

Untersuchungen zu manualmedizinischen Befunden des Krankheitsbildes „Low Back Pain“

**Dissertation zur Erlangung des akademischen Grades
doctor philosophiae (Dr. phil.)**

**vorgelegt dem Rat der Fakultät für Sozial- und
Verhaltenswissenschaften der Friedrich-Schiller-Universität Jena**

Steffen Derlien

geboren am 29.06.1969 in Homberg / Efze

Gutachter

1. ___Prof. Gabriel_____
2. ___Prof. Smolenski_____

Tag des Kolloquiums: ___30.06.2009_____

Inhaltsverzeichnis

1. Einleitung / Problemstellung	S. 01
2. Theoretische Grundlagen	S. 07
2.1. Literaturanalyse	S. 07
2.2. Definitionen	S. 12
2.2.1. Testgütekriterien	S. 12
2.2.2. Definitionen und Klassifikationen von Rückenschmerz	S. 13
2.2.3. Manuelle Medizin	S. 15
2.2.4. Entwicklung der manualmedizinischen Schulen in Deutschland	S. 16
2.2.5. Manualmedizinische Befunde bei „Low Back Pain“.....	S. 17
2.2.5.1. Orientierende und segmentale Schmerzuntersuchungen der Lendenwirbelsäule.....	S. 18
2.2.5.2. Orientierende und segmentale Funktionsuntersuchungen der Lendenwirbelsäule.....	S. 21
2.2.5.3. Weniger gebräuchliche Lendenwirbelsäulentests	S. 24
2.2.5.4. Orientierende und segmentale Schmerzuntersuchungen der Sakroiliakalgelenke.....	S. 25
2.2.5.5. Orientierende und segmentale Funktionsuntersuchungen der Sakroiliakalgelenke.....	S. 28
2.2.5.6. Weniger gebräuchliche Iliosakralgelenkstests	S. 30
3. Fragestellungen	S. 31
4. Methodik der Untersuchung	S. 34
4.1. Untersuchungsansatz zur Auswahl der Studienpopulation	S. 34
4.1.1. Deskription und Diskussion der Stichprobe	S. 37
4.1.2. Generisches Assessmentinstrument SF36.....	S. 39
4.2. Methodisch mögliche Ansätze zur Sicherstellung der Validität	S. 42
4.3. Methodisch mögliche Ansätze zur Sicherstellung der Objektivität	S. 44
4.4. Methodisch mögliche Ansätze zur Sicherstellung der Reliabilität	S. 46
5. Empirische Untersuchungen zur manualmedizinischen Diagnostik	
bei Low Back Pain	S. 48
5.1. Untersuchungen zur Validität	S. 48
5.1.1. Strukturierte Expertenbefragung	S. 48
5.1.2. Ergebnisanalyse und Diskussion zur Validität.....	S. 52

5.2. Untersuchung zur Objektivität	S. 55
5.2.1. Standardisierung der Untersuchungsdurchführung	S. 55
5.2.2. Testbeschreibungen	S. 55
5.3. Untersuchungen zur Reliabilität	S. 58
5.3.1. Testdurchführung in den Referenzzentren	S. 58
5.3.2. Ergebnisdarstellung der Reliabilitätsuntersuchungen	S. 61
5.3.3. Ergebnisanalyse der Schmerzuntersuchungen	S. 63
5.3.4. Diskussion zu den Reliabilitäten der schmerzdetektierenden Tests	S. 64
5.3.5. Ergebnisanalyse der Funktionsuntersuchungen	S. 67
5.3.6. Diskussion zu den Reliabilitäten der Funktionsuntersuchungen	S. 70
5.3.7. Ergebnisanalyse der Globalurteile.....	S. 74
5.3.8. Diskussion zu den Reliabilitäten der Globalurteile.....	S. 79
6. Ausblick und Methodenkritik.....	S. 80
7. Zusammenfassung	S. 83
8. Verzeichnisse	S. 86
8.1. Abbildungen und Tabellen.....	S. 86
8.2. Literatur.....	S. 88
9. Anhang	S. 99

1. Einleitung / Problemstellung

„Low Back Pain - a scientific enigma in the New Millennium“.

Diese Worte Nachemsons auf dem Hauptkongress der Deutschen Gesellschaft für Manuelle Medizin im Jahre 2000 zum Thema Rückenschmerz sollen erläutern, wie die Idee zu dieser Dissertationsschrift entstanden ist.

Tatsächlich stellen Rückenschmerzen ein wesentliches gesundheitliches und volkswirtschaftliches Problem dar. 80-90% aller Menschen erkranken mindestens einmal in ihrem Leben an Rückenschmerz, spontane Besserungen sind häufig. Gegenwärtig beträgt die Häufigkeit von Rückenschmerz zu jedem Zeitpunkt (Punktprävalenz) unter Erwachsenen in Deutschland etwa 40%, die Einjahresprävalenz überschreitet 70% und die Lebenszeitprävalenz 80% (Abbildung 1). Rückenschmerz dominiert, laut dem ersten und bis heute einzigen gesamtdeutschen Bundes-Gesundheitssurveys des Robert-Koch-Instituts, die meisten der Morbiditätsstatistiken (Ellert et al. 1999).

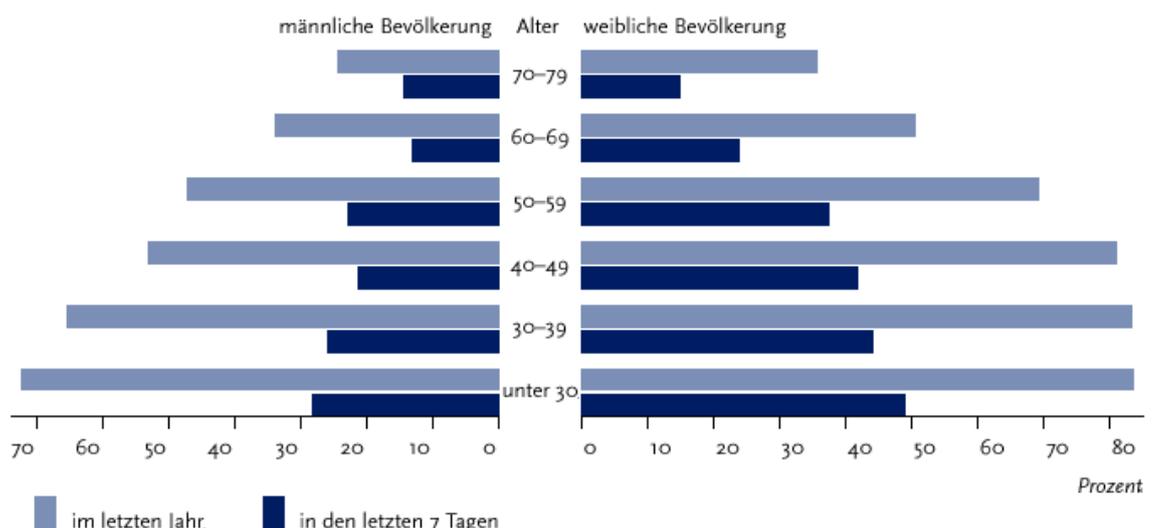


Abb. 1: Rückenschmerzhäufigkeit unterteilt nach Geschlecht und Altersklassen

Quelle: Internet - <http://www.gbebund.de/gbe10/ergebnisse>

Laut Gesundheitsberichterstattung des Bundes 2006 betragen die krankheitsbezogenen Gesamtkosten im Jahr 2004 über 3,3 Milliarden Euro. Das heißt, auf jeden Bundesbürger fallen ca. 40 Euro pro Jahr an Kosten an. Das durchschnittliche Zugangsalter auf Grund verminderter Erwerbsfähigkeit in der gesetzlichen Rentenversicherung sank von 54,6 Jahren (2000) bis auf 53,0 (2005). Dies bedeutet, dass aufgrund der Diagnose Rückenschmerz (M.54 nach ICD-10) eine weibliche Arbeitnehmerin heute rund 1,5 Jahre eher erwerbsunfähig geschrieben wird als noch vor sechs Jahren, bei Männern sind die Zahlen noch bedeutsamer, da der Austritt aus dem Berufsleben rund 2 Jahre früher erfolgt (Gesundheitsberichterstattung des Bundes 2006).

Die Diagnostik des Rückenschmerzes gestaltet sich oft schwierig, da eine Vielzahl möglicher Ursachen in Betracht zu ziehen ist. Vorhandene Befunde, wie z.B. die häufigen radiologischen Zeichen einer degenerativen Wirbelsäulenerkrankung, stehen nicht immer in kausalem Zusammenhang mit der Schmerzsymptomatik. Andererseits lassen sich bei rezidivierenden oder chronischen Kreuz- oder Rückenschmerzen oft keine Einzelursachen finden, welche die Schmerzen hinreichend erklären. Hier wird eine multikausale Genese vermutet, die mit Persönlichkeitsstruktur und Umwelt, insbesondere der Arbeitswelt, in Zusammenhang steht. Erschwerend kommt hinzu, dass durch bestimmte Risikofaktoren („yellow flags“) die Gefahr der Chronifizierung gefördert werden kann (Guzmán et al. 2002)

Die wichtigsten therapeutischen und prognostisch relevanten Ziele der Diagnostik von Rückenschmerz bestehen in der Differenzierung unspezifischer und unkomplizierter Rückenschmerzen von solchen mit radikulären Symptomen und insbesondere von Patienten mit alarmierenden Symptomen („red flags“) (Ehrlich et. al 1999), die weiterer diagnostischer Strategien bedürfen. Am Anfang der Diagnostik und Differenzialdiagnostik von Schmerzen im Rücken steht daher eine ausführliche Anamnese. Neben Fragen nach Anlass, Zeitpunkt des Auftretens, Lokalisation und Ausstrahlung der Schmerzen, sollte auch nach anderen möglichen ursächlichen Erkrankungen gefragt werden (Koes et al. 2001).

Psychische und soziale Momente sollten dabei nicht außer Acht gelassen werden. Hier können standardisierte Assessmentinstrumente wie der „Medical Outcome Study Short Form“ (SF-36) als generisches Instrument oder zum

Beispiel der „Funktionsfragebogen Hannover zum Rücken“ (FFbH-R) als spezifisches Instrument eingesetzt werden. Insbesondere bei Rückenschmerzen und deren bekannter Tendenz zu Rezidiven und zur Chronifizierung kann sich eine Fehldiagnose durchgängig durch die Krankengeschichte ziehen und zu fatalen Folgen für den weiteren Weg des Patienten führen. Das sichere Stellen einer Diagnose zu Beginn der Therapie ist daher sowohl hinsichtlich einer veralteten Terminologie als auch unterschiedlicher Klassifizierungsmodelle von großer Wichtigkeit. Einerseits aus formellen Gründen, da die Einleitung einer Therapie, z.B. Physiotherapie für die gesetzlich Krankenversicherten (GKV), in Deutschland nur laut Heilmittelkatalog möglich ist. Hier muss für die Wahl der Therapiemittel eine Leitdiagnose mit Leitsymptomatiken vorhanden sein. Andererseits aus logistischen Gründen, da die Heilmittelerbringung nicht in derselben Hand wie die Diagnosestellung liegt und somit ein einheitliches Kommunikations- und Informationssystem vorhanden sein sollte.

Diesbezüglich erscheinen jedoch die derzeitig verwendeten Klassifizierungs- und Verschlüsselungssysteme nicht unproblematisch. Verschiedene Autoren schätzen, dass bei bis zu 85% der Patienten, die sich mit Rückenschmerz im Bereich der LWS vorstellen, es nicht möglich ist, eine definitive Diagnose zu stellen. (vgl. Donelson 1997, Riddle 1998, Fritz 2000, Guzmán 2002, Cherkin 2003).

Manuelle Medizin und Evidenz Based Medicine

Laut FIMM Glossar (1998) befasst sich die Manuelle Medizin im Rahmen der üblichen diagnostischen und therapeutischen Verfahren mit reversiblen Funktionsstörungen am Haltungs- und Bewegungssystem. Sie nutzt dabei manuelle diagnostische und therapeutische Techniken, die an der Wirbelsäule und an den Extremitäten zur Auffindung und Behandlung dieser Störungen dienen.

Grundlage für das Verständnis der Manuellen Medizin ist neben solidem schulmedizinischem Wissen unter anderem die Kenntnis der funktionellen Anatomie, Biomechanik, Physiologie und Pathophysiologie. Ziel der diagnostischen Bemühungen ist es demnach, Funktionsstörungen am Haltungs- und Bewegungssystem, speziell der Wirbelsäule, zu erkennen.

Neben der allgemeinen Befunderhebung, der Anamnese, Auswertung bildgebender Verfahren und Sichtung von Laborbefunden bilden spezifische Tests eine Grundlage der manualmedizinischen Diagnostik und der daraus resultierenden Therapie.

Immer häufiger wird in den letzten Jahren der Begriff der „Evidenz basierten Medizin“, kurz EBM thematisiert. EBM stellt den bewussten, expliziten und urteilsfähigen Einsatz der derzeit besten Beweise in der klinischen Entscheidungsfindung für Patienten dar (Sackett et al. 1997). Die in einem Konzept postulierten Prinzipien sollten auf Wissen basieren, das in der Forschungsliteratur zusammengefasst vorliegt. Evidenz vergrößert das Wissen, was für die Gesamtheit der gesicherten Tatsachen und Erkenntnisse steht, die von verschiedenen Berufsgruppen in einem bestimmten Fachgebiet als allgemeingültig akzeptiert werden. Für das Fachgebiet der Manuellen Medizin heißt dies, alle Verfahren und theoretischen Gebäude zur Diagnostik und Therapie müssen sich den Bedingungen modernen wissenschaftlichen Arbeitens stellen.

Zur Evidenz der Manipulationsbehandlung und der Manipulationstherapie gibt es vielzählige Veröffentlichungen (vgl. Aure et al. 2003, Alexiev 1994, Descarreaux 2004, Koes et al. 2001). Stellvertretend sei eine neuere Arbeit von Woodhead und Clough (2005) genannt. Die Autoren geben einen systematischen Überblick über die Evidenz von Manipulationsbehandlungen in der Lendenwirbelsäulenregion. Die Übersicht berücksichtigt 62 „randomised controlled trials“ (RCT) unterschiedlicher Qualität. Von diesen Arbeiten wiesen ca. 25% eine gute Qualität, gemessen an einem international üblichen Scoresystem für RCTs, auf. (vgl. Scherfer 2006) In den Schlussfolgerungen berichten die Autoren über wenig Evidenz bezüglich der Wirksamkeit von Manipulationstherapie bei akuten Beschwerden, jedoch eine klare Evidenz bei chronischem Krankheitsverlauf.

Es ist festzustellen, dass für den Bereich der Manipulationsbehandlung dieses Forschen nach Evidenz stattfindet. Das Anliegen dieser Arbeit ist es nicht, diese Ergebnisse zu werten. Vielmehr ist es ein Anliegen der Arbeit, den Schritt vor der therapeutischen Intervention zu betrachten. Es bedarf eines komplexen diagnostischen Vorgehens, um den Patienten einer adäquaten Therapie zuzuführen. Folglich ist es verständlich, dass alle allgemeingültig akzeptierten

Diagnostikverfahren einem ähnlich kritischen Durchleuchten standhalten müssen.

Diese Tests werden vor allem hinsichtlich der formellen Testgüte unterschiedlich bewertet.

In der Literaturübersicht (4.1) zu dieser Arbeit finden sich in einem Zeitraum von 1992 bis 2006 über 20 Arbeiten, die sich mit Testgütekriterien von manualmedizinischen Untersuchungen zum Thema Low Back Pain beschäftigen. Es entsteht ein sehr heterogenes Bild über die wissenschaftliche Wertigkeit der einzelnen Tests.

Zur Validität der allgemein gültigen Tests existieren im angegebenen Zeitraum lediglich zwei Literaturreviews (vgl. Hestboeck et al. 2000, Patijn 2001). Hestboeck spricht einzelnen Palpations- und Schmerztests akzeptable Validitätswerte zu, während Patijn feststellt, dass den momentan eingesetzten diagnostischen Prozeduren der Nachweis der Validität fehlt. Vereinzelt sind Studien zu finden, die sich mit speziellen Sachverhalten beschäftigen, aber zum gleichen Ergebnis kommen (Mens et al. 2002)

Deutlich mehr Arbeiten existieren zum Sachverhalt der Reliabilität der Untersuchungen. Viele Autoren (vgl. dazu Hicks 2003, Boline 1993, Strender 1997, Laslett und Williams 1994, Sauer et al. 1996) kommen in ihren Arbeiten zum Schluss, dass Schmerztests gegenüber Funktionstests bessere Reliabilitäten erreichen. Die Vorgehensweisen in den Arbeiten sind dabei sehr heterogen. So wird in einigen Untersuchungen nur ein Test angewendet (vgl. Hoogen 1996, Vincent-Smith 1999, Meijne 1999). Andererseits sind Arbeiten zu finden, die Einzeltests aus einem Mehrtestsystem heraus bewerten (vgl. Kokmeyer 2002).

Es existiert eine Vielfalt von Tests zu einer Fragestellung. Wenn es zum Beispiel um die Beweglichkeit der Sakroiliakalgelenke (Bewegungsausmaße im Millimeterbereich) geht, finden sich mehr als sechs verschiedenen Tests in unterschiedlichen Ausgangsstellungen. Diese Heterogenität in der Testauswahl muss reduziert werden, um wissenschaftlich solide Studiendesigns erstellen zu können. Die mangelnde Standardisierung der einzelnen Tests stellt ein weiteres Problem dar. Es müssen klar definierte Kriterien für die Durchführung des Diagnoseinstrumentes vorliegen. Neben der Ausgangsstellung für den Patienten und den Untersucher, müssen auch die Details der

Untersuchungstechnik wie Auflagepunkte der Untersuchungshände, Druckstärke etc. definiert sein. Unterschiedliche Anwendungspräferenzen der manualmedizinischen Schulen erschweren zusätzlich eine qualifizierte Aussage zur diagnostischen Wertigkeit der Tests. Hier bedarf es für eine wissenschaftliche Studie einer Absprache und des Übereinkommens aller manualmedizinischer Schulen, um die Nachvollziehbarkeit von diagnostischen Schlüsselmethoden zu evaluieren. Dies wird in dieser Studie umgesetzt. Das Augenmerk wird dabei auf den Basistechniken liegen, die alle Schulen in gleichem Maße vermitteln. Derartige Techniken werden auf eine wissenschaftlich valide Basis gestellt.

Dem Problem der von vielen Autoren (vgl. van der Wurff et al. 2000, van den Hoogen 1996, Wilson 1999) angesprochenen Heterogenität der Patientengruppe in den einzelnen Studien und den daraus resultierenden Streuungen der Ergebnisse wird in dieser Arbeit Rechnung getragen. Ein besonderes Augenmerk wird daher auf die Einschluss- und Ausschlusskriterien gelegt. Die vorliegende Arbeit berücksichtigt und integriert die wissenschaftlichen Aspekte dieses mehrdimensionalen Themas. Ein Hauptaugenmerk wird die Reliabilitätsbestimmung von manualmedizinischen Tests sein, wobei die Validität dieser Untersuchungstechniken auf die Auswahl der Tests entscheidenden Einfluss haben wird.

Aus diesem Grund wird nach dieser Einleitung (Punkt 1) in der theoretischen Analyse (2.) genauer auf das Problem Low Back Pain, Manuelle Medizin und deren Struktur in Deutschland eingegangen. Im Anschluss daran findet eine Auseinandersetzung mit den manualmedizinischen Untersuchungstechniken bei unterem Rückenschmerz statt, aus denen die Fragestellungen abgeleitet werden (3.). Im Kapitel zur Methodik der Untersuchung (4.) wird die Studienpopulation charakterisiert und auf die Problematik der Validität, Reliabilität und Objektivität eingegangen. Der fünfte Teil der Arbeit zeigt die jeweiligen Untersuchungsansätze mit deren Ergebnissen und der Diskussion hierüber (5.). Im Ausblick und der Methodenkritik (6.) werden die Ergebnisse aufgegriffen und der Zugewinn an Wissen durch diese Arbeit herausgestellt sowie die gewählten Methoden kritisch beleuchtet.

2. Theoretische Grundlagen

2.1. Literaturanalyse

Hicks (2003) stellte fest, dass Schmerzprovokationen auf höherem Reliabilitätsniveau messbar waren als segmentale Mobilität. Als Maßstab gilt hier der sogenannte Kappa-Wert (K-Wert oder K). Das Niveau der Schmerzprovokationstestkappas lag aber mit 0.25 – 0.55 immer noch unter dem von Landis und Koch (1977) als gut postulierten Kappa-Wert von 0.6.

In der Arbeit von Hicks wird die Messbarkeit der segmentale Mobilität mit -0.02 - 0.26 als nicht ausreichend reliabel gewertet.

Zu der Feststellung, dass Schmerztests in Reliabilitätsstudien besser abschneiden, kommen auch andere Autoren, z.B. Boline 1993, Strender 1997 und Laslett und Williams 1994.

Laslett und Williams (1994) untersuchten sieben Schmerzprovokationstests für das Sakroiliakgelenk (SIG). Dabei konnte das Autorenteam für sechs der verwendeten Tests eine Übereinstimmung von 0.76-0.88 für die Interrater-Reliabilität feststellen.

Auch in einer Literaturstudie von Pescioli et al. (1997) weisen alle Provokationstests eine hohe Übereinstimmung auf, wobei einschränkend zu sagen ist, dass lediglich vier Arbeiten dazu gefunden und diskutiert wurden. Von diesen vier Arbeiten werden in einem Fall von 1989 durch Mc. Combe und Mitarbeiter vier von fünf Tests als unzuverlässig bezeichnet.. Die Aussage zu den Schmerzprovokationstests beruht bei Pescioli (1997) lediglich auf der Arbeit von Laslett et al. (1994), die ich vorher schon besprochen habe.

Die Palpations - und Beweglichkeitstests schneiden in dem Literaturreview von Pescioli (1997) schlecht ab. Es gibt keine gefundenen Arbeiten, die akzeptable K-Werte für die Beweglichkeits- und Palpationsmessung angeben. Die Autoren verzichteten aufgrund des schlechten Abschneidens der einzelnen Tests auf eine nähere Beschreibung und kommen klar zur Aussage, dass Beweglichkeitstests ungeeignet für die klinische Untersuchung und Diagnostik der SIG sind.

Hunt (2001) geht allgemeinen physiotherapeutischen oder ärztlichen Untersuchungstechniken an einem Kollektiv aus der Normalbevölkerung nach.

In dieser Arbeit wurde unter anderem das passive Bewegungsausmaß des gestreckten Beinhebens durch einen Untersucher zweimal im Abstand von einer Woche zur Beurteilung der Intrarater-Reliabilität und durch einen zweiten Untersucher innerhalb von 24 Stunden nach der ersten Messung zur Beurteilung der Interrater-Reliabilität gemessen. Unerwarteter Weise ist die Interrater-Reliabilität dieses einfachen orientierenden Tests unter 0.6, die Intrarater-Reliabilität mit 0.8 auf akzeptablen Niveau. Da dieser Test auch im Sportbereich, wie Sportverein, Gesundheitssport und Schulsport angewandt wird, um Verkürzungen der ischiokruralen Muskulatur zu identifizieren, bedarf dies einer gesonderten Überprüfung. Diese soll aber nicht im Rahmen dieser Dissertation vorgenommen werden.

Einige Autoren prüfen in ihren Arbeiten lediglich einen diagnostischen Test. Hier wäre beispielsweise van den Hoogen (1996) zu nennen, der für das *Laséque-Zeichen* bei Patienten ohne Begleitdiagnosen einen K-Wert von 0.33 und mit Begleitdiagnosen von 0.56 feststellt. Beides scheint nach methodologisch korrekten Maßstäben nicht im akzeptablen Bereich.

Vincent-Smith prüft in einer Arbeit von 1999 den *Standing Flexion Test*, im deutschsprachigen Raum als *Vorlauf-Phänomen* bekannt. Er kommt bei ca. 45% Übereinstimmung in den Untersuchungen auf einen Kappa von 0.05 für Interrater und 0.46 für Intrarater.

Meijne (1999) hat ebenfalls nur einen Test untersucht, den *Gillet-Test*, der Reliabilitäten von Kappa 0.08 aufwies.

Kokmeyer et al. prüften 2002 die Reliabilität eines Mehrtestsystems für SIG-Schmerztests. In diesem Testsystem wurde der *Patrick-Test* mit $K=0.62$ und der *Femoral Shear-Test* mit 0.69 gemessen. Der *Gapping-Test* und *Gaenslen* (bei anderen Autoren *Pelvic Torsion* genannt) sind unter der akzeptierbaren Grenze (dem „cut-off“) von 0.6. Immerhin konnten alle 19 asymptomischen Probanden aus der Gesamtzahl der Untersuchten von den Untersuchern diskriminiert werden. Das Auswahlkriterium war das Vorhandensein von drei positiven Teiltests.

Die umfangreichste und qualitativ aufwändigste Arbeit zum Thema erschien 2001 von Patijn im *Journal of Orthopaedic Medicine*. Schon in der Einleitung stellt Patijn fest, dass es zurzeit kein valides und reproduzierbares

Klassifikationssystem für Low Back Pain Patienten gibt. Als einen Grund nennt er das Fehlen von reproduzierbaren Diagnostiktechniken, als zweiten die fehlende Validität der momentan eingesetzten diagnostischen Prozeduren. Für neurologische Untersuchungen wie Muskelschwäche und Reflexauffälligkeiten findet er akzeptable Kappa-Werte von 0.65- 1.0. Auch die Werte für das *Laséque-Zeichen* bzw. *Straight leg raising-Test (SLR)* und den *Cross-SLR* als Zeichen für durale Irritation variieren von 0.66–0.97, also auf akzeptablem Niveau.

Ganz anders ist die Datenlage bei den Inspektionen in neutraler stehender Position. Keine der aufgeführten Untersuchungstechniken zeigt gute Kappa-Werte. Dies gilt auch für die im deutschsprachigen Raum oft untersuchten und diskutierten *Trigger points*. Hier zeigen sich in mehreren Untersuchungen lediglich schlechte Kappa-Werte von 0.12-0.38.

