW, Hopwood NJ. 1979. Size and skeletal maturation of the hand in children with hypothyroidism and hypopituitarism. *Am J Roentgenol*, 133(3):405–8.
- Jergas M, Schmid G. 2005. Radiologic diagnosis of osteoporosis. *Radiologie up2date*, (5):29–46.
- Johnson GF, Dorst JP, Kuhn JP, Roche AF, Davila GH. 1973. Reliability of skeletal age assessments. *Am J Roentgenol Radium Ther Nucl Med*, 118(2):320–7.
- Johnston FE, Jahina SB. 1965. The contribution of the carpal bones to the assessment of skeletal age. *Am J Phys Anthropol*, 23(4):349–54.

- Kemperdick HF. 1981. Die Skeletalter-Bestimmung bei Kindern mit normalem und abweichendem Wachstumsverlauf. *Fortschr Med*, 99(5):152–6.
- Kemperdick HF. 1986. Die Skeletalterbestimmung beim Kind. *Radiologe*, 26(5):216–21.
- Koc A, Karaoglanoglu M, Erdogan M, Kosecik M, Cesur Y. 2001. Assessment of bone ages: is the Greulich-Pyle method sufficient for Turkish boys? *Pediatr Int*, 43(6):662–5.
- Kreitner KF, Schweden FJ, Riepert T, Nafe B. 1998. Bone age determination based on the study of the medial extremity of the clavicle. *Eur Radiol*, 8(7):1116–22.
- Krummenauer F, Doll G. 2000. Statistical methods for the comparison of measurements derived from orthodontic imaging. *Eur J Orthod*, 22(3):257–69.
- Löffler G, Petrides PE. 1998. Biochemie und Pathobiochemie, Kap 29.2.7, 29.4.3, 29.4.8, 29.6.3. 6te Aufl. Springer-Verlag, 824, 837, 839, 851.
- Mainland D. 1954. Evaluation of the skeletal age method of estimating children's development. II. Variable errors in the assessment of roentgenograms. *Pediatrics*, 13(2):165–73.
- Minas K, Minas S. 2004. Bestimmung der Skelettreife nach Tanner-Whitehouse (TW2 und TW3) und Berechnung der prospektiven Erwachsenengröße. Papenburg: Papenburger Druck.
- Mohr A, Barkmann R, Mohr C, Romer FW, Schmidt C, Heller M, Gluer CC. 2004. Quantitativer Ultraschall zur Osteoporosediagnostik. *Röfo*, 176(4):610–17.
- Mora S, Boechat MI, Pietka E, Huang HK, Gilsanz V. 2001. Skeletal age determinations in children of European and African descent: applicability of the Greulich and Pyle standards. *Pediatr Res*, 50(5):624–8.
- Nachtigal C, Wirtz M. 1998. Wahrscheinlichkeitsrechnung und Inferenzstatistik, Kap II.A.1.1.2. Erste Aufl. Juventa-Verlag, 92.
- Nemet D, Dolfen T, Wolach B, Eliakim A. 2001. Quantitative ultrasound measurements of bone speed of sound in premature infants. *Eur J Pediatr*, 160(12):736–40.
- Nessi R, Garattini G, Bazzini E, Zaffaroni R, Lazzerini F. 1997. Ultrasonography assessment of ossification foci of the wrist and pubertal growth spurt. *Radiol Med (Torino)*, 94(1-2):43–6.
- Paesano PL, Vigone MC, Siragusa V, Chiumello G, Del Maschio A, Mora S. 1998. Assessment of skeletal maturation in infants: comparison between two methods in hypothyroid patients. *Pediatr Radiol*, 28(8):622–6.
- Poland J. 1898. Skiagraphic atlas showing the development of bones of the wrist and hand. Smith Elder, London.
- Prader A, Tanner JM, von Harnack G. 1963. Catch-up growth following illness or starvation. An example of developmental canalization in man. *J Pediatr*, 62:646–59.