Bewegungsausmaßuntersuchungen scheinen aufgrund der natürlichen Breite der normalen Werte eine limitierte Aussagekraft zu haben. Ebenso eingeschränkt sind die Aussagen bezüglich qualitativer Assessmenttests für lumbale Bewegungsmuster. Hier fehlen zurzeit die Validitätsuntersuchungen. (vgl. Patijn 2001)

Interessant scheint der *Standing Flexion Test (Vorlauf-Phänomen)*. Dieser zeigt in einer Untersuchung von Croft et al. (1998) als einzeln untersuchter Test einen Kappa von 0.03-0.11. In einer Untersuchungskombination mit anderen Tests von Cibulka und Koldehoff (1998) weist dieser Test einen Kappa-Wert von 0.88 auf.

Patijn (2001) kommt zur Aussage, dass einzelne SI-Tests keine oder nur schlechte Reliabilität aufweisen. In Kombination gibt es zwei, die akzeptable Reliabilität zeigen. Dies sind das *Vorlauf-Phänomen* und der *Lateralflexionstest*. Beide Tests sind in der allgemeinen Inspektion bei Rückenschmerzpatienten in der manualmedizinischen Untersuchung auch üblich.

Die Schmerzprovokationstests, die Patijn (2001) gefunden hat, zeigen überwiegend gute Kappa-Werte. Hier schneiden die Schmerzprovokation bei *Ante- und Retroflexion* gut ab, während die Schmerzauslösung bei *Lateralflexion* nicht reproduzierbar zu sein scheint.

Die Untersuchungen aus der stehenden Position werden in der sitzenden Position wiederholt. Hierfür fehlen aber Studien über die Reproduzierbarkeit.

Für die Tests in Rückenlage findet Patijn (2001) variierende Werte von niedrig bis akzeptabel. Das *Laséque-Zeichen* zeigt gute Kappa-Werte, aber der diagnostische Wert des Tests, also seine Validität, ist zweifelhaft. Auch das Autorenteam um van der Wurff (2000) spricht dem Test keinen Nachweis der Validität zu. Vereinzelt berichten Autoren, z. B. van den Hoogen (1996), auch von der schlechten Reproduzierbarkeit des Tests.

Der *Patrick-Test* wird als Einzeltest mit $K=0.23-0.62$ gewertet, bekommt aber als Kombinationstest einen Kappa-Wert von 0.7.

Der *Gaenslen-Test* wird auch von Patijn (2001) in mehreren Arbeiten mit Kappa-Werten von 0.6-0.7 gefunden. Diese Feststellung ist auch in den Arbeiten von Laslett 1994 und Pescioli 1997 zu finden.

Die Aussagen zu den Tests in Rückenlage treffen weitgehend auch für die Bauchlagetests zu. Viele Kappa-Werte sind zu niedrig, um den Test im klinischen Alltag zur Detektion des Einflusses von sakroiliakalen Dysfunktionen auf den Rückenschmerz einzusetzen. Hier wären der *Springing-Test*, *SI-Kompression* und der *Kibler Hautfaltentest* als Beispiele zu nennen. Der *Cranial shear Test*, ein nach kranial gerichteter Druck auf das kaudale Ende des Sakrums, wurde von Patijn (2001) wie auch in der Arbeit von Pescioli (1997) mit akzeptablem Reliabilitätsniveau erwähnt. Auffällig bei den Tests in Bauchlage ist die fehlende Datenlage für die Kombination dieser Einzeltests.

Für die *Seitneige* führt Pescioli (1997) nur noch wenige Tests an, nur einer wurde auf seine Reproduzierbarkeit geprüft. Die ältere Arbeit, aus der dieser Wert stammt, ist von Kirkaldy-Willis et al. (1979). Hier wird lediglich für die *SI-Kompression* ein Kappa-Wert von 0.73 ausgewiesen.

In der Diskussion kommt Patijn (2001) zu dem Schluss, dass die Mehrzahl der manualmedizinischen Untersuchungen im Überblick ein dürftiges Bild mit meistens schlechten und nur vereinzelt akzeptablen Reproduzierbarkeiten zeigt. Manche der Tests, die gebraucht werden, "are based on theories rather than on solid scientific work".

Methodologische Fehler sind eine Hauptquelle für die dokumentierten schlechten Ergebnisse.

Ein zweiter Hauptgrund zum Erklären der schlechten Reproduktionsergebnisse einiger Tests ist das Fehlen der Validität dieser Tests. Hierfür ist aber nach Meinung vieler Autoren (vgl. Pescioli (1997); Patijn (2001); Nachemson (2000))

ein Goldstandard von Nöten, um die verwendeten Tests daran zu messen. Diesen Goldstandard gibt es aber in der Mehrzahl der Fälle nicht.

Die gefundenen Validitätsstudien haben unterschiedliche Designs. In einem Teil der Studien wird mit einer Goldstandart-Messung verglichen. Hier zum Beispiel Röntgenuntersuchungen (Aspegren et al. 1987, Rhodes et al. 1995), oder mechanische Modelle zur Bewegungspalpation (Harvey et al. 1991, Jensen et al. 1993) . Ein anderer Teil der Untersuchungen vergleicht die Tests mit LBP-Fragebögen (Leboeuf-Yde 1990), oder behandelte mit unbehandelten Patienten (Leboeuf-Yde et al. 1988). Alle Ansätze zeigten aber unzureichende Testdurchführungen oder inakzeptable Werte für die Validität (Hestboek et al. 2000)

Den dritten und sicherlich wichtigsten Grund für die Ergebnisse liefert das Krankheitsbild des Low Back Pain selbst. Diese Patientengruppe kann als stark heterogen beschrieben werden, mit unterschiedlichen medizinischen Historien, verschiedenen klinischen Auffälligkeiten und Prognosen. Ein valides Klassifizierungssystem fehlt. (Donelson et al. 1997, Riddle 1998, Koes et al. 2001)

Die genannten Probleme sind zurzeit noch immer aktuell und nicht gelöst. Aus diesem Grund wird das methodische Ausgestalten dieser Arbeit einen wichtigen Stellenwert bekommen und entscheidend für die späteren Ergebnisse werden.

2.2 Definitionen

2.2.1 Testgütekriterien

Wie gut und wie genau diagnostische Tests das erfassen, was sie zu erfassen vorgeben, wird im Allgemeinen mit den Gütekriterien eines diagnostischen Instruments beschrieben. Brauchbare Testverfahren müssen diese bestimmten Anforderungen, die als Testgütekriterien bezeichnet werden, erfüllen. Ein diagnostischer Test wird mit diesen Kriterien ein standardisiertes wissenschaftliches Routineverfahren zur Messung von wichtigen, abgrenzbaren Merkmalen einer Person oder einer Diagnose. Der Test dient vornehmlich zur quantitativen Bestimmung des relativen Grades von individuellen Merkmalsausprägungen, kann aber auch qualitative Aussagen über individuelle Ausprägung von Merkmalen ermöglichen.

Die klassische Testtheorie basiert auf den fünf folgenden Axiomen (Bortz & Döring 2002)

1. Axiom: Das Testergebnis setzt sich additiv aus dem wahren Wert und dem Messfehler zusammen.
2. Axiom: Bei wiederholter Testanwendung kommt es zu einem Fehlerausgleich mit Reduktion des Mittelwertes des Messfehlers, so dass schließlich der wahre Wert repräsentiert wird.
3. Axiom: Wahrer Wert und Messfehler sind unabhängig voneinander.
4. Axiom: Die Höhe des Messfehlers ist unabhängig vom Ausprägungsgrad anderer Persönlichkeitsmerkmale
5. Axiom: Die Messfehler verschiedener Testwiederholungen sind voneinander unabhängig.

Auf der Basis der Axiome unterscheidet Lienert (1998) in Haupt- und Nebengütekriterien. Hauptgütekriterien sind die Objektivität, die Reliabilität und die Validität. Als Nebengütekriterien nennt er die Ökonomie, Nützlichkeit, Normierung und Vergleichbarkeit von Testverfahren.

Weist ein Test diese Gütekriterien nicht auf, kann man im eigentlichen Sinne nicht von einem Test sprechen, da ihm die wissenschaftlich überprüften Grundlagen und notwendigen Kontrolluntersuchungen fehlen.

Nachstehend wird auf die Hauptgütekriterien ausschließlich per Definition eingegangen. Die Spezifizierung der Kriterien und deren Anwendung auf die vorliegende Untersuchung werden im empirischen Teil näher erläutert und im Untersuchungsansatz herausgestellt.

Objektivität

Die Objektivität steht für das Ausmaß, in dem die Ergebnisse eines Tests unabhängig von der Person des Untersuchenden sind.

Reliabilität

Die Reliabilität gibt die Zuverlässigkeit einer Messmethode an. Ein Test wird dann als reliabel bezeichnet, wenn er bei einer Wiederholung der Messung unter denselben Bedingungen und an denselben Gegenständen zum selben Ergebnis kommt. Das Maß ist der Reliabilitätskoeffizient und definiert sich aus der Korrelation der beiden Testungen.

Validität

Die Validität ist das wichtigste Testgütekriterium, denn sie gibt den Grad der Genauigkeit an, mit dem eine Untersuchung das erfasst, was sie erfassen soll.

2.2.2 Definitionen und Klassifikationen von Rückenschmerz

Rückenschmerzen sind Schmerzen in der Region zwischen dem 7. Halswirbel (vertebra prominens) und den Glutealfalten. Nach dem genaueren Ort ihres Auftretens können thorakale, lumbale und sacrale Rückenschmerzen unterschieden werden. Bleiben die Beschwerden auf die eine betroffene Region beschränkt, spricht man von einem lokalen Syndrom, an der Lendenwirbelsäule von einem lokalen Lumbalsyndrom, was dem „simple backache“, dem einfachen unkomplizierten Rückenschmerz entspricht. Strahlen die Schmerzen durch Wurzelkompression nach distal in die untere Extremität aus, so bezeichnet man dies als lumbales Wurzelsyndrom bzw. Ischialgie. Üblicherweise wird zwischen einer Ausstrahlung bis oberhalb oder unterhalb des Kniegelenks unterschieden (vgl. Bigos 1995, Bogduk 1985, Brügger 2001).

Zudem kann je nach Anzahl der gereizten Nervenwurzeln weiterhin zwischen einem mono- und polyradikulären Syndrom unterschieden werden. Die daraus resultierenden unterschiedlichen neurologischen Ausfälle erfordern unterschiedliche therapeutische Konsequenzen.

Für den englischen „low back pain“ (in der Region zwischen dem Unterrand der 12. Rippe und den Glutealfalten) gibt es, bis auf den umgangssprachlichen Begriff „Kreuzschmerz“, keinen deutschsprachigen Begriff.

Oft wird pragmatisch zwischen unspezifischem und spezifischem Rückenschmerz unterschieden. Unspezifische Rückenschmerzen sind als solche definiert, bei denen sich im Gegensatz zu spezifischen Rückenschmerzen, keine Hinweise auf ursächliche Erkrankungen wie z.B. Frakturen, Tumore oder Entzündungsprozesse finden lassen.

Eine weitere Einteilung erfolgt aufgrund des zeitlichen Auftretens der Rückenschmerzen. Es lassen sich akute, rezidivierende und chronische Verläufe finden. Die akuten Schmerzen treten innerhalb weniger Stunden bis maximal einen Tag auf und bestehen weniger als 12 Wochen Gesamtdauer mit variierender Schmerzintensität. Akute Rückenschmerzen, die länger als 6 Wochen bestehen, können auch als subakut bezeichnet werden. (vgl. Neumann 1998)

Rezidivierende Rückenschmerzen sind akute Schmerzen, die nach einem symptomfreien Intervall von mindestens 6 Monaten wieder auftreten. Sie werden als erneute Episode akuter Rückenschmerzen betrachtet und behandelt.

Chronischer Rückenschmerz definiert Rückenschmerz, der an mindestens der Hälfte der Tage in einem Beobachtungszeitraum von 12 Monaten als Einzel- bzw. in multiplen Episoden auftritt. Sie können während dieser Zeit in Intensität und Ausprägung variieren, wobei sich das Schmerzgeschehen aufgrund psychosozialer Faktoren weitgehend verselbständigt hat. (vgl. Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften)

2.2.3 Manuelle Medizin

Wie bereits in der Einleitung erwähnt, befasst sich die Manuelle Medizin im Rahmen der üblichen diagnostischen und therapeutischen Verfahren mit reversiblen Funktionsstörungen am Haltungs- und Bewegungssystem. Sie nutzt dabei manuelle diagnostische und therapeutische Techniken, die an der Wirbelsäule und an den Extremitäten zur Auffindung und Behandlung dieser Störungen dienen.

Neumann (1998) definiert: „die Manuelle Medizin erkennt und behandelt Funktionsstörungen am Bewegungssystem, dem Nervensystem und den inneren Organen. Das Vorgehen wird der individuellen Situation des Patienten angepasst, Einflüsse der Umwelt (Arbeitsplatz, Lebensgewohnheiten) und präventive Maßnahmen werden in das Behandlungskonzept einbezogen.“

„Manuelle Medizin (MM) ist die medizinische Disziplin, in der unter Nutzung der theoretischen Grundlagen, Kenntnisse und Verfahren weiterer medizinischer Gebiete die Befundaufnahme am Bewegungssystem, dem Kopf, viszeralen (die Eingeweide betreffend) und bindegewebigen Strukturen sowie die Behandlung ihrer Funktionsstörungen mit der Hand unter präventiver, kurativer und rehabilitativer Zielsetzung erfolgt. Diagnostik und Therapie beruhen auf biomechanischen und neurophysiologischen Prinzipien.“ (vgl. Bischoff, 2005)

Das „Knochensetzen“ ist so alt, wie die Menschheit selbst. Schon in frühen Zeiten und bei vielen Völkern hat es Kundige gegeben, die in der Lage waren, durch Handgriffe Beschwerden an Wirbelsäule und Extremitäten zu lindern oder zu beseitigen. In Thailand zeigen 4000 Jahre alte Skulpturen entsprechende Techniken. Auch in den alten Kulturen der Chinesen, Inder, Ägypter, Babylonier und Inkas wurden Manipulationen beschrieben. Hippokrates (460-377 v. Chr.) beschrieb in seinem Werk „De articulis“ die Rolle der Wirbelsäule, in den Werken von Galen (129-199 n. Chr.) werden Manipulationstechniken beschrieben und aus den Schriften von Avicenna (980-1037 n. Chr.) und Hildegard von Bingen (1098-1179) wissen wir, dass Manipulationen ein Bestandteil der damaligen Heilkunst waren. Die Wurzeln der heutigen Manuellen Medizin, wie die Lehre von den Handgriffen gegenwärtig genannt wird, hat sich in den letzten 50 Jahren erheblich weiterentwickelt und ist von der

Chiropraktik und der Osteopathischen Medizin, beide aus den USA kommend, auf eine wissenschaftliche Grundlage gestellt worden (Neumann, 1998).

Als Behandlungstechniken kommen in der Manuellen Medizin folgende Techniken zur Anwendung: Mobilisation, Manipulation, Weichteiltechniken, Neuromuskuläre Therapien.

Neben dem Begriff Manuelle Therapie werden synonym die Begriffe Chirotherapie, Funktionspathologie des Bewegungssystems, muskuloskelettale Medizin und funktionelle Neuroorthopädie gebraucht (vgl. Bischoff 2005, Vernon 1999).

2.2.4 Entwicklung der manualmedizinischen Schulen in Deutschland

In Deutschland ließen sich nach 1945 im Ausland ausgebildete Osteopathen und Chiropraktiker nieder und erzielten mitunter spektakuläre Erfolge, die mit herkömmlichen Methoden nicht erreichbar waren. Verschiedene kleine Arztgruppen begannen sich mit der neuen Methode auseinanderzusetzen. Sie gründeten im Jahre 1953 die Forschungsgemeinschaft für Arthrologie und Chirotherapie (FAC) in Hamm mit heutigem Sitz in Boppard und die Gesellschaft für Manuelle Wirbelsäulen- und Extremitätenbehandlung (MWE – heute Ärzteseminar Dr. Karl Sell) in Neutrauchburg. Beide schlossen sich 1966 zur Deutschen Gesellschaft für Manuelle Medizin (DGMM) zusammen, welche im Rahmen der 1968 gegründeten Internationalen Gesellschaft für Manuelle Medizin (FIMM) eng mit den nationalen Gesellschaften für Manuelle Medizin von 25 anderen Ländern zusammenarbeitet. Seit der Wiedervereinigung ist auch die Ärztevereinigung für Manuelle Medizin – Ärzteseminar Berlin (ÄMM) Mitglied der DGMM. Gegründet wurde sie 1972 in Oberhof als Sektion Manuelle Therapie der Gesellschaft für Physiotherapie.

Das Ziel der DGMM ist es, jeden interessierten Arzt über die Manuelle Medizin zu informieren und jeden, der sich mit der funktionellen Erkrankung des Haltungs- und Bewegungssystems beschäftigt, zu instruieren. Erst Anfang der 90er Jahre führten die Ärzteseminare Konsensgespräche, welche Definitionen und Techniken auf eine gemeinsame Grundlage stellten. Die gemeinsame Qualitätssicherung wurde seither kontinuierlich weiterentwickelt.

Die einzelnen Schulen oder Seminare der Deutschen Gesellschaft haben aber dennoch ihre besonderen Schwerpunkte behalten. Während die ÄMM die muskulären und die myofascialen Techniken besonders betont, sind es in der MWE schwerpunktartig die sanften Manipulationen in die schmerzfreie Richtung, bei der FAC wiederum die Mobilisationstechniken.

Die aufgezeigten Unterschiede äußern sich vor allem in der Wirbelsäulenbehandlung, während bei der Behandlung der Extremitätengelenke zwischen den Schulen der DGMM nur noch unwesentliche Unterschiede bestehen.

2.2.5 Manualmedizinische Befunde bei „Low Back Pain“

In der schon erwähnten Arbeit von Patijn (2001) sind eine Vielzahl von Untersuchungen aufgelistet, die zur Diagnosefindung bei „low back pain“ international üblich sind. Patijn wählt die Einteilung der Tests nach deren Ausgangsstellung während der Untersuchung. Diese Unterteilung ist Inspektion in neutraler stehender Position, Palpation von Schmerzzonen in neutraler stehender Position, „Range of Motion“- Untersuchungen in stehender Position, Schmerzprovokationstests im Stehen, Inspektion in neutraler sitzender Position, Tests in Bauchlage und Tests in Rückenlage. Patijn hat insgesamt über 70 Tests gefunden, die von einer Varus/Valgus Inspektion der Knie bis hin zu den in Deutschland bekannten und verwendeten Tests, wie zum Beispiel, Patrick, Kreuzgriff oder segmentale Federung gehen. In Deutschland hat sich in den Schulen der DGMM eine Einteilung der Tests als sinnvoll erwiesen, welche Grundlage für die Testauswahl in dieser Studie sein soll.

Insgesamt sind über 36 Tests in Deutschland gebräuchlich und bekannt. Hierbei bleiben die schulinternen Weiterentwicklungen und die Abwandlungen durch einzelne Lehrgruppen ungezählt und spielen für diese Studie keine Rolle. Allen Lehrmeinungen konform lässt sich die Vielzahl der Untersuchungen in zwei große Blöcke unterteilen, die schmerzdedektierenden Untersuchungen und die Funktionsuntersuchungen. In jedem dieser Blöcke gibt es orientierende und segmentale Tests. Dieses Schema wird im Folgenden auf die Lendenwirbelsäule und auf die Sakroiliakalgenke angewendet (vgl. Patijn 2001).

Daraus ergibt sich folgendes Betrachtungsschema:

Schmerzuntersuchungen	Lendenwirbelsäule	Orientierend
		Segmental
	Sakroiliakalgelenke	Orientierend
		Segmental
Funktionsuntersuchungen	Lendenwirbelsäule	Orientierend
		Segmental
	Sakroiliakalgelenke	Orientierend
		Segmental

Tab. 1: Betrachtungsschema zur Schmerz- und Funktionsuntersuchung der LWS und SIG

2.2.5.1. Orientierende und segmentale Schmerzuntersuchungen zur Lendenwirbelsäule

Im Folgenden sollen alle gebräuchlichen Tests kurz erläutert werden. Die mit Foto und Metrik dokumentierten Tests sind die, die aus dem Kapitel 5.1.1 über das Expertenrating hervorgegangen sind. Aus Gründen der besseren Übersicht und der Verständlichkeit werden die Tests an dieser Stelle vorgestellt, obwohl das Auswahlverfahren erst später beschrieben wird.

Gebräuchliche orientierende Tests sind:

Lasegue oder Straight leg raising (SLR)

Der Untersucher hebt das im Knie gestreckte Bein des Patienten, der in Rückenlage liegt, langsam bis zu dem Punkt an, an dem der Patient Schmerzen angibt. (Buckup 2005)

Crossed SLR oder Lasegue-Moutaud-Martin-Zeichen

Der Untersucher hebt das im Kniegelenk gestreckte, nicht schmerzhaftes Bein bis zu dem Punkt an, an dem der Patient Schmerzen angibt (Buckup 2005).

Globale Federung

<i>Globale Federung</i>	
Ausgangsstellung: Bauchlage, Arme entlang des Körpers auf der Liege, keine Unterlagerung der Lordose (nur bei starker Hyperlordose kleiner Ausgleich) Durchführung: Thenar des Untersuchers liegt über den Dornfortsätzen, die ulnare Kante bei L4, Tiefenkontakt und Vorspannung herstellen und dosierten, ventralisierenden Schub auslösen.	
Metrik	
Schmerz:	Ja / Nein

Abb. 2: *Globale Federung* - Durchführung und Metrik

Hyperextensionstest

Der Patient liegt auf dem Bauch. Der Untersucher fixiert beide Beine und fordert den Patienten auf, den Oberkörper anzuheben.

3-Stufen-Hyperextensionstest

Der Patient liegt auf dem Bauch. Der Untersucher umfasst mit einer Hand das gestreckt liegende Bein und hebt es in der 1. Phase, unter Gegenhalt der anderen Hand am Becken in die Hyperextension. In der 2. Phase wird das Sakrum parallel zum Iliosakralgelenk fixiert und das Bein in Hyperextension gebracht. In der 3. Phase fixiert die eine Hand mit der Handwurzel den 5. Lendenwirbelkörper, die andere Hand führt das Bein in die Überstreckung. Durch Verschiebung der Fixationshand nach kranial lassen sich auch höher gelegene LWS-Segmente testen.

Gebräuchliche segmentale Tests sind:

Aufsuchen von *Triggerpoints*

Vor allem Professor Janet. G. Travell und Professor David. G. Simons haben in der zweiten Hälfte des letzten Jahrhunderts "muskuläre Triggerpunkte" als Ursache von Schmerzen im Bewegungsapparat beschrieben. Diese können lokale Verdickungen, Verhärtungen und Verkürzungen sowie auch Schmerzen verursachen. Diese Schmerzen können aber auch an anderen Orten die sich in der Nähe des Triggerpunktes, aber auch weit davon entfernt befinden können („referred pain“) auftreten. Auf Grund dieses "Auslösens" von entfernten Schmerzen haben sie auch von J.G.Travell den Namen "*Trigger-Points*" (englisch: trigger = Auslöser) erhalten (Travell & Simons 2004)

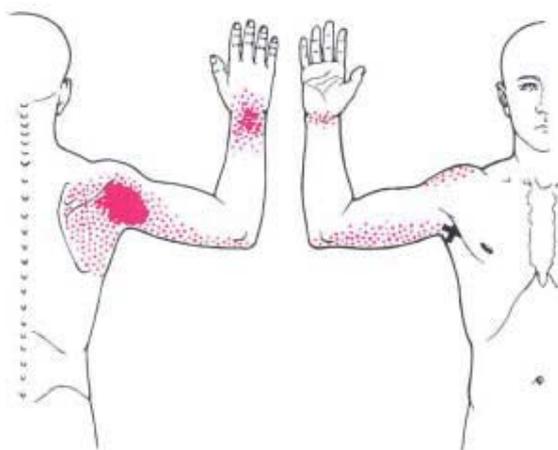


Abb. 3: Übertragender Schmerz (Referred Pain) bei Palpation eines aktiven *Trigger-Point* (modifiziert nach Travell und Simons) aus Pongratz 2003

Segmentale Federung

Segmentale Federung	
<p>Ausgangsstellung: Bauchlage, Arme entlang des Körpers auf der Liege, keine Unterlagerung der Lordose (nur bei starker Hyperlordose kleiner Ausgleich)</p> <p>Durchführung: Zeige- und Mittelfinger der „Spürhand“ liegen auf der Lamina, der Hypothena der „Druckhand“ übt einen ventralisierenden Druck aus. Bei Untersuchung L5 kommen Hände von kranial, bei L4 von caudal.</p>	
Metrik	
Schmerz L4:	Ja / Nein
Schmerz L5:	Ja / Nein

Abb. 4: *Segmentale Federung* – Durchführung und Metrik

Druckschmerz-Dornfortsatzreihe oder Dornfortsatz-Klopfest und Paravertebraler Druckschmerz

Der Patient liegt in leichter Vorbeugung. Mit dem Reflexhammer klopft man auf die Dornfortsätze der Lendenwirbelsäule und paraspinal auf die Muskulatur.

Kibler-Hautfaltentest

Der Patient liegt auf dem Bauch mit locker zurückgelegten Armen. Der Untersucher hebt zwischen Daumen und Zeigefinger eine Hautfalte ab und „rollt“ sie den Rumpf entlang ab (Buckup 2005).

2.2.5.2. Orientierende und segmentale Funktionsuntersuchungen der Lendenwirbelsäule

Gebräuchliche orientierende Tests sind:

Finger-Boden-Abstand Test (FBA) oder Ott-Schober -Zeichen

Der Patient steht. Es werden jeweils der Dornfortsatz C7 und ein Punkt 30 cm kaudal markiert. Des Weiteren wird Dornfortsatz S1 und 10 cm kranial ein Punkt markiert. Das Verschieben der Marker bei Vorneigung und maximaler Rückneigung wird registriert.

Seitneige im Stehen

<i>Seitneige im Stehen</i>			
<p>Ausgangsstellung: Stehen; Füße fußbreit auseinander; symmetrische Belastung</p> <p>Durchführung: Patient neigt sich zur rechten, dann zur linken Seite; Fingerführung am Tractus iliotib.; Untersucher kontrolliert mit Hand an der jeweiligen Schulter exakten Bewegungsablauf</p>			
<p>Metrik</p> <p>gestörte Bewegungssymmetrie :</p> <p>Bewegungsumfang gemindert rechts:</p> <p>Bewegungsumfang gemindert links:</p> <p>gestörte Bewegungsharmonie rechts:</p> <p>gestörte Bewegungsharmonie links:</p>	<p>Ja</p> <p>Ja</p> <p>Ja</p> <p>Ja</p> <p>Ja</p>	<p>/</p> <p>/</p> <p>/</p> <p>/</p> <p>/</p>	<p>Nein</p> <p>Nein</p> <p>Nein</p> <p>Nein</p> <p>Nein</p>

Abb. 5: *Seitneige* – Durchführung und Metrik

Gebräuchliche segmentale Tests sind:

Segmentale Mobilität bei Seitneigung (im Sitzen)

Testdurchführung analog dem Test „Seitneige im Stehen“, Durchführung im Sitzen.

Anteflexion in Seitlage

<i>Anteflexion in Seitlage</i>	
<p>Ausgangsstellung: Seitlage am Bankrand, Knie- und Hüftgelenk sind gebeugt; der Untersucher steht vor dem Patienten</p> <p>Durchführung: Der Untersucher übernimmt die Last der Beine mit beiden Oberschenkeln. Von den Beinen her leitet der Untersucher die Flexion der LWS ein. Ein Arm des Untersuchers fixiert mit dem Ellbogen von dorsal den Thorax des Patienten, zwei Finger halten den kranialen Partnerwirbel, während die Handwurzel der anderen Hand die Bewegung des Beckens unterstützt.</p>	
<p>Metrik</p> <p>Funktionseinschränkung L4: Ja / Nein</p> <p>Funktionseinschränkung L5: Ja / Nein</p>	

Abb. 6: *Anteflexion* – Durchführung und Metrik

Irritationspunkt in Bauchlage

<i>Irritationspunkt in Bauchlage</i>	
<p>Ausgangsstellung: Bauchlage, Arme entlang des Körpers auf der Liege, keine Unterlagerung der Lordose (nur bei starker Hyperlordose kleiner Ausgleich)</p> <p>Durchführung: Zeige- oder Mittelfinger der Hand tasten am unteren Rand des gleichseitigen Musculus gluteus med., ca. 4 Querfinger caudal der Christa iliaca und 3 Querfinger lateral des Gelenkspaltes. Auf druckdolente Verhärtungen ist zu achten.</p>	
Metrik	
Tonuserhöhung rechts :	Ja / Nein
Tonuserhöhung links :	Ja / Nein

Abb. 7: *Irritationspunkt in Bauchlage* – Durchführung und Metrik

2.2.5.3. Weniger gebräuchliche Lendenwirbelsäulentests

Die nachfolgenden Tests sind in Deutschland bekannt, werden aber seltener eingesetzt. Im Folgenden werden sie nur aufgrund des Anspruchs auf Vollständigkeit genannt, aber nicht näher erklärt. Betrachte ich alle Modifikationen der üblichen Tests, beziehe alle schulinternen Abwandlungen mit ein und zähle alle lehrerspezifischen Handgriffe mit, wird die Anzahl der diagnostischen Möglichkeiten unüberschaubar.

Die in der Literatur wiedergegebenen Tests sind: *Adam-Zeichen*, *Nachlasstest*, *Psoaszeichen*, *Loslasstest* nach *Laseque*, *Unterstützter Vorbeugetest* (*Gürteltest*), *Hoover-Zeichen*.

(vgl. Buckup 2005)

2.2.5.4. Orientierende und segmentale Schmerzuntersuchungen der Sakroiliakalgelenke

Orientierende Tests sind:

Patrick-Test

<i>Patrick-Test</i>																									
<p>Ausgangsstellung: Rückenlage; Behandler steht in Hüfthöhe-Blickrichtung fußwärts, Fixation des linken proximalen OS an der Unterlage durch kopfnaher Hand des Untersuchers, Durchführung: fußnahe Hand des Untersuchers greift von außen in die rechte Kniekehle, Hochziehen des entspannten Beins bis Fuß am medialen Kniegelenksspalt anliegt, anschließend führt Untersucher das Knie nach außen-unten</p>																									
<p>Metrik</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 60%;">gestörte Beweglichkeit rechts :</td> <td style="width: 10%; text-align: center;">Ja</td> <td style="width: 10%; text-align: center;">/</td> <td style="width: 20%; text-align: center;">Nein</td> </tr> <tr> <td>Schmerz rechts :</td> <td style="text-align: center;">Ja</td> <td style="text-align: center;">/</td> <td style="text-align: center;">Nein</td> </tr> <tr> <td>pathologisches Endgefühl rechts :</td> <td style="text-align: center;">Ja</td> <td style="text-align: center;">/</td> <td style="text-align: center;">Nein</td> </tr> <tr> <td>gestörte Beweglichkeit links :</td> <td style="text-align: center;">Ja</td> <td style="text-align: center;">/</td> <td style="text-align: center;">Nein</td> </tr> <tr> <td>Schmerz links :</td> <td style="text-align: center;">Ja</td> <td style="text-align: center;">/</td> <td style="text-align: center;">Nein</td> </tr> <tr> <td>pathologisches Endgefühl links :</td> <td style="text-align: center;">Ja</td> <td style="text-align: center;">/</td> <td style="text-align: center;">Nein</td> </tr> </table>		gestörte Beweglichkeit rechts :	Ja	/	Nein	Schmerz rechts :	Ja	/	Nein	pathologisches Endgefühl rechts :	Ja	/	Nein	gestörte Beweglichkeit links :	Ja	/	Nein	Schmerz links :	Ja	/	Nein	pathologisches Endgefühl links :	Ja	/	Nein
gestörte Beweglichkeit rechts :	Ja	/	Nein																						
Schmerz rechts :	Ja	/	Nein																						
pathologisches Endgefühl rechts :	Ja	/	Nein																						
gestörte Beweglichkeit links :	Ja	/	Nein																						
Schmerz links :	Ja	/	Nein																						
pathologisches Endgefühl links :	Ja	/	Nein																						

Abb. 8: *Patrick-Test* – Durchführung und Metrik

Drei-Bändertest

Der Patient befindet sich in Rückenlage. Zur Prüfung des Ligamentum iliolumbale wird das Bein sowohl im Kniegelenk wie auch im Hüftgelenk gebeugt und zur Gegenhälfte adduziert. Zur Prüfung des Ligamentum sakrospinale und L. iliosacrale wird das Bein in Hüft- und Kniegelenk maximal gebeugt und in Richtung der gegenseitigen Schulter adduziert. Zur

Untersuchung des Ligamentum sakrotuberale wird das Bein im Hüft- und Kniegelenk maximal gebeugt und zur gleichen Schulter bewegt. Bei allen Untersuchungen wird während der Bewegung über das Kniegelenk ein axialer Druck in Oberschenkellängsrichtung ausgeführt. Nach einigen Sekunden auftretende Dehnungsschmerzen sprechen für eine funktionelle Störung.

Gaenslen-Test

Der Patient liegt in Rückenlage mit der schmerzhaften Seite nah an der Bankkante, so dass das Bein nach unten hängen kann. Das andere Bein wird maximal gebeugt und über dem Bauch festgehalten (Thomas'sche Handgriff). Das heraushängende Bein wird nun vom Prüfer in die Überstreckung gezogen/geedrückt.

Treten dabei Schmerzen im Sakroiliakal-Gelenk auf oder werden bereits vorhandene verstärkt, geht man von einer Gelenksdysfunktion aus.

3-Stufen-Hyperextensionstest

Der Patient liegt auf dem Bauch. Der Untersucher umfasst mit einer Hand das gestreckt liegende Bein und hebt es in der 1.Phase, unter Gegenhalt der anderen Hand am Becken, in die Hyperxtension. In der 2. Phase wird das Sakrum parallel zum Iliosakralgelenk fixiert und das Bein in Hyperextension gebracht. In der 3. Phase fixiert die eine Hand mit der Handwurzel den 5. Lendenwirbelkörper, die andere Hand führt das Bein in die Überstreckung. Durch Verschiebung der Fixationshand nach kranial lassen sich auch höher gelegene LWS-Segmente testen

Segmentale Tests sind:

Federungstest (Springing Test)

Der Patient liegt in Bauchlage. Der Untersucher palpiert mit Zeige- und Mittelfinger die Gelenkfortsätze bzw. Laminae der zu untersuchenden Wirbelkörper. Mit der Ulnarkante der anderen Hand, die quer über den Palpationsfingern liegen soll, werden leichte, federnde Stöße dorsoventral ausgeführt, die durch die Palpationsfinger auf die Gelenkfortsätze bzw.Laminae der zu untersuchenden Wirbelkörper übertragen werden. (Buckup 2005)

Druckschmerz Untersuchungen

Der Patient liegt auf dem Bauch. Der Zeigefinger der Untersucherhand wird nacheinander die Druckschmerzpunkte auf den Sakroiliakalgelenken rechts und links (Lokaler Druckschmerz SIG), das Ligamentum Iliolumbale rechts und links (Druckschmerz Ligamentum Iliolumbale), sowie das Ligamentum Sakroiliakale (Druckschmerz Ligamentum Sacroiliacale) palpieren. Die Auslösung eines Druckschmerzes spricht bei diesen Tests für eine Irritation der getesteten Strukturen.

Irritationspunkte

Test ist in Abb. 7 dargestellt.

Kreuzgriff

<i>Kreuzgriff</i>			
Ausgangsstellung: Bauchlage, Arme entlang des Körpers auf der Liege, keine Unterlagerung der Lordose (nur bei starker Hyperlordose kleiner Ausgleich) Durchführung: Untersucher nimmt mit der ulnaren Handwurzel der von kaudal kommenden Hand Kontakt an einer SIPS, die Finger zeigen dabei nach kranial, mit der anderen von kranial kommenden Hand an die Sacrumspitze, Finger zeigen dabei nach kaudal. Leichter Druck aus den gestreckten Armen heraus, führt zur Federungsbewegung zwischen Sakrum und Ilium in Gegennutationsrichtung.			
Metrik			
Schmerz L4:	Ja	/	Nein
Schmerz L5:	Ja	/	Nein
Funktionseinschränkung rechts :	Ja	/	Nein
Funktionseinschränkung links :	Ja	/	Nein

Abb. 9: *Kreuzgriff* – Durchführung und Metrik

Segmentale Tests sind:

Irritationspunkte

Test ist in Abb. 7 dargestellt.

Federungstest (Springing Test)

Der Patient liegt in Bauchlage. Der Untersucher palpiert mit Zeige- und Mittelfinger die Gelenkfortsätze bzw. Laminae der zu untersuchenden Wirbelkörper. Mit der Ulnarkante der anderen Hand, die quer über den Palpationsfingern liegen soll, werden leichte, federnde Stöße dorsoventral ausgeführt, die durch die Palpationsfinger auf die Gelenkfortsätze bzw. Laminae der zu untersuchenden Wirbelkörper übertragen werden (Buckup 2005).

Spine-Test

<i>Spine-Test</i>	
Ausgangsstellung: Stehen; Patient mit Rücken zum Untersucher; Füße fußbreit auseinander; symmetrische Belastung Durchführung: Spinae iliacae post.sup. werden von unten palpiert, Patient hebt rechtes Bein durch Knieflexion vom Boden ab, Bewegung der Spinea werden beobachtet	
Metrik	
Positiv (Blockierung) rechts :	Ja / Nein
Positiv (Blockierung) links :	Ja / Nein

Abb. 11: *Spine-Test* – Durchführung und Metrik

2.2.5.6. Weniger gebräuchliche Iliosakralgelenkstests

Die nachfolgenden Tests sind in Deutschland bekannt, werden aber seltener eingesetzt. Im Folgenden werden sie nur aufgrund des Anspruchs auf Vollständigkeit genannt, aber nicht näher erklärt. Betrachte ich alle Modifikationen der üblichen Tests, beziehe alle schulinternen Abwandlungen mit ein und zähle alle lehrerspezifischen Handgriffe mit, wird die Anzahl der diagnostischen Möglichkeiten unüberschaubar.

Die in der Literatur wiedergegebenen Tests sind:

Stoddard, Yeoman-Test, Laguerre-Test und Abduktionsbelastungstest (vgl. Buckup 2005).

3. Fragestellungen

Aus den beschriebenen Definitionen und Entwicklungen der Manuellen Medizin in Deutschland ergeben sich grundlegende Fragen, die es zu überlegen und zu bearbeiten gilt. Seit über zwanzig Jahre wird in Deutschland, dies- und jenseits der ehemaligen innerdeutschen Grenze in mittlerweile über fünfzig unterschiedlichen Ausbildungsstätten Manuelle Medizin gelehrt. Jedes Ärzteseminar hat eigene Schwerpunkte und Präferenzen in der Anschauungsweise muskulo-skelettaler Probleme.

Kommen unterschiedlich diagnostizierende Ärzte zu einem Konsens, könnte der Weg dahin ohne Bedeutung sein. Wird also die Diagnose der Meniskopathie am Kniegelenk aus der einen oder anderen Sichtweise gestellt, ist dies für die weitere Behandlung, insofern es nur eine resultierende gibt, irrelevant. Da es aber in der Praxis diese eine Therapie nicht gibt, sondern über eine ganze Fülle von Therapien gesprochen werden kann, entsteht hieraus ein weiterer Problemkreis.

Das diagnostische Problem ist bei den Extremitätengelenken, bis auf kleine Ausnahmen in den unterschiedlichen Seminaren, im Konsens einheitlich gelöst worden.

Uneinig ist man sich aber, wenn es sich um Probleme an der Wirbelsäule handelt. Es macht einen erheblichen Unterschied, ob über Mobilisationen oder Manipulationen in Schmerzrichtung, um die beiden extremen Pole zu nennen, nachgedacht wird.

Die Grundlage einer jeglichen Therapie, egal ob wir passive oder aktive Therapien betrachten, ist die exakte Stellung der Diagnose. Die dafür notwendigen diagnostischen Mittel aus manualmedizinischer Sicht müssen zweifelsfrei von allen Erbringern dieser Leistung anerkannt und beherrscht werden.

Ein wichtiger Untersuchungsschwerpunkt für diese Studie sollen die Grundlagen der Diagnosestellung in der Ausbildung von Manualmedizinern sein. Die im Jahre 2003 vom Deutschen Ärztetag beschlossene neue Weiterbildungsordnung verlängerte die Ausbildungszeit für Manuelle Medizin auf 320 Stunden. Mit dieser Ausweitung der Ausbildungszeit regelt sich nur der zeitliche Rahmen für die Ausbildung, die Inhalte bedürfen weiterhin einer

Vereinheitlichung. Ein Weg dahin ist das in Zusammenarbeit zwischen DGMM und Bundesärztekammer entstandene Kursbuch, das die Weiterbildungsinhalte festlegt. Es besteht aber weiterhin das Problem, dass konkrete diagnostische Techniken von den einzelnen Seminaren und Ausbildungsstätten unterschiedlich gelehrt und „verfeinert“ werden.

Aus den genannten Gründen ist diese Studie über diagnostische Techniken angelegt und keine therapeutische Interventionsstudie. Auch die therapeutischen Interventionen müssen sich einer kritischen Überprüfung stellen. Hier fehlen noch zu den meisten Techniken fundierte Untersuchungen, die Theoriegebäude sind nicht mit „harten“ Zahlen und belegten Fakten untermauert. Diese Aufgabe steht aus meiner Sicht noch an, ist nachgeordnet auch zwingend notwendig.

Diese Studie ist auch im Zuge von Prozess- und Qualitätsoptimierung im Sinne der „Evidence Based Medicine“ entstanden, um die gewachsenen Strukturen der deutschen Manuellen Medizin mit den modernen Qualitätsansprüchen an Medizin und deren Ausübung zu verbinden.

Zu diesem Zweck werden die Testgütekriterien der gebräuchlichen manualmedizinischen Untersuchungen bei unterem Kreuzschmerz oder besser „Low Back Pain“ geprüft. Dieses Krankheitsbild ist, wie oben schon beschrieben, sowohl in seiner zeitlichen Ausdehnung wie auch in der Intensitätsausprägung als sehr heterogen gekennzeichnet. Es wird eine Studienpopulation zu finden sein, die repräsentativ hierfür ist.

Aus den genannten Überlegungen heraus wurden folgende Fragestellungen abgeleitet:

Frage I: Welche manualmedizinischen Diagnostiken erfüllen die wissenschaftlichen Kriterien der Validität, die an einen Test gestellt werden?

Frage II: Welche manualmedizinischen Diagnostiken erfüllen die wissenschaftlichen Kriterien der Reliabilität, die an einen Test gestellt werden?

Frage III: Wie kann für ausgewählte diagnostische Prozeduren die Objektivität sichergestellt werden.

4. Methodik der Untersuchung

4.1. Untersuchungsansatz zur Auswahl der Studienpopulation

Studien, die am Ende der 90er Jahre in der Bundesrepublik durchgeführt wurden, ergaben auf die Frage: „Leiden Sie heute an Rückenschmerz?“, dass etwa 40 % der Befragten mit „Ja“ antworteten. Rückenschmerzen im letzten Jahr gehabt zu haben, gaben etwa 70% der Befragten an und überhaupt schon einmal unter Rückenschmerz gelitten zu haben, ungefähr 80% (Raspe / Kohlmann, 1998).

Das heißt wiederum, es können keine altersmäßige Einschränkung im Krankheitsbild festgestellt werden. Aus diesem Grund wird bei der Patientenrekrutierung für den Altersabschnitt lediglich die Einschränkung der Volljährigkeit vorgenommen. Erfahrungsgemäß ist das mittlere Alter der Patienten, die wegen Rückenschmerz in der ärztlichen Praxis erscheinen, zwischen der 3. und 6. Lebensdekade (vgl. Koes et al. 2001, Göbel 2001).

Grund für den Arztbesuch des Patienten sollten unspezifische Schmerzen in der Lendenwirbelsäule oder im Beckenbereich sein. Dieser Schmerz sollte sich, subjektiv vom Patienten eingeschätzt, auf einer visuellen Analogskala über einen Bereich von 20 Punkten, bei einer Maximalzahl von 100 Punkten, bewegen. Diese 20-Punktlinie kennzeichnet die untere Grenze zum Einsatz von schmerzdetektierenden Tests. Des Weiteren stellt dieser Wert die Grenze zum chronifizierten Schmerz dar, der erfahrungsgemäß darunter liegt und sich so grundlegend vom akuten Rückenschmerz unterscheidet.

Aus den Ausführungen in Kapitel 2 ergeben sich hinsichtlich der Schmerzdauer folgende Einschränkungen. Die Schmerzen müssen in den letzten 3 Monaten persistierend oder rezidivierend gewesen sein und dürfen aufgrund der Chronifizierungsgefahr eine Gesamtdauer von 12 Monaten nicht überschritten haben.

Hinsichtlich neurologischer Untersuchungskriterien sollten die Patienten der Studienpopulation keine radikulären Symptomaten mit Hypermobilitäten und muskulären Defiziten aufweisen, da dies über den Low Back Pain hinaus auf strukturelle Probleme hinweist. Hier wäre z.B. ein Bandscheibenvorfall mit

Irritation der Nervenwurzel zu nennen. Aus Gründen der zweifelsfreien Abgrenzung der Rückenschmerzen darf der Patient auch an keinen anderen neurologischen Ausfallsymptomen oder Grunderkrankungen leiden. Hier wären exemplarisch Multiple Sklerose, Morbus Parkinson, Zustand nach Apoplex oder spastische Störungen verschiedener Genese zu nennen.

Viele Untersuchungstechniken erfordern eine sensorische Schulung des Untersuchers, aber auch Sensibilität vom Untersuchten. Deshalb ist es zwingend erforderlich, Patienten mit tonusregulierenden Medikamenten aus der Studie auszuschließen, da bei ihnen regulativ auf die Tiefensensibilität, auf die Schmerzwahrnehmung und auf den Ruhe- und Aktionstonus der wirbelsäulenumgebenden Muskulatur Einfluss genommen wurde.

Manualmedizinische Untersuchungstechniken an der Wirbelsäule basieren größtenteils auf der exakten Palpierung von anatomischen Strukturen und auf dem Auffinden von tief im Gewebe liegenden Strukturen. Diese Grundvoraussetzung für die Diagnosestellung ist bei einer überproportionalen Unterhautfettschicht nicht mehr gewährleistet. Um das Hauptgütekriterium Reliabilität prüfen zu können, erweist sich ein Body-Maß-Index (BMI) von über 32 (Grenze zur Adipositas) als kontraindiziert. In dieser Studie werden folglich nur Patienten mit einem BMI von weniger als 32 eingeschlossen.

Eine weitere Einschränkung im Patientengut stellen die vielfältigen Operationsmöglichkeiten an der Wirbelsäule oder in der Beckenregion dar. Stattgehabte Operationen an der Wirbelsäule können zwar mikroinvasiv durchgeführt worden sein, führen aber in vielen Fällen zum Verzicht von bestimmten Bewegungen, die wiederum für manuelle Untersuchungstechniken essentiell sein können. Als Beispiel hierfür wäre eine Wirbelfusion im Lendenwirbelsäulenbereich zu nennen, eine häufige Methode um Spondylolisthesis (Wirbelgleiten) operativ zu versorgen, nach der Mobilitätsuntersuchungen in diesem Segment nicht mehr indiziert sind. Gleiche operationsbedingte Einschränkungen sind nach Beckenverletzungen zu finden, hier ist zum Beispiel die Beckenringfraktur mit ihren unterschiedlichen Schweregraden A, B und C zu nennen. Dabei sind viele Techniken in Seit-, Bauch- oder Rückenlage nicht mehr möglich. Aus den genannten Gründen werden Patienten mit Operationen oder schweren Traumata an der Wirbelsäule oder dem Becken ebenfalls aus dieser Studie ausgeschlossen.

Daraus ergeben sich folgende neun Einschlusskriterien:

1. Schmerz LWS und/oder Beckenregion
2. in den letzten 3 Monaten persistierend oder rezidivierend
3. nicht länger als 12 Monate
4. VAS aktuell mind. 20/100
5. Keine radikuläre Symptomatik
6. Keine Operationen und /oder Traumata (WS, Becken,)
7. Keine tonusregulierende Medikation
8. Keine neurologische Ausfallsymptomatik (z.B. Spastik, MS, Z.n. Apoplex, Athetose)
9. BMI nicht über 32

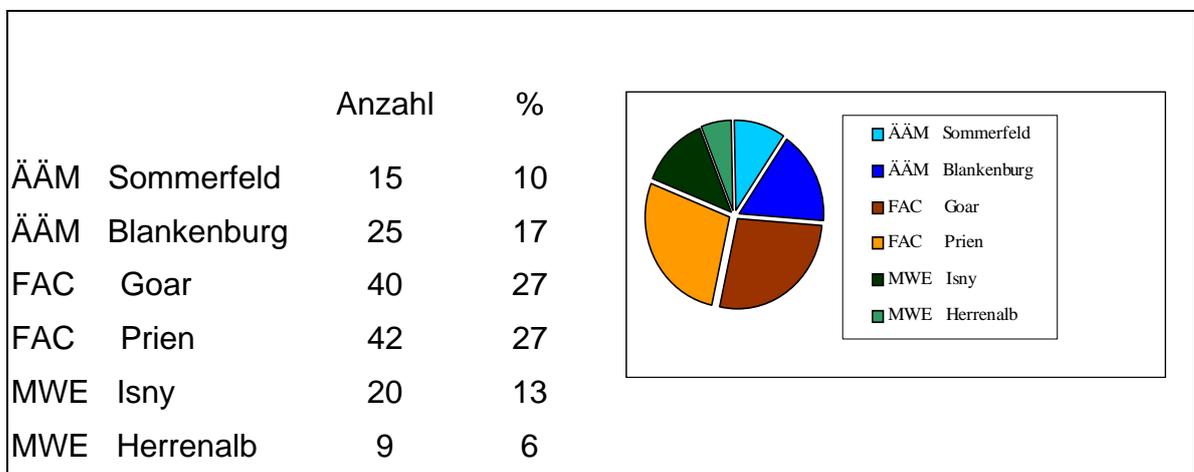
Tab.2: Einschlusskriterien

4.1.1. Deskription und Diskussion der Stichprobe

Insgesamt sind für die Studie 151 Patienten vorgesehen. Diese verteilen sich auf die einzelnen Untersuchungszentren wie in untenstehender Tabelle 3 dargestellt. Hierbei ist festzuhalten, dass etwas über 50% der Patienten aus den FAC assoziierten Kliniken rekrutiert werden konnten. Dieses Ungleichgewicht resultiert hauptsächlich aus Veränderungen in der Zuweisungspraxis für die Kliniken während des Studienzeitraums und kliniksinternen Umstrukturierungen. Aus diesem Grund wird die eigentlich geplante Subgruppenanalyse nach Schulen nicht durchgeführt.

Die Verteilung der Gesamtgruppe nach Geschlechtern entspricht der bundesweiten Verteilung wie sie in dem Report des Statistischen Bundesamtes nachzulesen ist. Mit einem Anteil von 52% Frauen zu 48% Männern bildet sich hier der bundesweite Durchschnitt ab. (Daten des Statistischen Bundesamtes. Zugriff auf das Internet am 19.01.09 unter: <http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Navigation/Statistiken/Bevoelkerung/Bevoelkerung.psml;jsessionid=2C09B503B5E412241E27FE9B1235A824.internet2>)

Die Rohwerte der soziodemographischen Daten sind in den Anhang gestellt.



Tab. 3: Verteilung der Studienpopulation auf die Untersuchungszentren

	Männlich (48%)	Weiblich (52%)
Größe (MW-Angabe)	178	165
Gewicht (MW-Angabe)	86	67
BMI (MW-Angabe)	27	25
Alter (MW-Angabe)	47	48

Tab. 4: Charakterisierung der Studienpopulation nach Größe, Gewicht, BMI und Alter

4.1.2. Generisches Assessmentinstrument „SF-36“

Neben den Identitätsdaten einschließlich der allgemein üblichen Variablen wie Größe, Gewicht und Alter, wurde als Kontrollvariable zum Vergleich der Stichproben der „*Medical Outcome Study Short Form 36*“ (*SF36*) eingesetzt. Der *SF36* ist das weltweit am häufigsten verwendete Instrument zur Erfassung des allgemeinen Gesundheitsstatus. Der selbstadministrierte Fragebogen umfasst 8 Skalen, die je aus 2-10 Items bestehen, sowie einer zusätzlichen Frage zur Veränderung im generischen Gesundheitsstatus. Der *SF-36* erlaubt die Bewertung von 8 Subskalen sowie die Berechnung von 2 Summenscores für eine physische und eine psychische Komponente (Bullinger 2000).