- Pryor JW. 1907. The hereditary nature of variation in the ossification of bones. *Anat Rec*, 1(4):84–88.
- van Rijn RR, Lequin MH, Robben SG, Hop WC, van Kuijk C. 2001. Is the Greulich and Pyle atlas still valid for Dutch Caucasian children today? *Pediatr Radiol*, 31(10):748–52.
- Roche AF, Davila GH, Pasternack BA, Walton MJ. 1970a. Some factors influencing the replicability of assessments of skeletal maturity (Greulich-Pyle). *Am J Roentgenol Radium Ther Nucl Med*, 109(2):299–306.
- Roche AF, Rohmann CG, French NY, Davila GH. 1970b. Effect of training on replicability of assessments of skeletal maturity (Greulich-Pyle). *Am J Roentgenol Radium Ther Nucl Med*, 108(3):511–5.
- Rose E, Kann P, Jonas IE. 2001. Speed of sound measurements in mandibular and phalangeal bone during growth. *Clin Oral Investig*, 5(1):2–5.
- Rotch TM. 1909. A study of the development of the bones in children by roentgen method, with the view of establishing a developmental index for grading of and the protection of early life. *Trans Assoc Am Phys*, 24:603–30.
- Schindler A. 2003. Schilddrüse, Klimakterium und Hormontherapie. *Journal Menopause*, 10(4):27–31.
- Schmidt RF, Thews G, Lang F. 2000. *Physiologie des Menschen*, Kap 20.4, 20.6. 28te Aufl. Springer-Verlag, 379–85, 388–90.
- Sherif H, Noureldin M, Bakr AF, Mahfouz AE. 2003. Sonographic measurement of calcaneal volume for determination of skeletal age in children. *J Clin Ultrasound*, 31(9):457–60.
- Sievanen H, Cheng S, Ollikainen S, Uusi-Rasi K. 2001. Ultrasound velocity and cortical bone characteristics in vivo. *Osteoporos Int*, 12(5):399–405.
- Sitzmann FC. 2002. *Pädiatrie*, Kap 9.1.2, 9.5.1. Zweite Aufl. Thieme-Verlag, 198–202, 217–21.
- Sproul A, Peritz E. 1971. Assessment of skeletal age in short and tall children. *Am J Phys Anthropol*, 35(3):433–9.
- Tanner J, Whitehouse RH, Cameron M, Marshall WA, Healy MJR, Goldstein H. 1983. *Assessment of skeletal maturity and prediction of adult height*. Academic Press, London.
- Tanner JM, Whitehouse RH, Healy MJR. 1962. Standards for skeletal age. A new system for estimating skeletal maturity from hand and wrist radiographies with standards derived from a study of 2600 healthy British children. *Centre International de L'enfance*, Paris.
- Tanner JM, Whitehouse RH, Marshall WA, Healy MJR, Goldstein H. 1975. *Assessment of skeletal maturity and prediction of adult height (TW 2-Method)*. Academic Press, London New York.

- Todd TW. 1937. Atlas of skeletal maturation (hand). Mosby, St. Louis.
- Vrahoriti H, Damilakis J, Papadokostakis G, Hadjipavlou A, Gourtsoyiannis N. 2004. Bilateral variation in radial bone speed of sound. *Eur Radiol*, 14(6):953–8.
- Wagner UA, Diedrich V, Schmitt O. 1995. Determination of skeletal maturity by ultrasound: a preliminary report. *Skeletal Radiol*, 24(6):417–20.
- Zadik Z, Price D, Diamond G. 2003. Pediatric reference curves for multi-site quantitative ultrasound and its modulators. *Osteoporos Int*, 14(10):857–62.
- Zerin JM, Hernandez RJ. 1991. Approach to skeletal maturation. *Hand Clin*, 7(1):53–62.