Wie in nachstehender Abbildung 12 zu entnehmen ist, sind Einschränkungen der Probanden in allen acht Subskalen des *SF-36* festzustellen. In der männlichen Population können wir im Item RP „Rollenverhalten wegen körperlicher Funktionsbeeinträchtigung“ (Bullinger 2000) und BP „Schmerzen“ eine 50% Verminderung der Scorewerte feststellen, was eine starke Abweichung von der Normpopulation (Ellert / Bellach 1999) darstellt. Eine mittlere Verschiebung der Scorewerte nach unten kann bei den Items RE „Rollenverhalten wegen seelischer Funktionsbeeinträchtigung“, SF „soziale Funktionsfähigkeit“ und PF „körperliche Funktionsfähigkeit“ festgestellt werden.

Leichte Veränderungen gegenüber der Normpopulation sind in den Items GH „allgemeiner Gesundheitszustand“ und VT „Vitalität“ und MH „seelische Funktionsfähigkeit“ festzustellen, die sich aber im 10-15% Bereich bewegen. Die Ergebnisse der Männer sind nicht überraschend, es ist bekannt, dass durch Schmerzaffektionen solche Einschränkungen zu erwarten sind (vgl. Gerbershagen 1995, Brüggjenjürgen 1994).

In der weiblichen Population sind die Scorewertverschiebungen wie in der männlichen Population zu finden, von der Ausprägung aber etwas geringer und statistisch nicht signifikant. Dies ist ebenfalls aus der Literatur bekannt (vgl. Bullinger 1995).

Einschränkungen in den Summenscores zur physischen (PCS) und mentalen Befindlichkeit (MCS) sind die Folge aus den Verschiebungen der Einzelitems.

Im physischen Score ist dies deutlicher festzustellen, während der mentale Score keine starken Einschränkungen gegenüber der Normpopulation aufweist. Abb. 12 stellt die erreichten Skalenwerte in dieser Population dar und vermittelt einen visuellen Gesamteindruck der erreichten Skalenwerte im Vergleich zur Normpopulation.

Die gesamten Rohwerte der Untersuchungspopulation sind in den Anhang gestellt.

Männlich (n=71)	MV	Normpop. BRD (Ellert / Bellach 1999)
PF	64,6	91,3
RP	46,5	88,4
BP	33,2	71,3
GH	50,7	68,1
VT	41,8	64,2
SF	63,6	89,2
RE	64,4	91,9
MH	59,7	75,2
PCS	37,1	
MCS	44,2	
Weiblich (n= 74)	MV	Normpop. BRD (Ellert / Bellach 1999)
PF	59,7	87,2
RP	42,6	82,5
BP	32,1	63,6
GH	51,7	66,4
VT	41,9	57,4
SF	59,6	84,1
RE	59,3	86,9
MH	57,4	69,4
PCS	36,6	
MCS	43,5	

Tab. 5: Geschlechter getrennte Stichproben- und Normverteilung
(Ellert / Bellach1999)

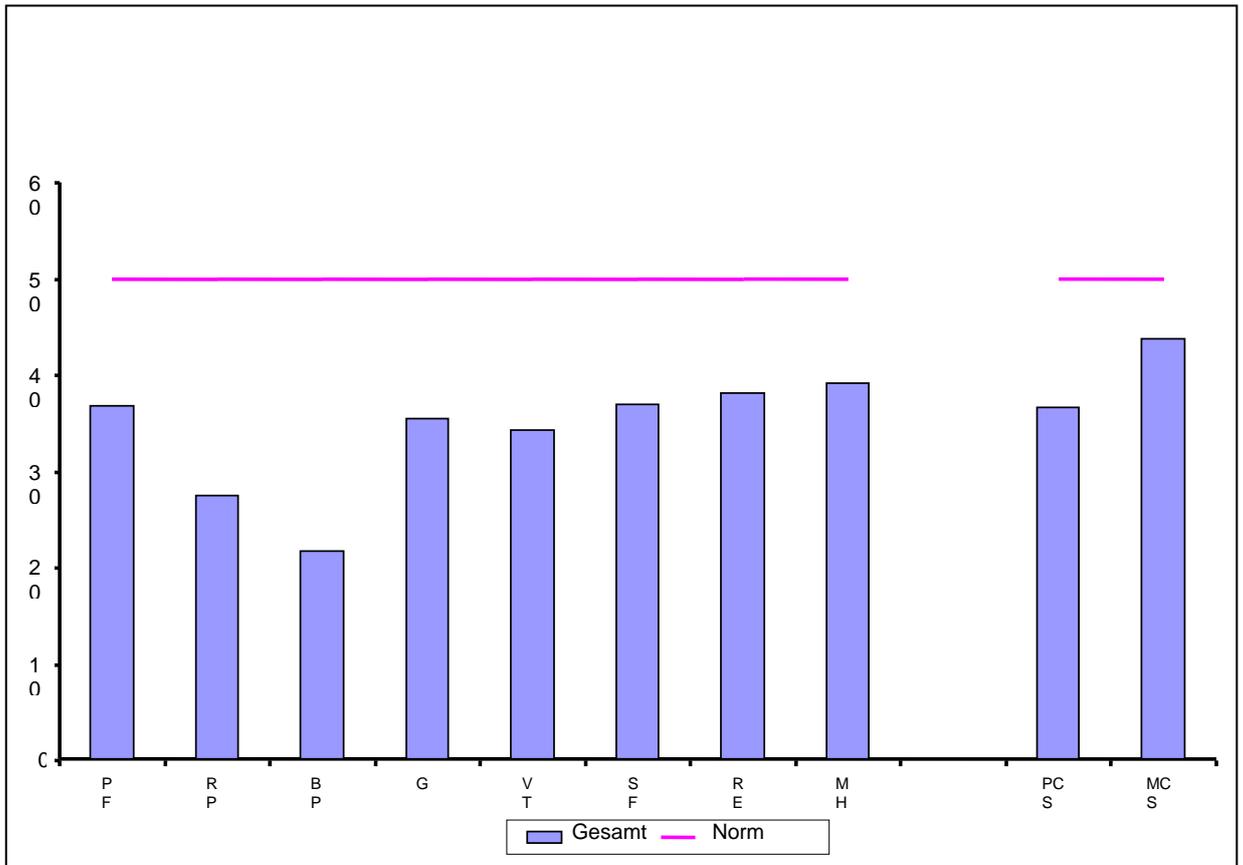


Abb. 12: Stichprobenverteilung der SF-36 Items adjustiert auf 50

4.2. Methodisch mögliche Ansätze zur Sicherstellung der Validität

Die Überprüfung der Genauigkeit wird mithilfe der Korrelation des Untersuchungsergebnisses mit einem Außenkriterium vorgenommen. Man unterscheidet drei verschiedene Arten von Validität: - die Konstruktvalidität, - die Kriteriumsvalidität und - die Inhaltsvalidität.

Konstruktvalidität liegt dann vor, wenn aus dem Konstrukt der Messung das erfasst wird, was erfasst werden soll. Werden aus einem Konstrukt Hypothesen abgeleitet, dann bedeutet eine hohe Konstruktvalidität die gute empirische Bestätigung dieser Hypothesen. Eine geringe jedoch spricht nicht zwangsläufig gegen die Messung, sondern könnte auch gegen das Konstrukt an sich sprechen. Diese Form der Validität setzt also gesichertes Wissen um das Konstrukt voraus, also Kenntnis der zugehörigen Theorien und der einschlägigen Befunde. Die Ergebnisse einer Testkonstruktion stimmen bei hoher Validität mit dem Gesamtwissen zum Konstrukt überein. Die Faktorenanalyse stellt ein besonderes Verfahren zur Bestimmung der Konstruktvalidität dar (vgl. Cook und Campbell 1976).

Die Kriteriumsvalidität ist ein spezieller Aspekt der Konstruktvalidität, bei dem die Messungen mit einer anderen konstruktvaliden Messung hoch korrelieren.

Die Inhaltsvalidität ist ein weiterer spezieller Aspekt der Konstruktvalidität. Sie liegt vor, wenn die durch die Messung erfassten Inhalte den Untersuchungsgegenstand darstellen, der gemessen werden soll. Anders gesagt bedeutet das, dass die Gültigkeit der Messung für jedermann mehr oder weniger einsichtig aus den einzelnen Teilen des Messinstruments hervorgeht. Letztlich beruht sie auf der Kenntnis von „Experten“ über den betreffenden Gegenstand. Die Behauptung ein Messinstrument habe Inhaltsvalidität, bedeutet in der Forschungspraxis aber oft nichts anderes, als dass der Entwickler des Instruments selbst glaubt, das Instrument sei valide. Gebräuchlich ist hier der Begriff der „face validity“, d.h. der „augenscheinlichen Validität“.

Dennoch ist die Idee der Inhaltsvalidität sehr wichtig: Es geht letztlich darum, dass eine Messung das relevante Phänomen möglichst in allen Aspekten erfasst, und dies kann nur durch Forschen, Nachdenken und Kommunikation zwischen Wissenschaftlern herausgefunden werden, nicht aber durch bestimmte „Techniken“ (vgl. Lienert 1998).

Im Allgemeinen ist es nicht üblich und häufig auch nicht möglich, Inhaltsvalidität numerisch zu bestimmen. Sie wird vielmehr aufgrund logischer und fachlicher Überlegungen mit oder ohne fachliche Einschränkungen akzeptiert oder verworfen (Michel/Conrad 1982).

Da wir für das Problem des „Low Back Pain“ keine Goldstandards, kein konstruktvalides Instrument vorlegen können, an dem wir die Validität der manualmedizinischen Untersuchungstechniken korrelieren und messen können, bleibt nur die „face-validity“ als einfachste Form der Validität. Die Aufgabe, die es zu lösen gilt, ist ein Expertengremium zu konstituieren. Die Personen des Gremiums mussten zweifelsohne die Untersuchungstechniken beherrschen und Erfahrungen im Umgang mit Patienten und dem „Low Back Pain“ haben.

Aus der beschriebenen Struktur der Manuellen Medizin in Deutschland liegt es nahe, ausgewählte Experten der drei Schulen der Deutschen Gesellschaft für Manuelle Therapie zur Findung einer inhaltvaliden Messung an einem Tisch zu vereinen. Diese Experten müssen in den Untersuchungstechniken einen hohen Stand an Fertigkeiten aufweisen, mehrjährige Erfahrung mitbringen und idealerweise auch Lehrkörper der jeweiligen Seminare sein. Die an der Untersuchung Teilnehmenden müssen von nationaler und internationaler Reputation sein und somit anerkannte Vertreter des Fachgebietes darstellen.

4.3. Methodisch mögliche Ansätze zur Sicherstellung der Objektivität

Die Objektivität ist durch ihre drei Unterordnungen Durchführungsobjektivität, Auswertungsobjektivität und Interpretationsobjektivität gekennzeichnet.

Durchführungsobjektivität bezeichnet den Grad, in dem mehr oder weniger zufällige Variationen im Verhalten des Testleiters und der von ihm hergestellten Durchführungsbedingungen zu Variation im Verhalten der Testperson führen. Um eine höchstmögliche Invarianz des Testleiterverhaltens zu gewährleisten, wird eine maximale Standardisierung der Testsituation angestrebt. Maßnahmen zur Optimierung der Durchführungsobjektivität sind schriftliche Vorgaben der Instruktion, ein neutrales Verhalten des Untersuchers gegenüber dem Patienten und die Notwendigkeit Interaktionen zwischen Patient und Untersucher zu minimieren.

Auswertungsobjektivität gibt das Ausmaß an, in dem numerische und kategoriale Auswertung des registrierten Testverhaltens nach ausgewählten Regeln stattfindet.

Maßnahmen zur Optimierung der Auswertungsobjektivität sind automatische bzw. computerisierte Kodierung, Schablonen zur Kodierung oder Kodiervorgaben. In der vorliegenden Untersuchung muss versucht werden, die Auswertungskategorie nominal zu skalieren. Das heißt, als Auswertung des registrierten Testverhaltens kommen nur die Entscheidungen „Ja /Nein, gestört / nicht gestört“ zur Wertung. Damit ist eine nachträgliche computerisierte Auswertung der Testergebnisse vereinfacht.

Interpretationsobjektivität ist der Grad, in dem die aus den numerischen Testergebnissen gezogenen Schlüsse unabhängig von der Person desjenigen sind, der die Interpretation vornimmt. Sie steht aber auch als Maß dafür, wie aus den gleichen Scores verschiedener Probanden identische Schlüsse gezogen werden. Maßnahmen zur Optimierung der Interpretationsobjektivität sind: Interpretationsspielräume möglichst eingrenzen, gleiche numerische oder kategoriale Kodierung gleich interpretieren und gleiche Schlüsse aus gleichen Ergebnissen ziehen.

Das Eingrenzen der Interpretationsspielräume ist jedoch bei dieser Untersuchung nicht möglich. Persönliche Erfahrung und die Komplexität der Erkrankung lassen hierbei viel Spielraum. Idealerweise dient ein Test zur Diagnosefindung. Dies ist

aber in der Medizin im Allgemeinen und bei Problemen am Bewegungssystem im Speziellen nicht möglich. Eine Vielzahl von Faktoren und nicht zuletzt auch die normale biologische Varianz machen es notwendig, mehrere Tests zum Erhärten eines Verdachtes oder zur finalen Diagnose zu verwenden. Ziel dieser Arbeit soll es daher auch sein, aus dem Spektrum der vielen Tests, die „Hochwertigen“, in Bezug auf die Testgütekriterien, herauszufiltern.

Durch ein Testmanual, welches erstellt werden muss und Bestandteil dieser Arbeit sein wird, können entsprechende Vorgaben zur Testdurchführung und -auswertung gemacht werden. Ein Test wäre vollkommen objektiv, wenn verschiedene Untersucher bei derselben Person zu gleichen Ergebnissen gelangen würden. Man spricht deshalb auch von „interpersoneller Übereinstimmung“ der Untersucher.

4.4. Methodisch mögliche Ansätze zur Sicherstellung der Reliabilität

Da das zweite Axiom der klassischen Testtheorie vorgibt, dass sich die wahre Merkmalsausprägung auch bei wiederholter Messung nicht ändert, dagegen der Fehler ausgeglichen wird, müsste also ein vollkommen reliabler Test nach wiederholter Untersuchung des selben Probanden zum gleichen Ergebnis führen. Abweichungen werden auf Messfehler zurückgeführt. Vereinfacht gesagt ist ein Test vollkommen reliabel, wenn die mit seiner Hilfe erzielten Ergebnisse den Probanden genau, d.h. fehlerfrei beschreiben bzw. auf der Testskala lokalisieren. Auf der Grundlage der Axiome 3 bis 5 sind die Messfehler vom wahren Wert unabhängig. Bei Messwiederholungen können sich somit nur unsystematische Abweichungen von den Messwerten ergeben. Diese werden als Fehlervarianz bezeichnet. Je größer die Fehlervarianz, desto mehr Messfehler beinhalten die beobachteten Werte.

Zur Bestimmung der Reliabilität wird außer der Varianz der beobachteten Werte eine Schätzung der wahren Varianz ermittelt. Diese Schätzung kann mittels verschiedener Methoden durchgeführt werden. So besteht die Möglichkeit eine Testwiederholung (Retest-Methode), einen anderen, gleichwertigen Test (Paralleltest-Methode), die Testhalbierung („split-half“-Methode) oder die Konsistenzanalyse zu wählen.

Die Vergabe paralleler Versionen eines Tests (z.B. Test A und Test B) an dieselbe Gruppe, gilt als hochwertigste Möglichkeit der Reliabilitätsbestimmung. Voraussetzung hierfür ist die Testäquivalenz, welche mit der Prüfgröße „Wilks Lambda“ ausgedrückt wird.

Für unseren Studienansatz, wie vorher beschrieben, lässt sich ein solcher äquivalenter Test zu den Einzeltests nicht finden, die Paralleltest-Methode ist somit nicht anwendbar.

Die Testhalbierung wird dann angewandt, wenn es keine Möglichkeit zur Wiederholung oder Paralleltestung gibt. Hierbei wird der vorhandene Test in zwei äquivalente Hälften aufgeteilt, z.B. durch systematische Teilung, Teilung in gerade und ungerade Items oder zufällige Teilung und nach Bearbeitung miteinander korreliert.

Die Konsistenz-Analyse ist die Verallgemeinerung des Testhalbierungsverfahrens.

Der Test wird in mehrere Teile zerlegt, im Extremfall in so viele, wie er Items hat. Es wird vorausgesetzt, dass die Teile parallel sind, d.h. sie erfassen dasselbe Merkmal und denselben „wahren“ Wert. Die Variation der Testteile wird dann als Fehlervarianz angenommen. Die daraus interpretierte Reliabilität nennt man auch innere Konsistenz.

Für die Split-half-Methode und die Konsistenzanalyse fehlen die nötigen Voraussetzungen im Rahmen der vorliegenden Studie. Die in der Validitätsphase ausgesuchten Untersuchungen können nicht in zwei Teile geteilt werden, da, bis auf wenige Ausnahmen, die einzelnen Untersuchungen einen unterschiedlichen Diagnosefokus besitzen und auch nicht in Items zerlegt werden können.

Aus den genannten Gründen ist der Studienansatz der vorliegenden Studie die Retest-Methode.

Bei dieser Reliabilitätsuntersuchung wird ein und derselbe Test in der Regel an der gleichen Testpopulation zweimal durchgeführt. Die Korrelation zwischen den ersten und zweiten Ergebnissen gibt das Ausmaß der Retest-Reliabilität an. Das Zeitintervall zwischen den beiden Tests sollte so groß gewählt werden, dass Erinnerungs- und Übungseffekte ausgeschlossen werden können, da sonst die Reliabilität überschätzt werden könnte. Gleichzeitig sollte der Abstand nicht so groß gewählt werden, dass es zu realen Schwankungen des erfassten Merkmals kommt. Dadurch erscheint die Messgenauigkeit des Instruments auf niedrigerem Niveau, als es tatsächlich der Fall ist. Die Retest-Reliabilität ist somit immer auch abhängig von der Stabilität des zu erfassenden Merkmals, weshalb auch synonym der Begriff Teststabilität für Retest-Reliabilität verwendet werden kann.

5. Empirische Untersuchungen zur manualmedizinischen Diagnostik bei low back pain

5.1. Untersuchungen zur Validität

5.1.1. Strukturierte Expertenbefragung

Die zwölf Experten der drei Ausbildungsschulen Deutschlands, die an der Studie teilnehmen (Tab.12), werden ausführlich über die Ziele der Studie informiert. Allen wird gemäß dem Schema in Kapitel 2.2.5. die Liste der Tests vorgelegt, die deutschlandweit zur Diagnostik und Klassifizierung von Low Back Pain genutzt werden. Jeder Experte muss die Einzeltests bewerten und sein Urteil über dessen Aussagekraft abgeben. Das heißt, jeder Experte muss eine Aussage über die Inhaltsvalidität der Einzeltest treffen. Dieses Urteil kann zustimmend (Test ist geeignet, um den Befund abzubilden) oder ablehnend (Test ist nicht geeignet, um den Befund abzubilden) erfolgen. In der nachfolgenden Tabelle sind die Ergebnisse dieser Expertenbefragung dargestellt. Zur Beurteilung der Wertigkeit des Befragungsergebnisses wird die nachfolgende Klassifikation der Konsensusstärke heran gezogen. (vgl. AWMF, 2001)

Klassifikation der Konsensusstärke		
Starker Konsens	Zustimmung von mehr als 95% der Teilnehmer	11 – 12 Teilnehmer
Konsens	Zustimmung von 75 – 95% der Teilnehmer	9 – 10 Teilnehmer
mehrheitliche Zustimmung	Zustimmung von 50 – 75% der Teilnehmer	6 – 8 Teilnehmer
kein Konsens	Zustimmung von unter 50% der Teilnehmer	Unter 6 Teilnehmer

Tab. 6: Klassifikation der Konsensusstärke

In den folgenden Tabellen sind die Ergebnisse der Befragung, unterteilt in Schmerztest und Funktionstests, dargestellt.

Schmerzuntersuchungen

		Test	Experten- überein- stimmung	Konsensusstärke
Lendenwirbelsäule	Orientierend	<i>Lasegue oder Straight leg raising (SLR)</i>	6	Mehrheitliche Zustimmung
		<i>Crossed SLR oder Lasegue-Moutaud- Martin-Zeichen</i>	5	Kein Konsens
		<i>Globale Federung</i>	11	Starker Konsens
		<i>Hyperextensionstest</i>	6	Mehrheitliche Zustimmung
		<i>3-Stufen- Hyperextensionstest</i>	6	Mehrheitliche Zustimmung
	Segmental	<i>Triggerpoints</i>	7	Mehrheitliche Zustimmung
		<i>Segmentale Federung</i>	11	Starker Konsens
		<i>Druckschmerz- Dornfortsatzreihe und Paravertebraler Druckschmerz</i>	6	Mehrheitliche Zustimmung
		<i>Kibler-Hautfaltentest</i>	7	Mehrheitliche Zustimmung

Tab. 7: Konsensusentscheidungen zu Schmerzuntersuchungen der Lendenwirbelsäule

		Test	Experten- überein- stimmung	Konsensusstärke
Sacroiliakalgelenke	Orientierend	<i>Patrick</i>	12	Starker Konsens
		<i>Drei-Bändertest</i>	5	Kein Konsens
		<i>Gaenslen-Test</i>	4	Kein Konsens
		<i>3-Stufen-Hyperextensionstest</i>	5	Kein Konsens
	Segmental	<i>Federungstest (Springing Test)</i>	12	Starker Konsens
		<i>Irritationspunkte</i>	9	Konsens
		<i>Kreuzgriff</i>	12	Starker Konsens

Tab. 8: Konsensusentscheidungen zu Schmerzuntersuchungen der Sacroiliakalgelenke

Funktionsuntersuchungen

		Test	Experten- überein- stimmung	Konsensusstärke
Lendenwirbelsäule	Orientieren d	<i>Finger-Boden- Abstand Test (FBA) oder Ott-Schober - Zeichen</i>	5	Kein Konsens
		<i>Seitneigung</i>	10	Konsens
	Segmental	<i>Segmentale Mobilität bei Seitneigung (im Sitzen)</i>	4	Kein Konsens
		<i>Segmentale Mobilität bei Ante-/Retroflexion</i>	4	Kein Konsens
		<i>Anteflexion in Seitlage</i>	9	Konsens
		<i>Irritationspunkt mit Rotation in Bauchlage</i>	9	Konsens

Tab. 9: Konsensusentscheidungen zu Funktionsuntersuchungen der Lendenwirbelsäule

		Test	Experten- überein- stimmung	Konsensusstärke
Sakroiliakalgelenke	Orientieren d	<i>Derbolowsky-Zeichen</i>	6	Mehrheitliche Zustimmung
		<i>Seitneigung</i>	10	Konsens
		<i>Hebetest</i>	11	Starker Konsens
	Segmental	<i>Mennell</i>	7	Mehrheitliche Zustimmung
		<i>Irritationspunkt</i>	9	Konsens
		<i>Spine-Test</i>	12	Starker Konsens
		<i>Federungstest</i>	11	Starker Konsens

Tab. 10: Konsensusentscheidungen zu Funktionsuntersuchungen der Sakroiliakalgelenke

5.1.2. Ergebnisanalyse und Diskussion zur Validität.

Für diese Studie kommen nur Tests zum Einsatz, die mindestens einen Konsens unter allen Experten erreicht haben. Tests, die keinen Konsens oder nur eine mehrheitliche Übereinstimmung erzielen, werden nicht berücksichtigt. Damit soll sichergestellt werden, dass der Validitätsaspekt gewahrt bleibt.

Als Ergebnis des Findungsprozesses für valide Untersuchungen an der Lendenwirbelsäule sind die nachfolgend beschriebenen Tests bestimmt worden.

Für die Schmerzmessung an der Lendenwirbelsäule ist als orientierende Untersuchung die „globale Federung“, als segmentale Untersuchung ist die „segmentale Federung“ hervorgegangen.

Für die orientierende Untersuchung der Lendenwirbelsäulenfunktion ist die „*Seitneige im Stehen*“ bestimmt worden, für die segmentale Untersuchung sind zwei weitere Untersuchungen von den Experten festgelegt worden. Zum einen die „*Anteflexion in Seitlage*“, zum anderen die „*Irritationspunktbestimmung mit Rotation in Bauchlage*“.

Für die Schmerzmessung an den Sakroiliakalgelenken ist als orientierende Untersuchung der „*Patrick-Test*“ hervorgegangen, für die segmentale Untersuchung der „*Kreuzgriff*“.

Als orientierende Funktionsmessung nehme ich den „*Spine-Test*“ und das „*Vorlauf-Phänomen*“ in die Studie auf, für die segmentale Funktionsmessung sind es der „*Kreuzgriff*“, „*Irritationspunktfindung an den SIG*“ und der „*Heben-Test*“.

Aus der teilweise doppelten Auswahl von Tests, einmalig auch dreifach für einen Untersuchungsbereich, ist zu schlussfolgern, dass in diesen Bereichen ein komplexeres Krankheitsgeschehen stattfindet. Dies könnte einerseits an der Komplexität der zu untersuchenden Strukturen hängen. Andererseits weist es auf eine größere Unsicherheit bei der Beurteilung nur eines Tests zum tatsächlichen Ist-Zustand in diesem Bereich hin.

Die Schmerzmessung scheint für die Lendenwirbelsäule wie für die Sakroiliakalgelenke eindeutig mit einem Test, unabhängig davon, ob orientierend oder segmental untersucht wird, durchführbar zu sein.

Für die Lendenwirbelsäule ist die orientierende Funktionsuntersuchung ebenfalls mit einem Test abbildbar. Schwieriger gestaltet sich die segmentale Zuordnung der Funktionsbestimmung, hier bedarf es zweier Tests.

Besonders komplex wirkt das Funktionsgeschehen an den Sakroiliakalgelenken. Hier haben sich die Experten auf zwei Tests für die orientierende Untersuchung geeinigt, für die segmentale müssen sogar drei Tests durchgeführt werden, um über das segmentale Funktionsgeschehen Aussagen treffen zu können.

Dies spricht meines Erachtens für die große Unsicherheit in der allgemeinen Beurteilung der SIG. Über die anatomischen und die Funktionseigenschaften der Sakroiliakalgelenke herrscht aus heutiger Sicht noch keine einheitliche Lehrmeinung. Am Beispiel der Beweglichkeit der SIG ist zu erkennen, dass die Studienlage sehr heterogen und bezeichnend für dieses Problem ist (vgl. van der Wurff et al. 2000).

Zur Fragestellung I dieser Arbeit kann gezeigt werden, dass nicht alle manualmedizinischen Untersuchungstechniken den wissenschaftlichen Kriterien an die Validität standhalten können. Der Expertenkonsensus diene in dieser Studie als durchführbare Form der Validitätsbestimmung. Er zeigt schon die notwendige Reduzierung der Techniken auf ein überprüfbares Maß. Unerwähnt darf allerdings nicht bleiben, dass es weiterer Studien und Bemühungen der Fachgesellschaft bedarf, um die Entwicklung der diagnostischen Prozeduren voran zu treiben. Hierauf werde ich im Kapitel 6 eingehen.

5.2. Untersuchung zur Objektivität

5.2.1. Standardisierung der Untersuchungsdurchführung

Die gesamte Untersuchung durch die Beobachter muss standardisiert durchgeführt werden. Insbesondere ist auf eine einheitliche Durchführung und Beurteilung der Testverfahren zu achten. Die Definition der durchgeführten Tests ist gemeinsam erarbeitet und dokumentiert worden. Zu diesem Zwecke wurde ein Trainingstreffen in Jena mit allen Studienteilnehmern durchgeführt. In diesem Treffen sind allen Techniken, die in der Studie eingesetzt wurden, gemeinsam geübt und die Durchführung diskutiert und beurteilt worden.

Die Reihenfolge der Untersuchung, das heißt in welcher Reihung die Tests zueinander gebracht werden, bedurfte weiterer Diskussion. Auch aus Gründen der objektiven Untersuchungsdurchführung in allen beteiligten Untersuchungskliniken mussten die ausgewählten Tests in einer Reihenfolge stattfinden, die in allen Zentren gleich ist.

5.2.2. Testbeschreibungen

Die Beschreibung der Untersuchungen findet nach gleichem Schema für alle statt. Die Ausgangsstellung wird beschrieben und die Testdurchführung anschließend so exakt wie nötig notiert. Hierbei wird auf die Beschreibung der Patiententätigkeit und auf die exakte Handhaltung des Untersuchers geachtet. Die nötigen Palpationspunkte werden mit Fachbegriffen oder fachüblichen Abkürzungen genannt.

Wichtiger Teil dieser Studienphase war die Vereinheitlichung der Metrik. Für alle Test wurden Bewertungsmaßstäbe gefunden, auf die schulübergreifende Lehrmeinungen wie auch seminarspezifische Aspekte Einfluss hatten. Diese Metrik wurde konsensuell von allen Teilnehmern erstellt. Zur besseren Reliabilitätsbestimmung wurde auf eine dichotome Antwortstruktur geachtet.

Testname	
Ausgangsstellung: Durchführung:	Foto der Durchführung
Metrik	

Tab. 11: Schema zur Erstellung der Testanweisungen für das Studienhandbuch

Einheitliche Beschreibungen der Techniken, die notwendige Metrik und eine fotografische Dokumentation wurden in einem Studienhandbuch zusammengefasst und allen Teilnehmern zur Verfügung gestellt.

Erarbeiten des Testprotokolls

Aus der täglichen Untersuchungspraxis und aus Gründen der Praktikabilität ergibt sich die Reihenfolge vom Stehen zum Liegen. Es wird hierbei zuerst die Seitneige geprüft, dann der Vorlauf-Test durchgeführt und schließlich der Spine-Test. Im Liegen muss erst in Rückenlage, dann in der Seitlage und dann in Bauchlage untersucht werden. Dabei wird mit dem Patrick-Test begonnen, dann die Anteflexion in Seitneige, dann die Prüfung der Irritationspunkte an der Lendenwirbelsäule, die globale und die segmentale Federung. Anschließend werden die Irritationspunkte an den SIG geprüft, dann der Kreuzgriff durchgeführt und letztlich das Heben.

Abschließend für die gesamte Untersuchung soll der diagnostizierende Arzt ein Globalurteil zur Funktionsstörung der Lendenwirbelsäulenregion L4 und L5 und zur Frage der Blockierung der SIG rechts oder links abgeben.

Die vorher beschriebenen Überlegungen zur Objektivität der Untersuchungen sind Basis für den für diese Studie entwickelten Untersuchungsbogen. Dieser zweiseitige Bogen enthält neben der Untersuchungsmatrix einen Kopfbogen, in dem die Nummer des Studienzentrums vermerkt wird, die Patientenummer und ein Feld für die Nummer des Untersuchers. Der gesamte Untersuchungsbogen ist im Anhang dargestellt.

5.3. Untersuchungen zur Reliabilität

Bedingt durch das sehr inhomogene Auftreten des Krankheitsbildes ist die zeitliche Stabilität des low back pains sehr gering. Das heißt, die Untersuchungen können an einem Patienten nicht eine Woche später wiederholt werden, da sich in der Woche die Merkmalsausprägung grundlegend geändert haben kann. Der Patient, der in der ärztlichen Praxis vorstellig wird und den Einschlusskriterien entspricht, muss mehrmals die Untersuchungsprozedur durchlaufen. Wenn ein Untersucher dies machen sollte, müssten die Zeiträume zwischen der ersten und zweiten Untersuchung relativ groß sein. Um Erinnerungseffekte zu minimieren, scheint ein Zeitraum von 3-4 Stunden angemessen. Dies ist aber in einer ärztlichen Praxis organisatorisch sehr schwer lösbar. Ein weiteres Problem besteht zudem auf der ethischen Seite, da sich der Patienten hilfeschend mit einem aktuellen Problem an den Arzt wendet. Hier kann es nicht vertreten werden, aus Studiengründen weitere 3-4 Stunden zu warten, um dann die erwartete Hilfe zu bekommen.

Das Problem muss folglich mit mehreren Untersuchern, die zeitnah hintereinander testen, gelöst werden. Die Anzahl der Untersucher ist eine weitere Hürde. Bei mehr als drei bis 4 Untersuchungen hintereinander, verschwimmen die Grenzen zwischen diagnostischer und therapeutischer Intervention. Es gibt Autoren (Patijn 2001), die davon ausgehen, dass repetitive Untersuchungen an einem Patienten, einen therapeutischen Effekt auslösen und damit die Reliabilität im Laufe der Untersuchungen schlechter wird. Die Wertung der schlechten Reliabilitätswerte sagt dann nichts über die Tests als solche, sondern mehr über methodischen Aufbau der Untersuchungsreihe aus.

Aus diesen genannten Gründen wird im Studienaufbau ein Test-Retest-Ablauf mit zwei Untersuchern gewählt.

5.3.1. Testdurchführung in den Referenzzentren

Die Struktur der Manuellen Medizin in Deutschland machte es notwendig, aus jedem manualmedizinischen Seminar zwei Referenzkliniken zu wählen, die mit ihren Untersuchern zur Verfügung stehen. Die Beteiligung wurde in Absprache mit

den Präsidenten der jeweiligen Gesellschaften getroffen, die Studienteilnahme war selbstverständlich freiwillig. Damit ist anzunehmen, dass die Motivation bei allen Untersuchern gegeben war.

Die nachfolgende Tabelle gibt die an der Studie teilnehmenden Kliniken der drei Ärzteseminare sowie die beteiligten Untersucher und die jeweilige Kodierungen an.

Manualschule	Studien-Klinik	Untersuchung s-zentrum	Untersucher	Codierung s-nummer
<u>ÄMM</u> (I)	Helmuth-Ulrici-Kliniken	1	Sturm	I/1-1
			Schröder	I/1-2
	Teufelsbad Fachklinik Blankenburg	2	Kluge	I/2-1
			Nagel	I/2-2
<u>FAC</u> (II)	Loreley-Kliniken	3	Psczolla	II/3-1
			Galeazzi	II/3-2
	Reha-Klinik f. Orthopädie	4	Wilde	II/4-1
			Wagner	II/4-2
<u>MWE</u> (III)	Argentalklinik	5	Biskupek	III/5-1
			Müller	III/5-2
	Klinik Falkenburg	6	Frey	III/6-1
			Klaas	III/6-2

Tab. 12: Teilnehmende Studienseminare, Kliniken und Untersucher

Im vorliegenden Studienfall wird die Untersucherreihenfolge randomisiert. Um systematische Verzerrungen durch die Reihenfolge der Untersucher auszuschließen, man spricht vom Bias, wird die Reihenfolge, in der der Patient von Untersucher A oder B getestet wird, nach dem Zufallsprinzip zugeordnet.

Jeder Untersucher bekommt eine Codierungsnummer, die im Testprotokoll einzutragen ist und außerdem zur Randomisierung dient. Die Randomisierung wird mit Hilfe der Internetpräsenz www.Random.org durchgeführt. Es werden alle 40 Patienten plus fünf Reservepatienten pro Studienzentrum randomisiert. Diese kompletten Randomisierungslisten liegen über den gesamten Studienzeitraum der Studienleitung in Jena vor. An die jeweilige Untersuchungsstelle wird zum Zweck der Geheimhaltung immer nur Unterlagen für fünf Patienten übermittelt. Erst wenn ein Studienzentrum die Unterlagen der vorletzten 5 Patienten übersendet, werden die Studienunterlagen inklusive der Randomisierungsliste für die nächsten fünf Patienten weggeschickt. Ein Studienzentrum hat also maximal für zehn Patienten

Studienunterlagen und die dazu gehörige Randomisierung vorliegen. Dies ist für unsere Studie die höchstmögliche Art der zufälligen Untersucherzuweisung.

In dem Referenzzentrum werden die ausgewählten Patienten dann in der festgelegten Reihenfolge, von dem jeweiligen Arzt A und B untersucht. Der erste Untersucher ist angehalten, die Untersuchungsergebnisse dem Patienten nicht zu übermitteln. Dies ist als weitere als Blindierungsmaßnahme vorgesehen, um den Patienten nicht vor der anschließenden zweiten Untersuchung zu beeinflussen. Dem zweiten Arzt werden selbstverständlich die Ergebnisse der ersten Untersuchung ebenfalls verschwiegen. Auch ein späteres Auswerten der gleichen oder gegebenenfalls unterschiedlichen Untersuchungsergebnisse wird untersagt. Diese beschriebenen Maßnahmen stellen sicher, dass keine systematische Anpassung während der gesamten Studiendauer durch die beiden Untersucher stattfinden kann. Dies wird auch in dem Treffen in Jena so kommuniziert und von den Anwesenden getragen, so dass davon auszugehen ist, dass sich alle Studienteilnehmer an diesen Ablauf halten.

5.3.2. Ergebnisdarstellung der Reliabilitätsuntersuchungen

		ohne Seitigkeit		Rechts		Links	
Test	Bewertungsdimension	Kappa	% Agreement	Kappa	% Agreement	Kappa	% Agreement
<i>Seitneige</i>	gestörte Bewegungssymmetrie	0,39	72				
	Bewegungsumfang gemindert			0,45	72	0,36	68
	gestörte Bewegungsharmonie			0,33	66	0,23	61
<i>Vorlauf</i>	Positive Blockierung			0,39	76	0,41	81
<i>Spine</i>	Positive Blockierung			0,50	81	0,39	80
<i>Patrick</i>	gestörte Beweglichkeit			0,61	81	0,64	83
	Schmerz			0,64	85	0,64	85
	pathologisches Endgefühl			0,48	75	0,61	81
<i>Anteflexion</i>	Funktions-einschränkung L4	0,34	74				
	Funktions-einschränkung L5	0,61	81				
<i>Irritationspunkt</i>	Tonuserhöhung L4			0,54	77	0,37	70
	Tonuserhöhung L5			0,41	71	0,31	66
<i>Globale Federung</i>	Schmerz	0,58	79				

Tab. 13a: Ergebnisdarstellung der Reliabilitätstest

		ohne Seitigkeit		Rechts		Links	
Test	Bewertungsdimension	Kappa	% Agreement	Kappa	% Agreement	Kappa	% Agreement
<i>Segmentale Federung</i>	Schmerz L4	0,53	77				
	Schmerz L5	0,49	75				
<i>Irritationspunkt ISG</i>	Tonuserhöhung			0,51	76	0,47	74
<i>Kreuzgriff</i>	Schmerz SIG			0,48	78	0,42	76
	Funktionseinschränkung			0,53	81	0,44	79
<i>Heben</i>	Funktionseinschränkung			0,46	79	0,39	77
<i>Globalurteil</i>	LWS Funktionsstörung	0,44	77				
	SIG Blockierung			0,50	79	0,36	73

Tab. 13b: Ergebnisdarstellung der Reliabilitätstest

5.3.3 Ergebnisanalyse der Schmerzuntersuchungen

Patrick Test

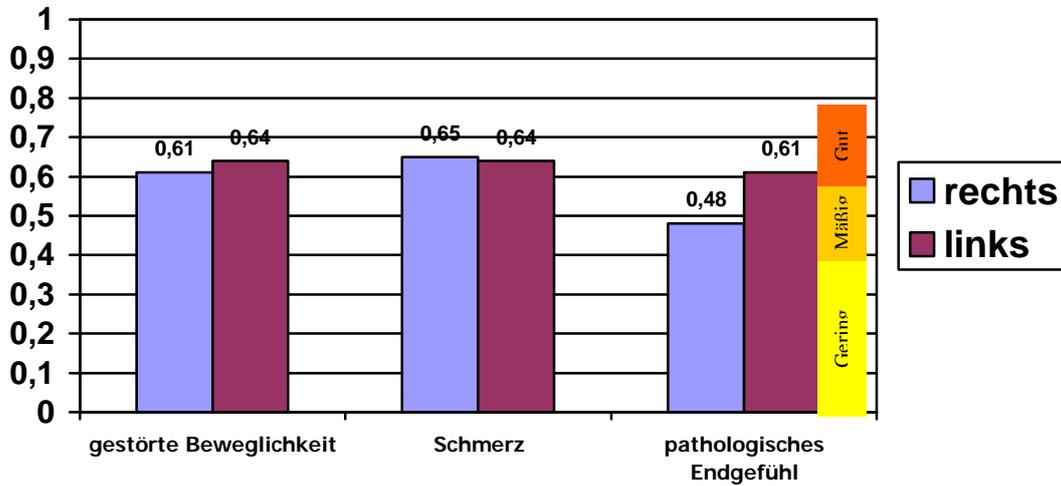


Abb. 13 : Kappawerte der Reliabilität des *Patricktests*

Analyse: Gestörte Beweglichkeit (.61 und .64) und Schmerzbefund (.65 und .64) sind auf gutem Reliabilitätsniveau. Das pathologische Endgefühl hat gutes Reliabilitätsniveau für die linke Seite (.61), rechts nur mäßiges Niveau (.48)

Kreuzgriff

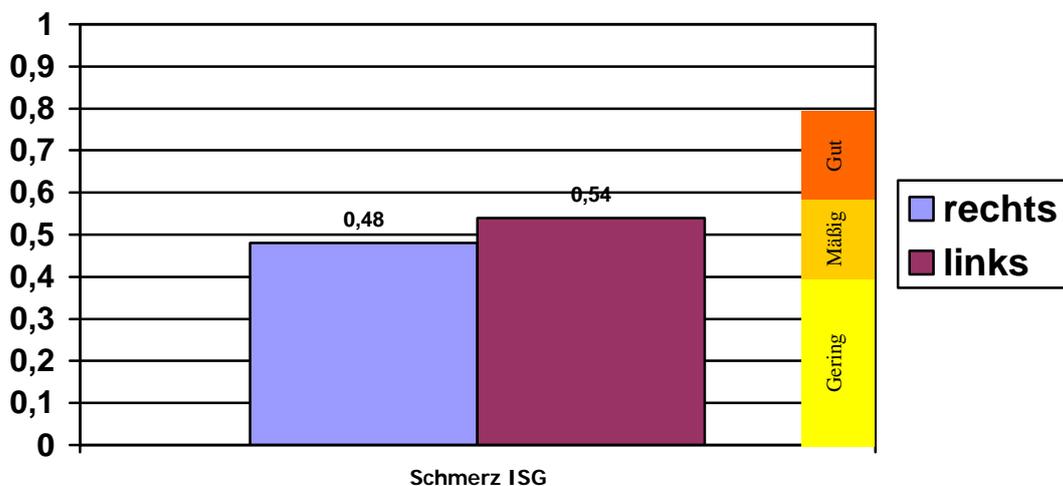


Abb. 14: Kappawerte der Reliabilität des *Kreuzgriffs*

Analyse: Die Schmerzmessung der SIG ist mit .48 und .54 auf mäßigem Niveau.

Globale und segmentale Federung

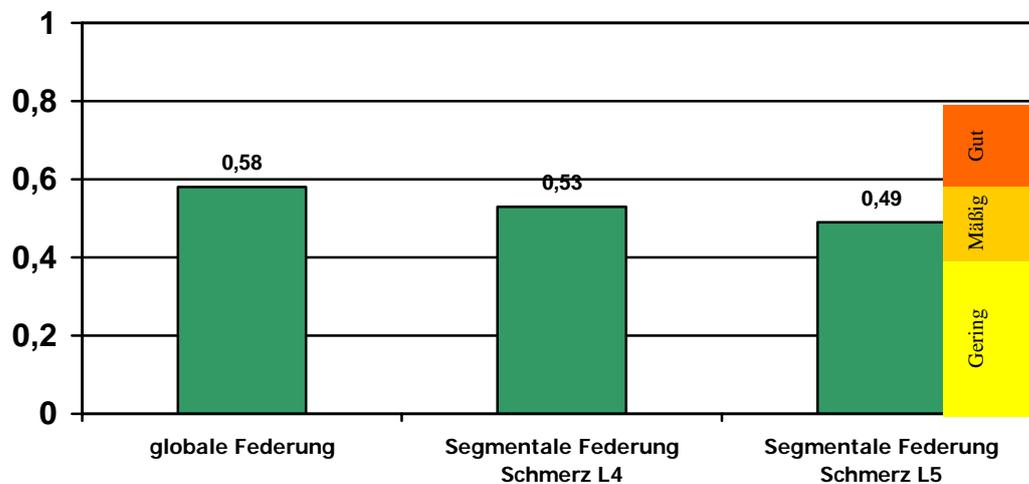


Abb. 15: Kappa-Werte der Reliabilität der *globalen und segmentalen Federung*

Analyse: Die *globale Federung* ist mit .58 nahe am guten Reliabilitätsniveau während die *segmentale Federung* für L4 (.53) und für L5 (.49) lediglich ein mäßiges Niveau aufweist.

5.3.4. Diskussion zu den Reliabilitäten der schmerzdetektierenden Tests

Die Schmerzprovokationstests in der vorliegenden Untersuchung schneiden meist auf mäßigem Reliabilitätsniveau ab. Für die drei Facetten des *Patrick-Tests* Endgefühl, Beweglichkeit und Schmerz kommt diesem die größte Bedeutung zu.

Zu gleichen Ergebnissen kommen Kokmeyer et al. (2002) in einer Untersuchung zu einem Mehrtestsystem bestehend aus vier Tests. Einer der verwendeten Untersuchungstests war der *Patrick-Test* mit einem guten Kappa-Niveau von .62 bei einem Studiendesign mit zwei Untersuchern. Dieses Niveau ist mit unserem zu vergleichen, wobei das Ziel der Studie ein anderes war. Es ging um die Sensivität und Spezifität eines Testsystems. Die Autoren bestimmten dennoch die Übereinstimmungen und daraus ergebenden Reliabilitätswerte der Einzeltests.

Laslett und Williams (1994) stellten fest, dass bei zwei gleichen Untersuchern das Reliabilitätsniveau für den *Kreuzgriff* bei .45 liegt. Auch dies ist in der vorliegenden Studie reproduziert, unsere Übereinstimmung ist etwas höher für die linke Seite, durchschnittlich aber bei beiden Seiten auf ähnlich moderatem Niveau.

Hicks (2003) stellte fest, dass Schmerzprovokationen reliabler messbar waren als segmentale Mobilität. Das Niveau der Schmerzprovokationstestkappas war aber mit 0.25 – 0.55 immer noch unter dem von Landis und Koch (1977) postulierten 0.6.

Zu der Feststellung, dass Schmerztests in Reliabilitätsstudien besser abschneiden, kommen auch Boline (1993) Strender (1997) und van der Wurff et al. (2006).

Boline et al. (1993) untersuchte acht Dimensionen der lumbalen Abnormalität auf deren Reliabilität. Als lumbale Abnormalität bezeichnet er unter anderem knöchernen Schmerz, Weichteilschmerz, visuelle Observation, OEMG-Ableitungen relevanter Muskeln und Hauttemperaturmessungen. Für alle diese Messungen oder Observationen ergibt sich kein gutes Niveau der Reliabilität. Lediglich der knöcherne Schmerz mit .49 - .90 und der Weichteilschmerz mit .40 - .79 weisen minimal einen moderaten Wert aus. Die Untersuchung von 1993 hat allerdings methodisch viele Schwächen. Die Spannweite an Kappawerten kommt durch die Analyse der einzelnen Untersuchungspaare. Insgesamt beteiligten sich 3 Untersucher bei einer Populationszahl von 28 Patienten. Des Weiteren sind in der Veröffentlichung keine Testdurchführungen beschrieben, so dass Untersuchungstechniken im Unklaren bleiben.

Strender et al. (1997) stellten nur für den Extensions- und Flexionsschmerz eine gutes Reliabilitätsniveau (.71 bzw .79) fest. In der Studie der Autoren bekommen andere verwendete Schmerztests keine gute Übereinstimmung. Dies steht nur für den *Patrickstest* und den *SIG-Compression-Pain-Test* im Gegensatz zu unseren Ergebnissen. Die Begründung liegt im methodischen Aufbau der Studie. Einerseits werden die Untersuchungen von zwei Berufsgruppen, Ärzten und Physiotherapeuten, durchgeführt. Hierbei kommen sie zu unterschiedlichen Ausprägungen der Reliabilität. Das kann wiederum an der unterschiedlichen Ausbildung liegen oder vielmehr in der Patientenzuweisung. Die untersuchten Patienten sind nicht randomisiert zugeordnet worden und die Ärztesgruppe umfasste nur 21 Patienten. In diesen Tatsachen liegt der Grund für den hohen Bias der Studie.

Die Analyse dieser Literaturstelle vor Beginn der Studie folgte eine Beschränkung der Studie auf eine Untersuchergruppe mit randomisierter Zuteilung der Patienten auf den Erstuntersucher.

Auch in einer Literaturstudie von Pescioli et al. (1997) weisen alle Provokationstests eine hohe Übereinstimmung auf, wobei einschränkend zu sagen ist, dass lediglich vier Arbeiten dazu gefunden und diskutiert wurden.

Von diesen vier Arbeiten werden in einem Fall von 1989 durch Mc. Combe und Mitarbeiter vier von fünf Tests als unzuverlässig bezeichnet. Eine weitere Arbeit von 1985 weist lediglich die Übereinstimmung in Prozent aus und ist damit für eine Interpretation unbrauchbar. Auf den Schwachpunkt der Aussage zu den Schmerzprovokationstests wurde schon in Kapitel 2.1 hingewiesen.

Robinson et al. (2007) stellen in ihrer Untersuchung auch die schmerzdedektierenden Untersuchungen mit einem Kappa von 0.43 – 0.84 als moderat bis gut dar. In deren Untersuchung beziehen sie sich lediglich auf die Schmerztests für die SIGs. Unterschiede in der Studienpopulation zu den Patienten in der vorliegenden Studie bestehen, so dass die Frage der externen Validität zu unserer Studie gestellt werden muss. Ein hoher Anteil an akuten Patienten kann zu solcher Verschiebung der Wertigkeit führen, ist aber aus unserer Studie so nicht belegbar, da dies nicht getestet wurde.

Abschließend ist zu den Schmerztests festzuhalten, dass lediglich der *Patrick-Test* für die Schmerzdetektion der SIG gute Werte für die Reliabilität bei der Messung durch zwei Untersucher aufweisen kann. *Die globale Federung* kommt knapp an die Werte und wird für die Schmerzfindung der LWS als orientierende Untersuchung empfohlen. Beide Tests dienen der globalen Beurteilung des Schmerzes. Begündet scheint es in der Komplexität des Krankheitsbildes zu sein.

Die segmentalen Tests zur Schmerzprovokation schneiden in unserer Untersuchung deutlich schlechter ab. Daraus schlussfolgernd ist zu sagen, dass Schmerzgeschehen in der untersuchten Population nur auf globaler Ebene wissenschaftlich gesichert messbar ist, die segmentalen Schmerzuntersuchungen bleiben diesen Beweis schuldig.

Eine Schmerzwahrnehmung im Bereich des unteren Rückens lässt sich polysegmental testen, monosegmental kann dies nicht selektiv getestet werden.

5.3.5. Ergebnisanalyse der Funktionsuntersuchungen

Seitneige

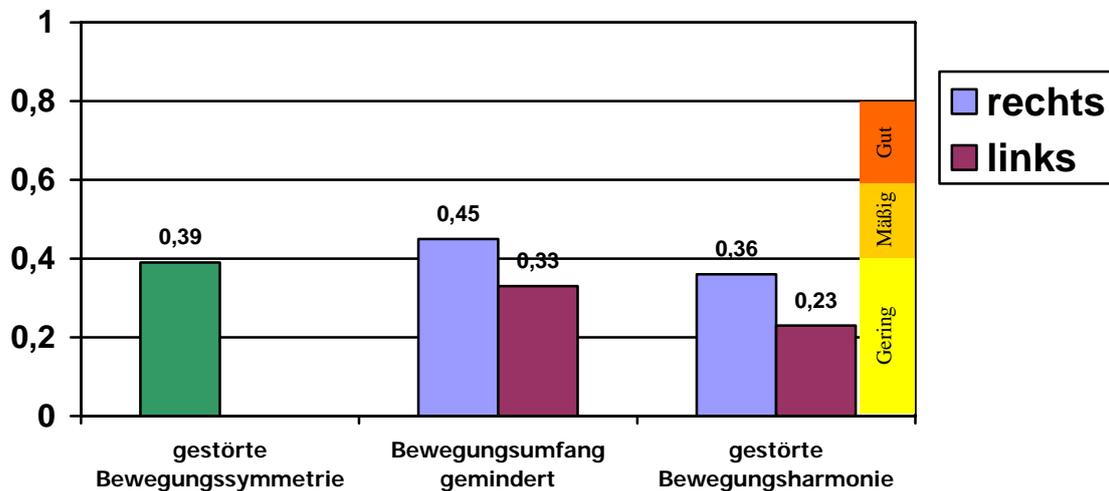


Abb. 16: Kappa-Werte der Reliabilität des *Seitneigetests*

Analyse: Der *Seitneige-Test* zeigt in allen drei Items, Bewegungssymmetrie Bewegungsumfang und Bewegungsharmonie geringes Reliabilitätsniveau. Lediglich für die rechte Seite der Bewegungsumfangsmessung konnten mäßige Kappawerte erzielt werden.