Lebenslauf

Persönliche Daten

Name: Marcus Eulenstein
Anschrift: Ortsstraße 22
07646 Rausdorf
Geburtsdatum: 27.08.1980
Geburtsort: Gera
Familienstand: ledig

Schulbildung

09/1987 - 07/1991 Polytechnische Oberschule, Stadtroda
09/1991 - 07/1999 Pestalozzi-Gymnasium, Stadtroda
07/1999 Abitur, Stadtroda

Hochschulausbildung

10/2000 Beginn des Medizinstudiums an der Friedrich-Schiller-Universität
Jena
09/2002 Ärztliche Vorprüfung
09/2003 Erster Abschnitt der Ärztlichen Prüfung
09/2005 Zweiter Abschnitt der Ärztlichen Prüfung

Promotion

01/2004 - 02/2006 Promotionsarbeit am Institut für Diagnostische und
Interventionelle Radiologie der Friedrich-Schiller-Universität Jena

Jena, 17. Januar 2007

Marcus Eulenstein

Danksagungen

Ich bedanke mich ganz herzlich bei Herrn Prof. Dr. med. Dipl.-Chem. Werner A. Kaiser, meinem Doktorvater, für die Vergabe des Themas dieser Promotion. Seiner Unterstützung und seiner Bereitschaft, dieses Promotionsthema zu fördern, verdanke ich die Fertigstellung meiner Promotionsarbeit.

Ganz besonderer Dank gebührt meinem unmittelbaren wissenschaftlichen Betreuer, Herrn Dr. med. H.-J. Mentzel. Sein unerschöpflicher Elan, sein Interesse an dieser Arbeit, seine Anregungen sowie seine konstruktive Kritik haben maßgeblich zu deren Gelingen beigetragen. Bei allen Problemen konnte ich mir seiner Hilfe gewiss sein und mich auf seine Unterstützung verlassen. Für die hervorragende Betreuung - Vielen Dank!

Bei Herrn Dr. Vollandt vom Institut für Medizinische Statistik, Informatik und Dokumentation bedanke ich mich für die ausführliche und geduldige Beratung bei der statistischen Auswertung der Ergebnisse.

Frau Dr. med. S. Vogt möchte ich dafür danken, dass sie mir bei der Einarbeitung in die Thematik behilflich war. Des Weiteren möchte ich mich bei Ihr und Frau Dr. med. C. Vilser für die Mitarbeit an dem Projekt bedanken, die damit ebenfalls einen maßgeblichen Anteil der Arbeit beitrugen.

Der Firma Sunlight Medical Ltd. (Tel Aviv, Israel) gilt mein Dank für die Überlassung des Bonage-Systems und der zuverlässigen Hilfestellung bei allen technischen Fragestellungen.

Nicht zuletzt gilt mein Dank den zahlreichen freiwilligen Patientinnen und Patienten für die geopfert Zeit und die Erduldung der Unannehmlichkeiten.

Neben meiner Familie und meinen Freunden bin ich vor allem einem ganz besonderen Menschen für die liebevolle Betreuung in stressigen Zeiten zu Dank verpflichtet. Ich möchte mich bei euch für das mir entgegengebrachte Vertrauen, die moralische Unterstützung und für die Tatsache bedanken, dass ihr immer für mich da wart, wenn ich euch gebraucht habe.

Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass

- mir die Promotionsordnung der Medizinischen Fakultät der Friedrich-Schiller-Universität Jena bekannt ist,
- ich die Dissertation selbst angefertigt habe und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben sind,
- mich folgende Personen bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts unterstützt haben:
 - Prof. Dr. med. Dipl.-Chem. W.A. Kaiser
 - Dr. med. H.J. Mentzel
 - Dr. med. S. Vogt
 - Dr. med. C. Vilser
 - Dr. R. Vollandt,
- die Hilfe eines Promotionsberaters nicht in Anspruch genommen wurde,
- Dritte weder mittelbar noch unmittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen
- und ich die gleiche, eine in wesentlichen Teilen ähnliche oder eine andere Abhandlung nicht bei einer anderen Hochschule als Dissertation eingereicht habe.

Jena, 17. Januar 2007

Marcus Eulenstein, Verfasser