Anteflexion

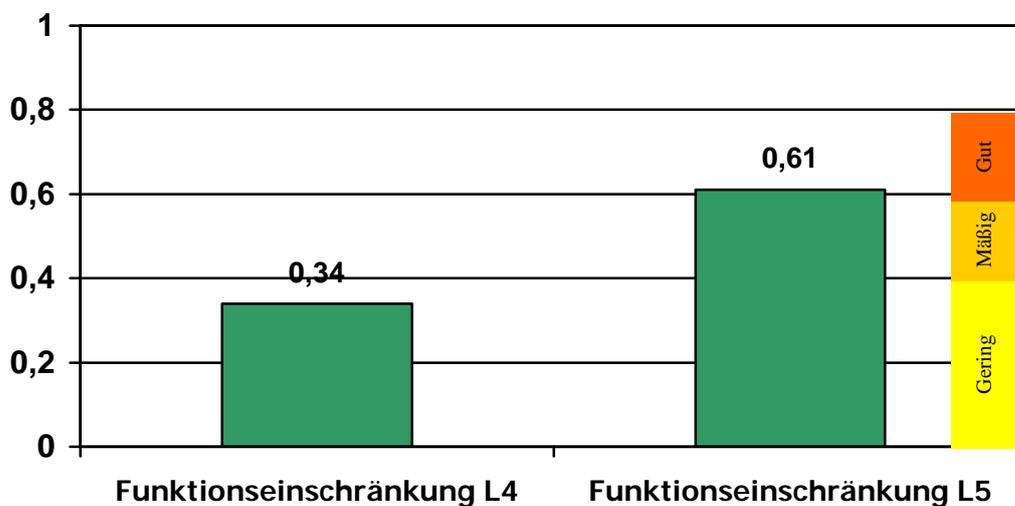


Abb. 17: Kappa-Werte der Reliabilität des *Anteflexionstests*

Analyse: Die Beurteilung der Funktionseinschränkung mittels *Anteflexion* für L4 zeigt Reliabilitäten auf geringem Niveau (.34). Für L5 sind mit diesem Test gute Reliabilitäten nachweisbar (.61).

Spine

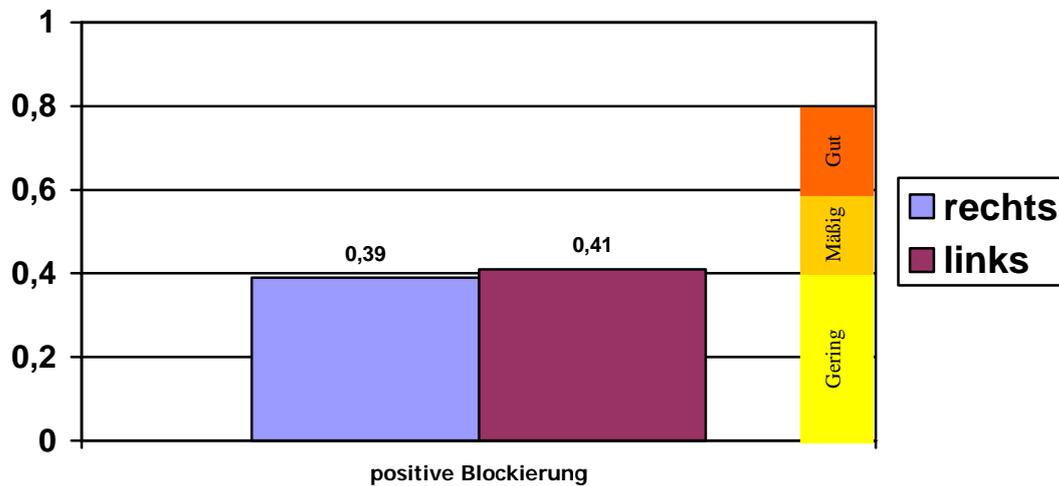


Abb. 18: Kappa-Werte der Reliabilität des *Spinetests*

Analyse: Im *Spine-Test* ist für beide Seiten nur ein geringes bis mäßiges Reliabilitätsniveau für das Auffinden einer Blockierungssituation festzustellen.

Vorlauf

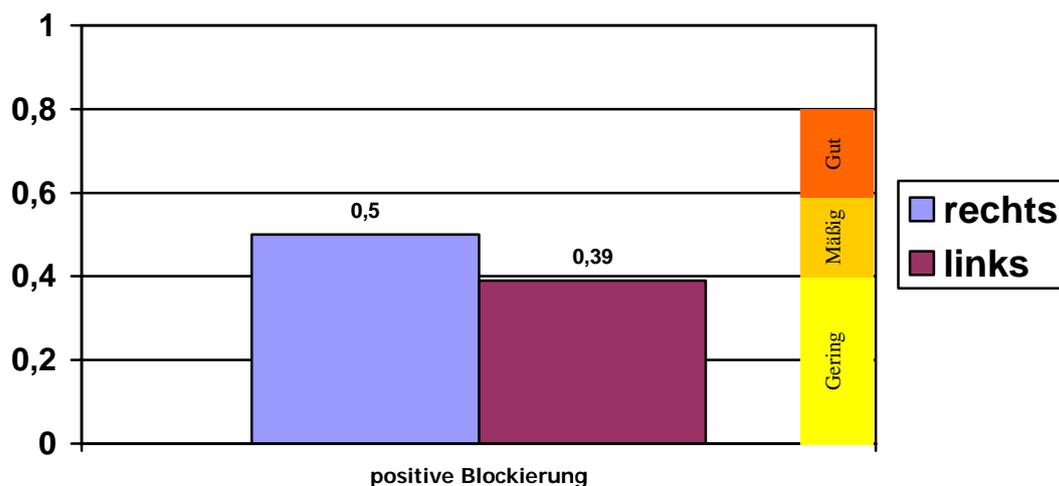


Abb. 19: Kappa-Werte der Reliabilität des *Vorlauftests*

Analyse: *Der Vorlauftest* zeigt für die rechte Seite ein mäßiges und auf der linken Seite ein geringes Reliabilitätsniveau.

Kreuzgriff

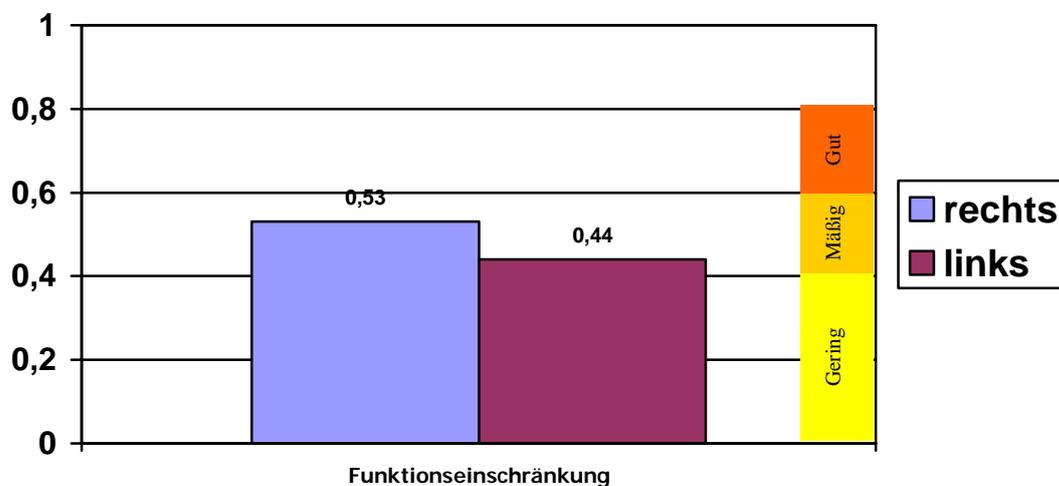


Abb. 20: Kappa-Werte der Reliabilität des *Kreuzgriff*tests

Analyse: Der *Kreuzgriff* misst die Funktionseinschränkung der SIG auf mäßigem Niveau für beide Seiten. Auch hier ist eine bessere Messung auf der rechten (.53) gegenüber der linken Seite (.54) möglich.

Heben

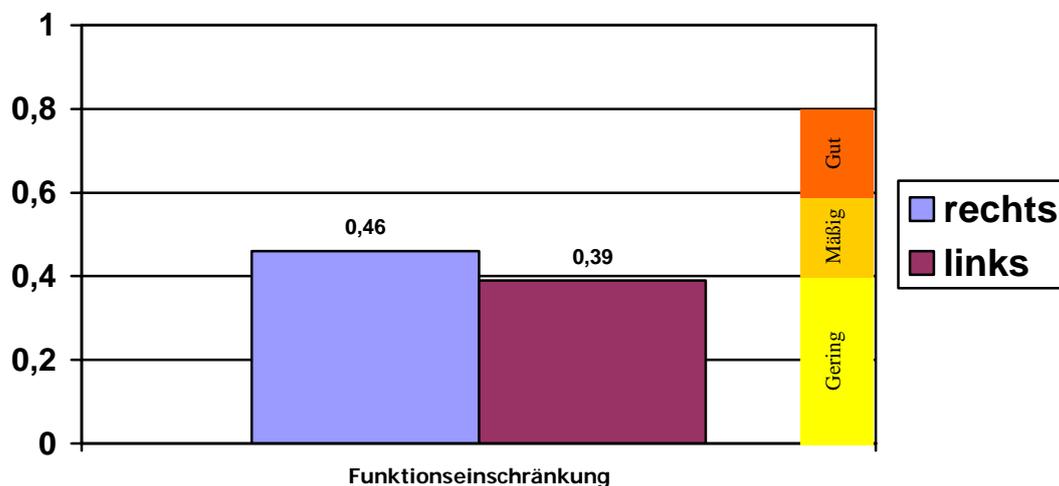


Abb. 21: Kappa-Werte der Reliabilität des *Heben-Tests*

Analyse: Die Funktionseinschränkungen des ISG sind auf geringem bis schwach mäßigem Reliabilitätsniveau messbar. Wie in den vorangegangenen Test ist auch hier die rechte Seite mit geringfügig besseren Werten.

Irritationspunkte

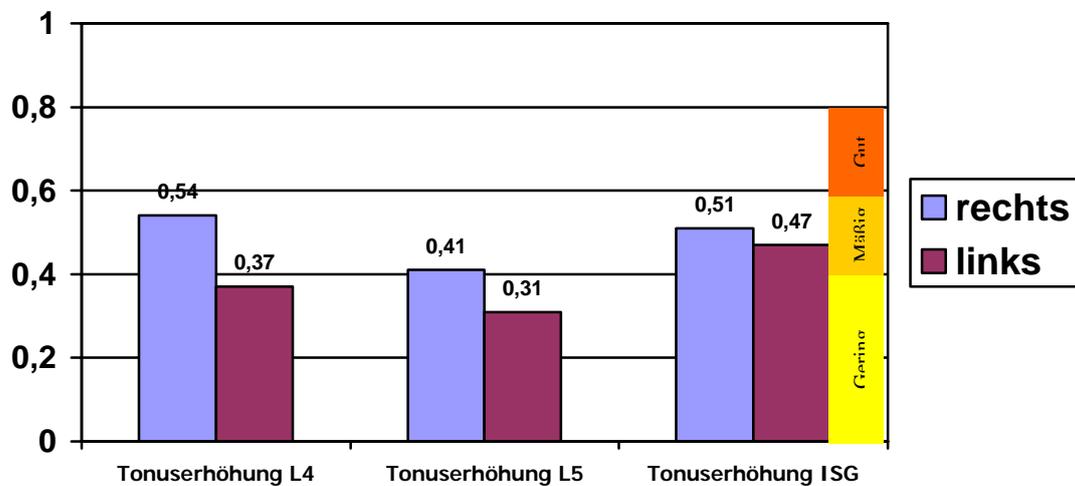


Abb. 22: Kappa-Werte der Reliabilität der Irritationspunktmessungen

Analyse: Die *Irritationspunktmessung* ist auf mäßigem (L4 und ISGs) bis geringem (L5) Reliabilitätsniveau durchführbar. Die Tonuserhöhung ist wiederum rechts (.41 - .54) und links (.31 - .47) unterschiedlich messbar, die Unterschiedsausprägung ist aber nicht so groß wie bei den vorangegangenen Tests.

5.3.6. Diskussion zu den Reliabilitäten der funktionsuntersuchenden Tests

Die in dieser Arbeit verwendeten Funktionsuntersuchungen sind alle durch ein Expertenrating als valide gekennzeichnet worden, die Objektivität wird durch die beschriebenen Maßnahmen sichergestellt. In der Reliabilitätsprüfung jedoch zeigen alle Tests nur mäßige Kappawerte. Lediglich der *Anteflexionstest* für L5 zeigt gute Werte, so dass prinzipiell nach der Wertigkeit der Funktionsuntersuchungen gefragt werden muss.

In der Literatur finden sich wenige wissenschaftlich verwertbare Studien. In einer von Vincent-Smith et al. (1999) veröffentlichten Studie zeigen die Autoren, dass die Interrater-Reliabilität des *Standing-Flexion-Tests* (*Vorlauf-Test*) nur .05 beträgt. Selbst die Intrarater-Reliabilität ist mit .46 auf dem Niveau unserer Studie. Zu ähnlichen Ergebnisse, 0.03 bis 0.11 für die Reliabilitäten, kommen Croft et al. (1998) in ihrer Arbeit.

Der *Spine-Test*, der von der gleichen Arbeitsgruppe beurteilt wurde, kommt nur auf -.16 -.10.

Mit dieser Arbeit kann belegt werden, dass es möglich ist, den *Spine-Test* sicherer zu messen.

Die dafür hohen Kappa-Werte in der vorliegenden Studie sprechen sehr für die methodische Qualität und Konsequenz in der Planung und Durchführung. Einerseits durch eine hohes Maß an Homogenität in der Patientenauswahl, andererseits durch die vielen Maßnahmen zur Steigerung der Validität und Objektivität. Die Diskussion zum deutlichen Unterschied in der Beurteilerreliabilität zwischen rechts und links wird in dieser Arbeit an späterer Stelle geführt werden. Dies stellt sich nicht als ein Problem des *Vorlauf-Tests* an sich dar, sondern zieht sich als Phänomen durch alle Untersuchungsergebnisse im Funktionstestteil.

In den drei Testitems des *Seitneige-Tests* geht es um Symmetrie, Umfang und Harmonie der Bewegung. Hier zeigt sich, dass der Bewegungsumfang als ein Item, das eher metrisch beobachtet werden kann, besser abschneidet, als die beiden anderen. Für die Beurteilung des Tests kommt der Sichtinspektion klar zum Vorteil, dass der Umfang in Radiussegmenten der Bewegung beobachtet wird, was einen Vergleich zu einem Normbild oder zur Gegenseite deutlich vereinfacht. Der Manualmediziner hat bei Testdurchführung des zu testenden Körperabschnitts nicht, wie sonst üblich, den Kontakt zum eigenen Körper. Er muss aus einer Inspektion der Bewegung heraus entscheiden, ob eine Einschränkung besteht oder nicht. Es ist anzunehmen, dass dies mit einem Inklinometer als Hilfsapparat einfacher ist. Das bleibt aber eine Vermutung, da es in diesem Rahmen nicht geprüft wurde.

Zur Beurteilung der Bewegungssymmetrie und -harmonie fließen sehr viele Informationen in die Meinungsbildung ein. Hier bildet sich ein Problem ab, da die Fülle an Information leicht zur Reizüberflutung führen kann und somit ein objektives Urteil schwer fällt. Dies ist der Erklärungsansatz, warum diese beiden Items mit nur geringer Reliabilität abschließen.

Ein weiterer Ansatz ist die Bewegungsführung. In diesem mehrsegmentalen Test ist die Ausführung durch den Patienten nicht genug standardisiert. Initialbewegungen können aus drei Abschnitten der Wirbelsäule kommen. Jede unterschiedliche terminale Bewegung führt auch zwangsläufig zur Veränderung in der Beurteilung der gesamt Bewegungsausführung.

Im *Anteflexionstest* zeigt die Beurteilung der Funktion in L4 nur geringe Reliabilitäten. Dies ist hauptsächlich damit zu begründen, dass in L4 zur Bewegungsausführung mehr Freiheitsgrade zur Verfügung stehen als im Segment darunter (Kapandji 2006).

Deshalb kann eine Funktionseinschränkung mehr oder weniger gut in dieser Ebene kompensiert werden, bzw. ist für einen Untersucher deutlicher zu sehen als für den nächsten Untersucher.

In der Beurteilung der Funktionseinschränkung in L5 ist die Beobachtungssituation grundlegend anders. Hier lässt sich das Sakrum als caudal angrenzendes Segment mit der Hand des Untersuchers fassen und feststellen. Damit fehlt für die Funktion in L5 eine Kompensationsmöglichkeit und lässt sich besser beurteilen.

Boline et al (1993) stellten für die visuelle Observation von „lumbaler Abnormalität“ eine große Spanne von Kappa-Werten der Reliabilität zwischen .34 -.84 fest. Die niedrigen Werte können bestätigt werden, für die hohe Reliabilität findet sich in der vorliegenden Studie keine Bestätigung.

Im *Heben-* und *Kreuzgriff-Test* können mäßige Reliabilitäten festgestellt werden. Mit .46 und .53 für die rechte Seite bzw. .39 und .44 liegen die Werte über den von van Deursen (1990 und 1993) mit .15 bzw .30. Dies ist wieder ein Indiz für den Erfolg der flankierenden Objektivitätsmaßnahmen, da diese in der van Deursen Studie nicht in dem Maße berücksichtigt wurden.

Den Sellschen *Irritationspunkten* liegen folgende physiologische Aspekte zu Grunde. Das Iliosakralgelenk (ISG) des Menschen hat wichtige Funktionen bei der Statik und Dynamik des Bewegungsapparates. Beim ISG-Syndrom hat Sell einen Irritationspunkt in der Regio glutea beschrieben. Es zeigte sich, dass die Ansatzportion des M. piriformis unterhalb des auf der Haut markierten Irritationspunkts im Spatium subglutaeum gelegen ist. Das ISG und der M. piriformis werden aus den Rückenmarksegmenten S1 und S2 versorgt. Beide Strukturen sind deshalb als Stationen eines Reflexbogens verknüpft. Die Befunde aus der Untersuchung von Christ et al (2001) machen es sehr wahrscheinlich, dass die unter dem M. gluteus maximus palpable Gewebsverhärtung auf eine reflektorische Verspannung des M. piriformis zurückzuführen ist.

Diese *Irritationspunkte* können als sekundäre Phänomene zum Rückenschmerz genannt werden. In der internationalen Literatur liegen keine vergleichbaren Techniken vor. Die Trigger points sind nicht identisch mit den Irritationspunkten

am SIG, so dass ein Vergleich nicht möglich ist. Auch fehlen in der deutschen Literatur wissenschaftliche Belege für die Testgütekriterien der Sellschen *Irritationspunkte*.

Das Problem der Übereinstimmung mehrerer Untersucher bei Funktionsuntersuchungen ist nicht auf den lumbalen Teil der Wirbelsäule begrenzt. Stellvertretend für viele Arbeiten sei die von Schöps et al. (2000) genannt, die die Güte der Übereinstimmung mehrerer Untersucher in ihrer methodisch guten Studie über HWS-Untersuchungstechniken mit gering bis mäßig einschätzen. Auch in dieser Studie wird schlussfolgernd bemerkt, dass die Standardisierung des Untersuchungsablaufs, Festlegung von Dokumentationsrichtlinien und häufige Trainingsphasen entscheidend zur Verbesserung des Problems beitragen könnten. Dies ist leider in der hier vorliegenden Studie trotz sorgfältiger Umsetzung nur zum Teil zu belegen gewesen.

Das Studienprojekt konnte zeigen, dass die Maßnahmen zur Objektivität und Validität einen positiven Einfluss auf die Studienergebnisse genommen haben. Die Tatsache, dass die Reliabilität nicht ausreichend vorhanden ist, bleibt trotzdem.

5.3.7. Ergebnisanalyse der *Globalbeurteilungen*

Globalbeurteilungen

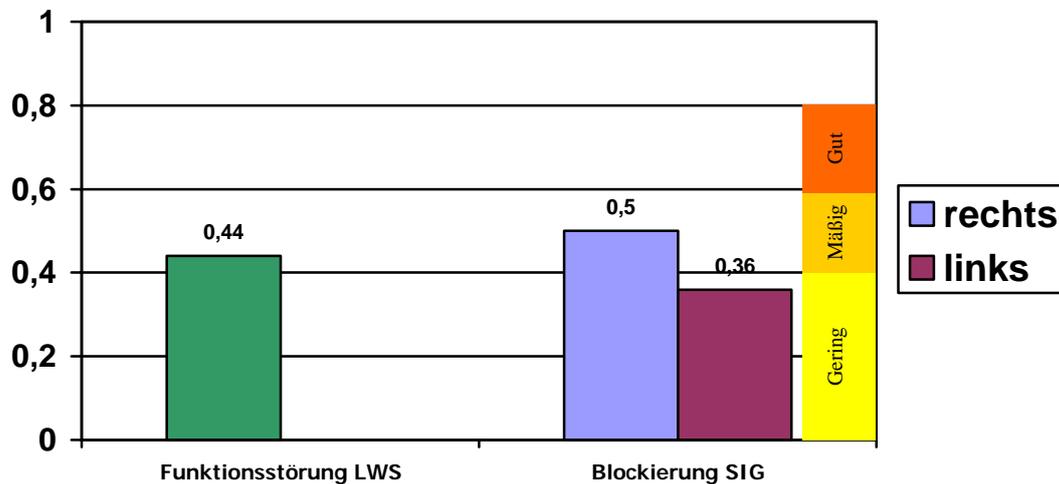


Abb. 23: Kappa-Werte der Reliabilität der *Globalbeurteilungen* zur Funktionsstörung der LWS und zur Blockierung der SIG

Analyse: Die *Globalbeurteilung* zur LWS-Funktionsstörung ist mäßig reliabel (.44). Die *Globalbeurteilung* zur SIG- Blockierung hat ebenfalls mäßiges Reliabilitätsniveau für die rechte Seite (.50) und geringes für die linke Seite (.36). Die Seitendifferenzen aus den Einzeltests spiegeln sich im Globalurteil wider.

Einflussgrößen auf das Globalurteil LWS-Funktionsstörung

Die Analyse zur Prädiktion zwischen den einzelnen Tests und dem Globalurteil „LWS-Funktionsstörung“ (als Ausdruck hierfür das „Odds ratio“, dargestellt in der Tabelle 14 im Anschluss) zeigt folgendes: Einen nachweisbar hohen Einfluss (OR >10) hat die Bewertung der Funktionseinschränkung in L5 bei Anteflexion. Einen minderen, aber immer noch deutlichen Einfluss haben die Beurteilung der Funktionseinschränkung in L4 ebenfalls in Anteflexion, die Symmetriebewertung in Seitneige und der segmentale Federungsschmerz in L5. Kein starker Zusammenhang konnte zwischen Umfang und Harmonie der Bewegung in Seitneige, den Irritationspunkten, dem Schmerz bei globaler und segmentaler Federung in L4 gefunden werden.

Globalurteil LWS –Funktionsstörung

		OR
SN gestörte Bewegungssymmetrie		11,5
SN Bewegungsumfang gemindert	re	9,8
SN Bewegungsumfang gemindert	li	6,4
SN gestörte Bewegungsharmonie	re	8,3
SN gestörte Bewegungsharmonie	li	7,7
Anteflexion Funktionseinschränkung	L4	51,7
Anteflexion Funktionseinschränkung	L5	15,1
Irritationspunkte Tonuserhöhung L4	re	8,9
Irritationspunkte Tonuserhöhung L4	li	5,7
Irritationspunkte Tonuserhöhung L5	re	5,0
Irritationspunkte Tonuserhöhung L5	li	5,6
Globale Federung Schmerz		9,6
Segmentale Federung Schmerz	L4	13,7
Segmentale Federung Schmerz	L5	7,9

Tab. 14: Analyse zum Zusammenhang zwischen dem Globalurteil LWS-Funktionsstörung und den Testergebnissen

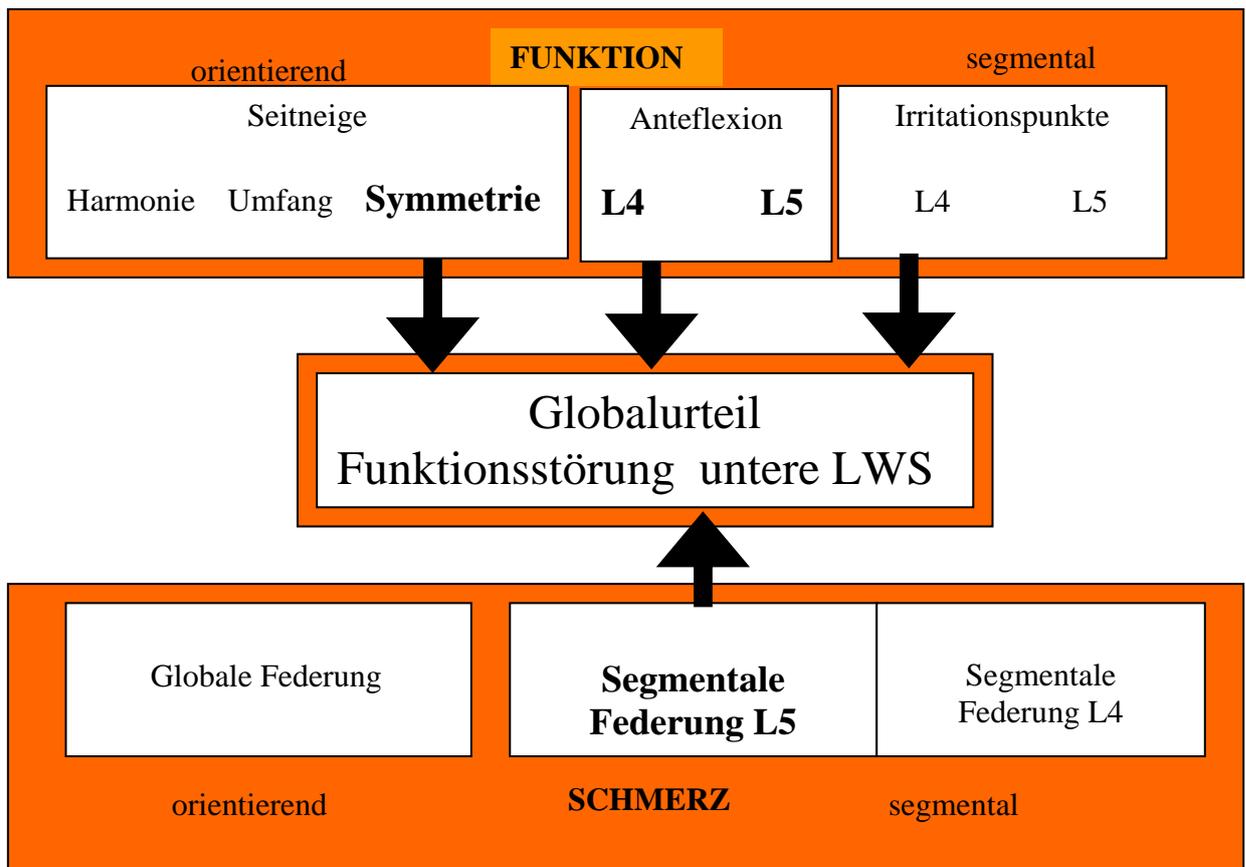


Abb. 24: Schematische Darstellung des Zusammenhangs der Einzeltests auf das Globalurteil Funktionsstörung LWS

Einflussgrößen auf das Globalurteil SIG-Blockierung

Die Analyse zur Prädiktion zwischen den einzelnen Tests und der SIG-Blockierung (Tab. 15) zeigt folgendes: Einen nachweisbar hohen Einfluss auf die Bewertung der SIG-Blockierung hat der *Hebe-Test*. Weiterhin haben die Tests *Vorlauf*, *Spine* und die Funktionsbeurteilung aus dem *Kreuzgriff* einen deutlichen Einfluss auf die Globalbeurteilung. Keinen nachweislich starken Einfluss haben die *Irritationspunkte*, der *Patrick-Test* und die Schmerzangabe aus dem *Kreuzgriff*.

Globalurteil SIG-Blockierung rechts		OR
Vorlauf Positiv (Blockierung)	re	23,0
Spine Positiv (Blockierung)	re	35,9
Patrick gestörte Beweglichkeit	re	3,4
Patrick Schmerz	re	1,8
Patrick pathologisches Endgefühl	re	4,7
Irritationspkt. ISG Tonuserhöhung	re	6,2
Kreuzgriff Schmerz ISG	re	5,4
Kreuzgriff Fkt.-Einschränkung	re	36,8
Heben Fkt.-Einschränkung	re	56,0
Globalurteil SIG-Blockierung links		OR
Vorlauf positiv (Blockierung)	li	39,5
Spine positiv (Blockierung)	li	24,5
Patrick gestörte Beweglichkeit	li	3,6
Patrick Schmerz	li	2,4
Patrick pathologisches Endgefühl	li	4,9
Irritationspkt. ISG Tonuserhöhung	li	7,1
Kreuzgriff Schmerz ISG	li	8,6
Kreuzgriff Fkt.-Einschränkung	li	58,0
Heben Fkt.-Einschränkung	li	74,3

Tab. 15: Analyse zum Zusammenhang zwischen dem Globalurteil LWS-Funktionsstörung und den Testergebnissen

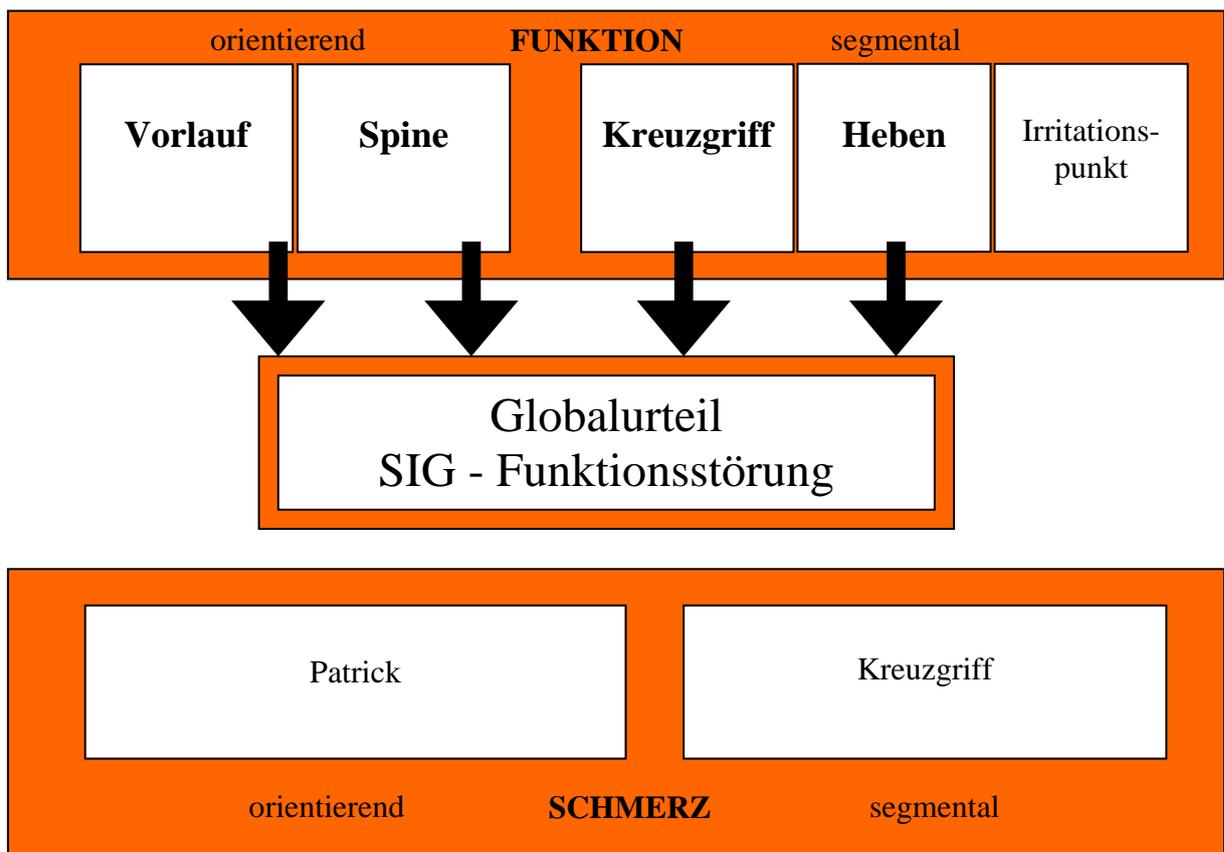


Abb. 25: Schematische Darstellung des Zusammenhangs der Einzeltests auf das Globalurteil Funktionsstörung SIG

5.3.8. Diskussion zu den Reliabilitäten der Globalurteile

Die Reliabilitäten der Globalbefunde zeigt ein extrahiertes Bild der Einzeltestbefunde. Obwohl als Einzelitem in diese Untersuchung eingegangen, spiegeln sich doch die Einzelbefunde in den Globalurteilen wieder. Das Gesamturteil ist nur Ausdruck der Heterogenität der Einzelwertungen.

Interessante Aufschlüsse bietet die Zusammenhangsanalyse der Einzeltestbefunde zum Globalurteil. Den größten Einfluss auf das Urteil hat nicht der Test mit der größten Reliabilität (*Patrick*) und damit der wissenschaftlich höchstwertige Test. Vielmehr nehmen zur Beurteilung der LWS-Funktion der Symmetriem der Seitneige, Anteflexion L4 und L5 sowie die segmentale Federung einen großen Stellenwert ein. Lediglich die Anteflexion in L5 ist aus wissenschaftlichen Gesichtspunkten heraus ein hochwertiger Test.

Bei der Beurteilung auf SIG-Blockierung werden vier Tests zur Beurteilung herangezogen. Auffällig ist das Fehlen von schmerzdedektierenden Tests. Die Schmerzangabe des Patienten scheint keinen Einfluss auf das Behandlerurteil zu haben, obwohl auch hier der Schmerzitem des „*Patrick*“ gute Werte der Übereinstimmung zwischen den Beurteilern zeigt. Auch die mäßig reliable Schmerzmessung des *Kreuzgriffs* zeigt keinen Einfluss, im Gegensatz hierzu nimmt die Funktionsmessung deutlichen Einfluss auf die Meinungsbildung.

Wie ist das zu deuten? In der internationalen Literatur fehlen Aussagen und Studien zu diesem konkreten Thema, so dass an dieser Stelle nicht adäquat diskutiert werden kann. Lediglich die Mutmaßung, dass die Globalbeurteilung als finales Statement des Arztes eine Multitest-Aussage ist, die wesentlich von persönlicher Erfahrung und Funktionsverständnis des Untersuchers abhängt, kann geäußert werden. Es gibt nicht einen Test, der als „goldener Standard“ angesehen werden kann, vielmehr gilt es, ein Multitest-Regime zur Diagnosestellung bei Low-Back-Pain zu finden, das allen Anforderungen an die Wissenschaftlichkeit gerecht wird.

6. Ausblick und Methodenkritik

Ausblick

Mit dem Vorliegen dieser Arbeit ist ein Schritt in die Richtung zum Lüften des „scientific enigma“ getan worden. Es konnte gezeigt werden, dass eine wissenschaftliche Überprüfung der allgemein üblichen manualmedizinischen Untersuchungen zur Problematik „Low Back Pain“ notwendig und möglich ist. Die Vielzahl der Tests konnte mit einer sauberen Methodik reduziert werden, ohne an validem Inhalt zu verlieren. Dies ist aus zwei Hauptgründen essenziell gewesen. Alle Testprozeduren müssen einer Validitätsprüfung unterzogen werden, damit zweifelsfrei gemessen wird, was vorgegeben wird zumessen. Dies trifft besonders für schulinterne Variationen von Tests, aber auch für selbstverständlich durchgeführte Tests zu. Beispiele hierfür sind der Stoddard-Test und das Derbolowsky-Zeichen, die im Konsensus nicht genug Stärke zeigten und aus der Untersuchung gestrichen wurden.

Ein weiterer Grund zur Reduktion der Tests war die Praktikabilität. In unserem Untersuchungssetting spielte die tägliche Routine eine entscheidende Rolle. Die Patienten wurden im Klinikalltag rekrutiert und von zwei Ärzten untersucht. Diese Untersuchung kann aus verschiedenen Gründen nicht länger als 20 Minuten dauern. Einerseits sollten keine Laborbedingungen für dieses Projekt geschaffen werden. Das heißt, die Untersuchungsbedingungen des Projekts sollten die gleichen sein, wie sie im manualmedizinischen Alltag vorkommen. Damit stellen wir die externe Validität der Ergebnisse sicher. Aus den genannten Gründen musste im Vorfeld eine Auswahl an Tests getroffen werden.

Für die Zukunft muss sichergestellt werden, dass schulinterne Variationen zunächst auf deutscher Ebene diskutiert werden, bevor sie in die Curricula Einzug halten.

Die Angleichung der curriculären Voraussetzungen zur Erlangung der Fortbildung „Manuelle Medizin“ für Ärzte muss in allen drei Studienseminaren stattfinden. Die vorliegende Untersuchung hat einen ersten Schritt in diese Richtung unternommen. Dennoch ist diese Untersuchung nicht frei von schulinterne Variationen. Die Sell'schen Irritationspunkte sind als solche zu betrachten.

Die Angleichung der Ausbildungsinhalte, das gemeinsame Erarbeiten eines Curriculums dient der Sicherstellung der wissenschaftlich erbrachten Ergebnisse dieser Arbeit. Die Weiterentwicklung der diagnostischen Prozeduren kann meines Erachtens nur über ein Netzwerk der Studienseminare der FAC, MWE und ÄMM bewerkstelligt werden. Dies gilt gerade im Hinblick auf den nächsten Schritt in der Rehabilitation von Low Back Pain, denn nach der Untersuchung der Tests muss das weitere therapeutische Vorgehen untersucht werden. Hier eignen sich besonders randomised clinical trials (RCT) zur Klärung. Der RCT gilt heute als der Goldstandard zur Erbringung der höchsten Evidenz.

Mit der vorliegenden Arbeit wurde eine wichtige Grundlage für therapeutische Interventionsstudien geschaffen.

Methodenkritik

Externe Faktoren

Diese vorliegende Arbeit muss aus mehreren Blickwinkeln kritisch betrachtet werden.

Trotz aller Anstrengungen im methodischen Bereich der Arbeit ist es zu einigen Problemen in der Durchführung gekommen. An erster Stelle ist die teilweise als sehr schwierig zu bezeichnende Zusammenarbeit mit den Untersuchungskliniken zu nennen. Auf der einen Seite gab es engagierte Kliniken mit Ärzten, die in regelmäßigen Abständen Patienten untersuchten. Deren Vorgehen war mir als Studienkoordinator transparent und nachvollziehbar. Die Maßnahmen zur Blindierung der Ergebnisse und zur Randomisation der Patienten stießen auf volle Übereinkunft und Verständnis. Diese Kliniken hatten den wissenschaftlichen Sinn der getroffenen Maßnahmen verstanden und konnten ihn voll umsetzen.

Auf der anderen Seite gab es eine Klinik, die ca. 18 Monate brauchte, um einen Bruchteil der Patienten zu untersuchen, die eigentlich vorgesehen waren (Tab.3).

Anderen wiederum machte die Randomisation Probleme. Diese Probleme konnten von mir nicht zufriedenstellend zentral gelöst werden. Hier hätte mehr Kontakt durch die Seminarleiter oder durch mich als Studienkoordinator gewährleistet werden müssen. In nachfolgenden Studien wird darauf geachtet werden müssen, dass ein regelmäßiger persönlicher Kontakt zu den Kooperationskliniken und

Ärzten stattfinden muss. Vorstellbar sind periodische Treffen mit Trainingscharakter oder vor Ort Besuche der Studienleiter zur Klärung von organisatorischen und methodischen Fragen, die sich während der einzelnen Untersuchungsphasen ergeben. Diese Maßnahmen sollen aus wissenschaftlicher Sicht die methodische Qualität der Studien hochhalten und die Untersuchungsergebnisse transparenter gestalten.

Die in dieser Studie verwendete postalische Übermittlung der Randomisationslisten kann in Zukunft über online Randomisationen stattfinden. Diese Möglichkeit ist für die Untersucher schneller und jederzeit abrufbar. Ein geeignetes Netz von Anbietern solcher Online-Angebote besteht in Deutschland und sollte genutzt werden.

Dies waren externe Faktoren der Studie, die in zukünftigen Forschungsprojekten optimiert werden können.

Interne Faktoren

Zu den internen Faktoren der Studie gehört u.a. die Zusammenstellung der Tests bei Low-Back-Pain. Die gewählte Testanordnung muss optimiert werden um Overflow-Effekte zu verhindern. Es ist bekannt, dass zu viele Einzeltests eine systematische Verzerrung in der Testbeurteilung mit sich zieht (Patijn 2000). Diesem Umstand muss ein Studiendesign folgen, das sich in Einzelabschnitten mit verschiedenen Zusammenstellungen der Einzeltests auseinander setzt um Teilfragen zu den Testgütekriterien zu klären. In der vorliegenden Studie konnte auf diesen Aspekt kein Augenmerk gelegt werden, da nicht nur die Reliabilität im Fokus des Interesses war, sondern auch die Testgütekriterien Validität und Objektivität. Es konnte gezeigt werden, dass mit den geeigneten Maßnahmen ein hohes Maß an Zuverlässigkeit in der Testauswahl erfüllt werden kann. Ohne diese Vorleistung für die genannten Gütekriterien würde einer Fortführung der Reliabilitätsuntersuchungen die notwendige Grundlage fehlen.

Trotz aller Methodenkritik glaube ich, dass die vorliegende Arbeit den Fachgesellschaften auf ihrem Weg zur wissenschaftlichen Evaluierung ihrer täglichen Arbeit geholfen hat. Sie stellt die Grundlage für weitere Forschungsvorhaben dar.

7. Zusammenfassung

Spezifische Tests bilden eine Grundlage der manualmedizinischen Diagnostik und der daraus resultierenden Therapie. Diese Tests werden vor allem hinsichtlich der formellen Testgüte unterschiedlich bewertet. Die Vielfalt von Tests, mangelnde Standardisierung und die unterschiedlichen Anwendungspräferenzen der manualmedizinischen Schulen erschweren eine qualifizierte Aussage zur diagnostischen Wertigkeit der Tests. Insgesamt sind über 36 Tests in Deutschland gebräuchlich und bekannt. Hierbei bleiben die schulinternen Weiterentwicklungen und die Abwandlungen durch einzelne Lehrgruppen ungezählt und spielen für diese Studie keine Rolle. Ohne Daten zur testtheoretischen Nachvollziehbarkeit von diagnostischen Schlüsselmethoden sind keine Aussagen zum therapeutischen Ergebnis der manualmedizinischen Therapie und deren Erfolgskriterien für eventuelle Interventionsstudien möglich.

So erscheint die Bestimmung der Testgütekriterien Validität, Reliabilität und Objektivität der verbreiteten manualmedizinischen Tests als Schlüsselproblem dieses Gebietes.

Zu diesem Zweck werden die Testgütekriterien der gebräuchlichen manualmedizinischen Untersuchungen bei unterem Kreuzschmerz oder besser „low back pain“ geprüft. Dieses Krankheitsbild ist sowohl in seiner zeitlichen Ausdehnung wie auch in der Intensitätsausprägung als sehr heterogen gekennzeichnet. Es wurde durch strenge Ein- und Ausschlusskriterien eine Studienpopulation gefunden, die repräsentativ hierfür ist.

Aus den genannten Überlegungen heraus wurden folgende Fragestellungen abgeleitet:

Welche manualmedizinischen Diagnostiken erfüllen die wissenschaftlichen Kriterien der Validität, die an einen Test gestellt werden?

Welche manualmedizinischen Diagnostiken erfüllen die wissenschaftlichen Kriterien der Reliabilität, die an einen Test gestellt werden?

Wie kann für ausgewählte diagnostische Prozeduren die Objektivität sichergestellt werden?

Bei dieser Studie handelt es sich um eine prospektive, kontrollierte, randomisierte Studie mit 151 Probanden.

Pro Studienzentrum werden gemäß der Einschlusskriterien LBP-Patienten in die Studie aufgenommen und durch beide Untersucher des jeweiligen Zentrums

unabhängig manualmedizinisch untersucht. Als Untersucher dienten die Experten der drei deutschen DGMM-Schulen, ÄMM, FAC und MWE, die in den Untersuchungskliniken tätig sind. Die Reliabilität wird mit Hilfe der Kappa-Statistik berechnet.

Die in dieser Arbeit verwendeten schmerzdetektierenden Untersuchungen sind alle durch ein Expertenrating als valide gekennzeichnet worden, die Objektivität wird durch die beschriebenen Maßnahmen sichergestellt. Die Resultate zeigen heterogen verteilt Kappa-Werte. Die Schmerzprovokationstests in der vorliegenden Untersuchung schneiden meist auf mäßigem Reliabilitätsniveau ab. Für die drei Facetten des Patrick-Tests Endgefühl, Beweglichkeit und Schmerz (.61-.65) kommt diesem die größte Bedeutung zu. Des Weiteren schnitten die Test wie folgt ab: Kreuzgriff .48 und .54, globale Federung .58, segmentale Federung L4 .53 und segmentale Federung L5 .49.

Die segmentalen Tests zur Schmerzprovokation schneiden in unserer Untersuchung deutlich schlechter ab. Daraus schlussfolgernd ist zu sagen, dass Schmerzgeschehen in der untersuchten Population nur auf globaler Ebene wissenschaftlich gesichert messbar ist, die segmentalen Schmerzuntersuchungen bleiben diesen Beweis schuldig.

Eine Schmerzwahrnehmung im Bereich des unteren Rückens lässt sich polysegmental testen, monosegmental kann dies nicht selektiv getestet werden.

Die in dieser Arbeit verwendeten Funktionsuntersuchungen sind ebenfalls alle durch ein Expertenrating als valide gekennzeichnet worden, die Objektivität wird auch hier durch die beschriebenen Maßnahmen sichergestellt. In der Reliabilitätsprüfung jedoch zeigen alle Tests nur mäßige Kappawerte. Lediglich der Anteflexionstest für L5 zeigt gute Werte, so dass prinzipiell nach der Wertigkeit der Funktionsuntersuchungen gefragt werden muss. Im Einzelnen schnitten die Tests wie folgt ab: Seitneige-Test .23 - .45, Anteflexion L4 .34, Anteflexion L5 .61, Spine-Test .39 und .41, Vorlauf-Test .39 und .50, Kreuzgriff .44 und .53, Heben-Test .39 und .46, Irritationspunkte.31 - .54

Die Standardisierung des Untersuchungsablaufs, Festlegung von Dokumentationsrichtlinien und häufige Trainingsphasen können entscheidend zur Verbesserung des Problems der Reliabilität beitragen. Dies ist leider in der hier vorliegenden Studie trotz sorgfältiger Umsetzung nur zum Teil zu belegen gewesen. Es konnte gezeigt werden, dass die Maßnahmen zur Objektivität und

Validität einen positiven Einfluss auf die Studienergebnisse genommen haben. Die Tatsache, dass die Reliabilität nicht ausreichend vorhanden ist, bleibt trotzdem.

8. Verzeichnisse

8.1 Abbildungen und Tabellen

Abb. 1: Rückenschmerzhäufigkeit unterteilt nach Geschlecht und Altersklassen

Quelle: Internet - <http://www.gbebund.de/gbe10/ergebnisse>

Abb. 2: Globale Federung - Durchführung und Metrik

Abb. 3: Übertragender Schmerz (Referred Pain) bei Palpation eines aktiven

Trigger-Points (modifiziert nach Travell und Simons) aus Pongratz 2003

Abb. 4: Segmentale Federung – Durchführung und Metrik

Abb. 5: Seitneige – Durchführung und Metrik

Abb. 6: Anteflexion – Durchführung und Metrik

Abb. 7: Irritationspunkt in Bauchlage – Durchführung und Metrik

Abb. 8: Patrick-Test – Durchführung und Metrik

Abb. 9: Kreuzgriff – Durchführung und Metrik

Abb. 10: Hebetest – Durchführung und Metrik

Abb. 11: Spine-Test – Durchführung und Metrik

Abb. 12: Stichprobenverteilung der SF-36 Items adjustiert auf 50

Abb. 13 : Kappawerte der Reliabilität des Patricktests

Abb. 14: Kappawerte der Reliabilität des Kreuzgriffs

Abb. 15: Kappa-Werte der Reliabilität der globalen und segmentalen Federung

Abb. 16: Kappa-Werte der Reliabilität des Seitneigetests

Abb. 17: Kappa-Werte der Reliabilität des Anteflexionstests

Abb. 18: Kappa-Werte der Reliabilität des Spinetests

Abb. 19: Kappa-Werte der Reliabilität des Vorlauftests

Abb. 20: Kappa-Werte der Reliabilität des Kreuzgrifftests

Abb. 21: Kappa-Werte der Reliabilität des Heben-Tests

Abb. 22: Kappa-Werte der Reliabilität der Irritationspunktmessungen

Abb. 23: Kappa-Werte der Reliabilität der Globalbeurteilungen zur
Funktionsstörung der LWS und zur Blockierung der SIG

Abb. 24: Schematische Darstellung des Zusammenhangs der Einzeltests auf das
Globalurteil „Funktionsstörung LWS“

Abb. 25: Schematische Darstellung des Zusammenhangs der Einzeltests auf das
Globalurteil „Funktionsstörung SIG“

-
- Tab. 1: Betrachtungsschema zur Schmerz- und Funktionsuntersuchung der LWS und SIG
- Tab. 2: Einschlusskriterien
- Tab. 3: Verteilung der Studienpopulation auf die Untersuchungszentren
- Tab. 4: Charakterisierung der Studienpopulation nach Größe, Gewicht, BMI und Alter
- Tab. 5: Geschlechter getrennte Norm- (Ellert/Bellach1999) und Stichprobenverteilung
- Tab. 6: Klassifikation der Konsensusstärke
- Tab. 7: Konsensusentscheidungen zur Untersuchungen der Lendenwirbelsäule
- Tab. 8: Konsensusentscheidungen zu Schmerzuntersuchungen der Sakroiliakalgelenke
- Tab. 9: Konsensusentscheidungen zu Funktionsuntersuchungen der Lendenwirbelsäule
- Tab. 10: Konsensusentscheidungen zu Funktionsuntersuchungen der Sakroiliakalgelenke
- Tab. 11: Schema zur Erstellung der Testanweisungen für das Studienhandbuch
- Tab. 12: Teilnehmende Studienseminare, Kliniken und Untersucher
- Tab. 13a: Ergebnisdarstellung der Reliabilitätstest
- Tab. 13b: Ergebnisdarstellung der Reliabilitätstest
- Tab. 14: Analyse zum Zusammenhang zwischen dem Globalurteil LWS-Funktionsstörung und den Testergebnissen
- Tab. 15: Analyse zum Zusammenhang zwischen dem Globalurteil LWS-Funktionsstörung und den Testergebnissen

8.2 Literaturverzeichnis

Aure OF, Hoel Nilsen J ; Vasseljen O: Manual Therapy and Exercise Therapy in Patients With Chronic Low Back Pain: A Randomized, Controlled Trial With 1-Year Follow-up. *Spine* 28, 525-531, 2003.

Airaksinen O , Brox JI, Cedraschi C: European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J* 15(Suppl.2) :192 – 300, 2006.

Alexiev AR: Some differences of the electromyographic erector spinae activity between normal subjects and low back pain patients during the generation of isometric trunk torque. *Electromyography and clinical neurophysiology* 34, 495-499, 1994.

Aspegren DD, Cox JM, Trier KK: Short leg correction: A clinical trail of radiographic vs. non-radiographic procedures. *J Manipulativ Physiol Ther* 10, 232-238, 1987.

Assendelft WJ, Morton SC, Yu EI, Suttorp MJ, Shekelle PG. Spinal manipulative therapy for low back pain. A meta-analysis of effectiveness relative to other therapies. *Ann Intern Med* 138, 871-881, 2003.

Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). Das Leitlinienmanual: Entwicklung und Implimentierung von Leitlinien in der Medizin. *Z Ärztl. Fortbild Qualitätssich* 95 (Suppl. 1), 2001.

Bigos S, Bowyer O, Braen G: Acute low back problems in adults. Rockville, Md.: U.S. Dept. of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research, *AHCPR publication no. 95-0642*, 1994.

Binkley J, Finch E, Hall J, Black T, and Gowland C: Diagnostic classification of patients with low back pain. Report on a survey of physical therapy experts. *Phys Ther* 73, 138-150, 1993.

Bischoff HP, Moll H: Kurz gefasstes Lehrbuch der Manuellen Medizin: Chirodiagnostische und chirotherapeutische Technik. Balingen, Spitta, 2007.

Brügger A: Lehrbuch der funktionellen Störungen des Bewegungssystems. Brügger Verlag, Zolingen und Benglen, 2001.

Brüggenjürgen B: Lebensqualität und volkswirtschaftliche Kosten der Migräne in Deutschland. Asgard-Verlag, Sankt Augustin, 1994.

Buckup K: Klinische Tests an Knochen, Gelenken und Muskeln. 3.Aufl., Stuttgart, Thieme, 2005.

Bullinger M: Erfassung der gesundheitsbezogenen Lebensqualität mit dem SF-36 Health Survey. *Bundesgesundheitsblatt* 43, 190-197, 2000.

Bogduk, N: The anatomical basis for spinal pain syndromes. *J Manipulativ Physiol Ther* 18, 603-605, 1995.

Boline PD, Haas M and Meyer JJ: Interexaminer reliability of eight evaluative dimensions of lumbar segmental abnormality: Part II. *J Manipulativ Physiol Ther* 16, 363-374, 1993.

Bortz J, Döring N: Forschungsmethoden und Evaluation: Für Human und Sozialwissenschaftler. 3. Auflage. Berlin und Heidelberg. Springer 2002.

Cherkin DC, Sherman KJ, Deyo RA, Shekelle PG. A review of the evidence for the effectiveness, safety, and cost of acupuncture, massage therapy, and spinal manipulation for back pain. *Ann Int Med* 138, 898-906, 2003.

Christ B, Günther J, Frölich E, Huang R und Flöel H: Morphologische Grundlage des Sellschen Irritationspunktes für das Iliosakralgelenk. *Man Med* 39, 241-245, 2001.

Cibulka MT, Delitto A, Koldehoff RM: Changes in innominate tilt after manipulation of the sacroiliac joint in patients with low backpain. An experimental study. *Phys Ther* 68, 1359-1363, 1988.

Cook, TD, Campbell DT: The Design and Conduct of Quasi-Experiments and True Experiments in Field Settings, in Dunnette MD (Hrsg.): Handbook of Industrial and Organizational Research, 223-326, Chicago, McNally College Publishing 1976.

Descarreaux M, Blouin J, Drolet M, Papadimitriou S and Teasdale N: Efficacy of Preventive Spinal Manipulation for Chronic Low-Back Pain and Related Disabilities: A Preliminary Study, *J Man Phys Ther* 27, 509-514, 2004.

Donahue MS, Riddle DL, and Sullivan MS :Intertester reliability of a modified version of McKenzie´s lateral shift assessments obtained on patients with low back pain. *Phys Ther* 76, 706-716, 1996.

Donelson R, Aprill C, Medcalf R: A prospective study of centralization of lumbar and referred pain. A predictor of symptomatic discs an anular competence. *Spine* 22, 1115-1122, 1997.

Ehrlich GE, Khaltaev NG. Low back pain initiative. Geneva, World Health Organization; 1999.

Ellert U, Bellach BM: Der SF-36 im Bundes-Gesundheitssurvey - Beschreibung einer aktuellen Normstichprobe. *Das Gesundheitswesen*, Sonderheft 61, 1999.

FIMM Glossar 2002 : abzurufen auf der Internetseite der Internationalen Gesellschaft für Manuelle Medizin unter „http://www.fimm-online.com/pub/de/data/objects/fimm_glossary_v4_1.pdf“

French SD, Green S, and Forbes A: Reliability of chiropractic methods commonly used to detect manipulable lesions in patients with chronic low back pain. *J Manipulativ Physiol Ther* 23, 231-238, 2000.

Fritz JM, Delitto A, Vignovic M, and Busse RG: Interrater reliability of judgements of the centralization phenomenon and status change during movement testing in patients with low back pain. *Arch Phys Med Rehab* 81, 57-61, 2000.

Fritz JM, George S: The use of a classification approach to identify subgroups of patients with acute low back pain. Interrater reliability and short-term treatment outcomes. *Spine* 25, 106-114, 2000.

Gerbershagen HU: Quality of life research in pain patients. In I. Guggenmoos-Holzmann, L. Bloomfield, H. Brenner and U. Flick (Eds.): *Quality of life and health* Vol. 1, 107-124. Berlin und Wien, Blackwell, 1995.

Gesundheitsberichterstattung des Bundes 2006: abzurufen unter der Internetseite „<http://www.gbe-bund.de>“

Göbel H: Epidemiologie und Kosten chronischer Schmerzen. Spezifische und unspezifische Rückenschmerzen. *Schmerz*. 15, 92-98, 2001.

Gröbli, C., Dejung, B. Nichtmedikamentöse Therapie myofascialer Schmerzen, *Schmerz* 17, 475-480, 2003.

Gröbli, C., Dommerhold, J.: Myofasciale Triggerpunkte, Pathologie und Behandlungsmöglichkeiten, *Man Med* 35 (6), 295-303, 1997.

Guggenmoos-Holzmann: Modelling covariate effects in observer agreement studies: the case of nominal scale agreement. *Stat med* 14 (20), 2285-2288, 1995.

Guzmán J, Esmail R, Karjalainen K, Malivaara A, Irvin E, and Bombardier C: Multidisciplinary Bio-Psycho-Social Rehabilitation for Chronic Low Back Pain (Cochran Review). *Cochran Library* 2, 2002.

Harvey D, Byfield D: Preliminary studies with a mechanical model for the evaluation of spinal motion palpation. *Clin Biomech* 6, 79-82, 1991.

Hestboeck L and Leboeuf-Yde C: Are chiropractic tests for the lumbo-pelvic spine reliable and valid? *J Manipulativ Physiol Ther* 23, 258-274, 2000.

Hoogen, HJM, van den Koes BW, Devillé W, Eijk JTM, van Bouter LM: The inter-observer reproducibility of Lasègue's sign in patients with low back pain in general practice. *The British Journal of General practice* 46, 727-730 1996.

Hunt DG, Zuberbier OA, Kozlowski AJ, Robinson J, Berkowitz, J, Schultz IZ, Milner RA, Crook JM and Turk DC: Reliability of the lumbar Flexion, lumbar extension and passiv straight leg raise test in normal populations embedded within a complete physical examination. *Spine* 26, 2714-2718, 2001.

Hicks GE, Fritz JM, Delitto A, Mishock J: Interrater reliability of clinical examination measures for identification of lumbar segmental instability. *Arch Phys Med Rehabil*, 84, 1858–64, 2003.

Jacob T, Baras M, Epstein L: Low Back Pain: Reliability of a set of pain measurement tools. *Arch Phys Med Rehab* 82, 735-742, 2001.

Jensen KJ, Gemmell H, Thiel H: Motion palpation accuracy with a mechanical spinal model. *Eur J Chiropractic* 41, 67-73, 1993.

Kayser R, Moll H, Harke G: Sakroiliakalgelenk. Diagnostik und Behandlung einer sakroiliakalen Dysfunktion – Technikvorstellung. *Man Med* 46, 73–76, 2008.

Kapandji IA: Funktionelle Anatomie der Gelenke. Schematisierte und kommentierte Zeichnungen zur menschlichen Biomechanik: Untere Extremität - Rumpf und Wirbelsäule. 4 Aufl. Stuttgart, Thieme, 2006.

Kirkaldy-Willis WH, Hill RJ.: A more precise diagnosis for low-back pain. *Spine* 4(2):102-9, 1979.

Koes, BW.; van Tulder, M.; Ostelo, R.; Burton A and Waddell,G.: Clinical Guidelines fro the management of Low Back Pain in Primary Care: An International Comparison *Spine*. 26 (22), 2504-2513, 2001.

Kokmeyer DJ, Van-der-Wurff P, Aufdemkampe, G, Fickenscher, TCM:The reliability of multitestregimes with sacroiliac pain provocation test *J Manipulativ Physiol Ther* 25, 42-48, 2002.

Laslett M, Williams M: The reliability of selected pain provocation tests for sacroiliac joint pathology. *Spine* 19,1243-1249, 1994.

Landis JR and KochGG: The measurement of observer agreement for categorical data. *Biometrics* 33, 159-174, 1977.

Leboeuf-Yde C, Gardner V, Carter AL, and Scott TA: Chiropractic examination procedures: A reliability and consistency study. *J Austr Chiropr Assoc* 19, 101-104, 1989.

Leboeuf-Yde C: The sensitivity and specificity of seven lumbo-pelvic orthopaedic tests and the arm-fossa test. . *J Manipulativ Physiol Ther* 13, 138-143, 1990.

Leboeuf-Yde C, van Dijk J, Franz C, Hustad SA, Olsen D, Pihl T, Robech R, Vendrup SS, Bendix T and Kyvik KO: Motion palpation findings and self-reported low back pain in a population-based study sample. *J Manipulativ Physiol Ther* 23, 160-167, 2002.

Lienert GA., Raatz U: Testaufbau und Testanalyse.;6. Aufl.,BeltzPVU, 1998.

Mc.Combe PF: Reproducibility of physical signs in low back pain. *Spine* 14, 908-918, 1989.

Meijne W, van Neerbos, Aufdemkampe G, van-der-Wurff P: Intraexaminer and interexaminer reliability of the Gillet-test. *J Manipulativ Physiol Ther* 22, 4-9, 1999 1999.

Mens JMA, Vleeming A, Snijders CJ, Koes BW and Stam HJ: Validity of the active straight leg raise test for measuring disease severity in patients with posterior pelvic pain after pregnancy. *Spine* 27, 196-200, 2002.

Michel L, Conrad W: Theoretische Grundlagen psychometrischer Tests. In Karl-Josef Groffmann & Lothar Michel, Psychologische Diagnostik. Grundlagen psychologischer Diagnostik (1-129). Göttingen, Hogrefe, 1982.

Nachemson AL, Jonsson E (Hrsg.): Neck and back pain: the scientific evidence of causes, diagnosis and treatment. Swedish Council of Technology Assessment in Health Care (SBU). Philadelphia, Lippincott, Williams & Willkins 2000.

Neumann HD: Manuelle Medizin. Eine Einführung in Theorie, Diagnostik und Therapie. 3. Auflage, Springer, Berlin und Heidelberg, 1998.

Papageorgiou A, Croft P, Thomas E, Ferry S, Jayson M, Silman A. Influence of previous pain experience on the episodic incidence of low back pain. Results from the South Manchester back pain study. *Pain* 66,181- 186, 1996.

Patijn J: Low Back Pain- Reproducibility of diagnostic procedures in manual/musculoskeletal medicine. *J orth Medicine* 23, 36-42, 2001.

Pescioli A, Kool, J: Die Zuverlässigkeit klinischer Iliosakralgelenktests- eine Literaturstudie. *Manuelle Therapie* 1, 3-10,1997.

Pongratz D, Zierz S: Neuromuskuläre Erkrankungen: Diagnostik, interdisziplinäre Therapie und Selbsthilfe Köln, Deutscher Ärzteverlag, 2003.

Raspe H, Kohlmann T: Measuring patient related outcome criteria in medical rehabilitation: how well do "indirect" and "direct" methods of measuring change agree? *Rehabilitation* 37, 30-37 1998.

Riddle DL: Classification and low back pain: a review of the literature and critical analysis of selected systems. *Phys Ther* 78 (7), 708-737, 1998.

Robinson H, Brox J, Robinson R, Bjelland E, Solem S and Telje T: The reliability of selected motion- and pain provocation tests for the sacroiliac joint. *Manual Therapy*, 12, 72-79, 2007.

Rhodes DW, Mansfield ER, Bishop PA, Smith JF : The validity of the prone leg check as an estimate of standing leg length inequality measured by x-ray. . *J Manipulativ Physiol Ther* 18, 343-346, 1995.

Sackett DL, Richardson WS, Rosenberg W, Haynes RB: Evidence based medicine: how to practice and teach EBM. New York, Churchill Livingstone, 1997.

Sauer PM, Pflingsten M, Ensink FB, Heinemann R, Koch D, Seeger D, und Hildebrandt J : Interrater-Untersuchung zur Reliabilitätsprüfung somatischer Befunde. *Rehabilitation* 35, 150-160, 1996.

Scherfer E: Forschung verstehen. Ein Grundkurs in Evidenzbasierter Praxis. München, Pflaum, 2006.

Schöps P, Pflingsten M, Siebert U: Reliabilität manualmedizinischer Untersuchungstechniken an der Halswirbelsäule. Studie zur Qualitätssicherung in der manuellen Diagnostik. *Z Orthop Ihre Grenzgeb* 138, 2-7, 2000.

Simons, D.G., Mense, S.: Diagnostik und Therapie myofaszialer Triggerpunkte. *Schmerz* 17, 419-425, 2003.

Strender LE, Sjöblom A, Sundell K, Ludwig R, and Taube A: Interexaminer Reliability in physical Examination of patients with Low Back Pain. *Spine* 22, 814-820, 1997.

Travell JG , Simons DG: Handbuch der Muskel-Triggerpunkte, 2. Aufl., Urban Fischer bei Elsevier, 2001.

-
- Van Deursen L, Snijders C, Patijn J: Influence of daily life activities on pain in patients with low back pain. *J O M* 24, 74-76, 2002.
- Van den Hoogen HJ, Koes BW, Deville W, van Eijk JT and Bouter LM: The interobserver reproducibility of Laseque´s sign in patient with low back pain in general practice. *Br J Gen Pract* 46, 727-730, 1996.
- Van der Wurff P, Hagmeijer RH, Meyne W: Clinical tests of the sacroiliac joint. A systematic methodological review Part 1: reliability. *Man Ther* 5, 30-36, 2000.
- Van der Wurff P, Meyne W, Hagmeijer RH: Clinical tests of the sacroiliac joint. A systematic methodological review Part 2: validity. *Man Ther* 5, 89-96, 2000.
- Van Der Wurff P, Buijs EF, Groen GJ: A multitest regimen of pain provocation tests as an aid to reduce unnecessary minimally invasive sacroiliac joint procedures. *Arch Phys Med Rehabil* 87, 10-14, 2006.
- Van-Dillen LR, Sahrman SA, Norton BJ, Caldwell CA, Fleming DA, McDonnell MK and Woolsey NB: Reliability of physical examination items used for classification of patients with low back pain. *Phys Ther* 78, 979-988, 1998.
- Vernon HT: Spinal manipulation for chronic low back pain: a review of the evidence. *Top Clin Chiropr* 6, 8-12, 1999.
- Vincent-Smith B, Gibbons P: Interexaminer and intraexaminer reliability of the standing flexion test. *Man ther* 4, 87-93, 1999.
- Wilson L, Hall H, McIntosh G and Melles T: Intetester reliability of a Low Back Pain classification System. *Spine* 24, 248-254, 1999.

Woodhead T, Glough A: A systematic review of the evidence for manipulation in the treatment of low back pain, *JOM* 27, 99-121, 2005.

9. Anhang

Entwickelter Studienbogen zur Erfassung der Testergebnisse

Fragebögen der Outcome-Measurements

SF-36

Ffb-H-R

Soziodemographischer Fragebogen

Rohdaten der Outcome-Measurements

SF-36

Ffb-H-R

Charakterisierung der Stichprobe

Soziodemographischer Datensatz

Lebenslauf

Ehrenwörtliche Erklärung

Entwickelter Studienbogen zur Erfassung der Testergebnisse

- Seite 1 -

Testergebnisse						
Studienzentrum:	[1]	[2]	[3]	[4]	[5]	[6]
Untersucher:	[1]	[2]	Patienten- Nr.:		<input type="checkbox"/> <input type="checkbox"/>	
			Datum:		<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	

Im Stehen

Seitneige	gestörte Bewegungssymmetrie :	Ja	[]	Nein	[]
	Bewegungsumfang gemindert rechts:	Ja	[]	Nein	[]
	Bewegungsumfang gemindert links:	Ja	[]	Nein	[]
	gestörte Bewegungsharmonie rechts:	Ja	[]	Nein	[]
Vorlauf	gestörte Bewegungsharmonie links:	Ja	[]	Nein	[]
	Positiv (Blockierung) rechts :	Ja	[]	Nein	[]
Spine	Positiv (Blockierung) links :	Ja	[]	Nein	[]
	Positiv (Blockierung) rechts :	Ja	[]	Nein	[]

Im Liegen

Patrick RL	gestörte Beweglichkeit rechts :	Ja	[]	Nein	[]
	Schmerz rechts :	Ja	[]	Nein	[]
	pathologisches Endgefühl rechts :	Ja	[]	Nein	[]
	gestörte Beweglichkeit links :	Ja	[]	Nein	[]
	Schmerz links :	Ja	[]	Nein	[]
Anteflexion SL	pathologisches Endgefühl links :	Ja	[]	Nein	[]
	Funktionseinschränkung L4:	Ja	[]	Nein	[]
	Funktionseinschränkung L5:	Ja	[]	Nein	[]

Im Liegen

Irritationspunkt BL	Tonuserhöhung L4 rechts :	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
	Tonuserhöhung L4 links :	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
	Tonuserhöhung L5 rechts :	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
	Tonuserhöhung L5 links :	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
Globale Federung BL	Schmerz:	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
Segmentale Federung BL	Schmerz L4:	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
	Schmerz L5:	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
Irritationspunkte ISG BL	Tonuserhöhung rechts :	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
	Tonuserhöhung links :	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
Kreuzgriff BL	Schmerz SIG rechts:	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
	Schmerz SIG links:	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
	Fkt-einschränkung rechts :	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
	Fkt-einschränkung links :	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
Heben BL	Fkt-einschränkung rechts :	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
	Fkt-einschränkung links :	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
Globalurteile	LWS-Funktionsstörung	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
	SIG-Blockierung rechts	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>
	SIG-Blockierung links	Ja	<input type="checkbox"/>	Nein	<input type="checkbox"/>

Unterschrift

Fragebogen SF-36

<p style="text-align: center;">Manualmedizinische Untersuchung</p> <p style="text-align: center;">Fragebogen zum Gesundheitszustand (SF-36 V2.)</p>

Datum ()

Patientennummer ()

In diesem Fragebogen geht es um Ihre Beurteilung Ihres Gesundheitszustandes. Der Bogen ermöglicht es, im Zeitverlauf nachzuvollziehen, wie Sie sich fühlen und wie Sie im Alltag zurechtkommen.

Bitte beantworten Sie jede der folgenden Fragen, in dem Sie bei den Antwortmöglichkeiten die Zahl ankreuzen, die am besten auf Sie zutrifft.

1. Wir würden Sie Ihren Gesundheitszustand im Allgemeinen beschreiben?

(Bitte kreuzen Sie nur eine Zahl an!)

- | | |
|----------------------|---|
| Ausgezeichnet: | 1 |
| Sehr gut: | 2 |
| Gut: | 3 |
| Weniger gut: | 4 |
| Schlecht: | 5 |

2. Im Vergleich zum vergangenen Jahr, wie würden Sie Ihren derzeitigen Gesundheitszustand beschreiben?

(Bitte kreuzen Sie nur eine Zahl an!)

- | | |
|--|---|
| Derzeit viel besser als vor einem Jahr: | 1 |
| Derzeit etwas besser als vor einem Jahr: | 2 |
| Etwa so wie vor einem Jahr: | 3 |
| Derzeit etwas schlechter als vor einem Jahr: | 4 |
| Derzeit viel schlechter als vor einem Jahr: | 5 |
-

3. Im folgenden sind einige Tätigkeiten beschrieben, die Sie vielleicht an einem normalen Tag ausüben. Sind Sie durch Ihren derzeitigen Gesundheitszustand bei diesen Tätigkeiten eingeschränkt? Wenn ja, wie stark?

(Bitte kreuzen Sie in jeder Zeile nur eine Zahl an!)

TÄTIGKEITEN	Ja, stark eingeschränkt	Ja, etwas eingeschränkt	Nein, überhaupt nicht eingeschränkt
a) anstrengende Tätigkeiten, z. B. schnell laufen, schwere Gegenstände heben, anstrengenden Sport treiben	1	2	3
b) mittelschwere Tätigkeiten, z. B. einen Tisch verschieben, staubsaugen, kegeln, Golf spielen	1	2	3
c) Einkaufstaschen heben oder tragen	1	2	3
d) mehrere Treppenabsätze steigen	1	2	3
e) einen Treppenabsatz steigen	1	2	3
f) sich beugen, knien, bücken	1	2	3
g) mehr als 1 Kilometer zu Fuß gehen	1	2	3
h) mehrere Straßenkreuzungen weit zu Fuß gehen	1	2	3
i) eine Straßenkreuzung weit zu Fuß gehen	1	2	3
j) sich baden oder anziehen	1	2	3

4. Hatten Sie in den vergangenen 4 Wochen aufgrund Ihrer körperlichen Gesundheit irgendwelche Schwierigkeiten bei der Arbeit oder anderen alltäglichen Tätigkeiten im Beruf bzw. zu Hause? (Bitte kreuzen Sie in jeder Zeile nur eine Zahl an!)

SCHWIERIGKEITEN	Immer	Meistens	Manchmal	Selten	Nie
a) Ich konnte nicht so lange wie üblich tätig sein	1	2	3	4	5
b) Ich habe weniger geschafft als ich wollte	1	2	3	4	5
c) Ich konnte nur bestimmte Dinge tun	1	2	3	4	5
d) Ich hatte Schwierigkeiten bei der Ausführung (z.B. Ich mußte mich besonders anstrengen)	1	2	3	4	5

5. Hatten Sie **in den vergangenen 4 Wochen aufgrund seelischer Probleme** irgendwelche Schwierigkeiten bei der Arbeit oder anderen alltäglichen Tätigkeiten im Beruf bzw. zu Hause (z. B. weil Sie sich niedergeschlagen oder ängstlich fühlten)?

(Bitte kreuzen Sie in jeder Zeile nur eine Zahl an!)

SCHWIERIGKEITEN	Immer	Meistens	Manchmal	Selten	Nie
a) Ich konnte nicht so lange wie üblich tätig sein	1	2	3	4	5
b) Ich habe weniger geschafft als ich wollte	1	2	3	4	5
c) Ich konnte nicht so sorgfältig wie üblich arbeiten	1	2	3	4	5

6. Wie sehr haben Ihre körperliche Gesundheit oder seelische Probleme **in den vergangenen 4 Wochen** Ihre normalen Kontakte zu Familienangehörigen, Freunde, Nachbarn oder zum Bekanntenkreis beeinträchtigt?

(Bitte kreuzen Sie nur eine Zahl an!)

- Überhaupt nicht: 1
- Etwas: 2
- Mäßig: 3
- Ziemlich: 4
- Sehr: 5

7. Wie stark waren Ihre Schmerzen **in den vergangenen 4 Wochen**?

(Bitte kreuzen Sie nur eine Zahl an!)

- Ich hatte keine Schmerzen: 1
- Sehr leicht: 2
- Leicht: 3
- Mäßig: 4
- Stark: 5
- Sehr stark: 6

8. Inwieweit haben die Schmerzen Sie in den vergangenen 4 Wochen bei der Ausübung Ihrer Alltagstätigkeit zu Hause und im Beruf behindert?

(Bitte kreuzen Sie nur eine Zahl an!)

- Überhaupt nicht: 1
- Ein bißchen: 2
- Mäßig: 3
- Ziemlich: 4
- Sehr: 5

9. In diesen Fragen geht es darum, wie Sie sich fühlen und wie es Ihnen in den vergangenen 4 Wochen gegangen ist. (Bitte kreuzen Sie in jeder Zeile die Zahl an, die Ihrem Befinden am ehesten entspricht!) Wie oft waren Sie in den vergangenen 4 Wochen ...

(Bitte kreuzen Sie in jeder Zeile nur eine Zahl an)

BEFINDEN	Immer	Meistens	Manchmal	Selten	Nie
a) ... voller Schwung?	1	2	3	4	5
b) ... sehr nervös?	1	2	3	4	5
c) ... so niedergeschlagen, daß Sie nichts aufheitern konnte?	1	2	3	4	5
d) ... ruhig und gelassen?	1	2	3	4	5
e) ... voller Energie?	1	2	3	4	5
f) ... entmutigt und traurig?	1	2	3	4	5
g) ... erschöpft?	1	2	3	4	5
h) ... glücklich?	1	2	3	4	5
i) ... müde?	1	2	3	4	5

10. Wie häufig haben Ihre körperliche Gesundheit oder seelischen Probleme in den vergangenen 4 Wochen Ihre Kontakte zu anderen Menschen (Besuche bei Freunden, Verwandten usw.) beeinträchtigt?

(Bitte kreuzen Sie nur eine Zahl an!)

- Immer: 1
- Meistens: 2
- Manchmal: 3
- Selten: 4
- Nie: 5

11. Inwieweit trifft **jede** der folgenden Aussagen auf Sie zu?

(Bitte kreuzen Sie in jeder Zeile nur eine Zahl an!)

AUSSAGEN	Trifft ganz zu	Trifft weitgehend zu	Weiß nicht	Trifft weitgehend nicht zu	Trifft überhaupt nicht zu
a) Ich schein etwas leichter als andere krank zu werden	1	2	3	4	5
b) Ich bin genauso gesund wie alle anderen, die ich kenne	1	2	3	4	5
c) Ich erwarte, daß mein Gesundheit nachläßt	1	2	3	4	5
d) Ich erfreue mich ausgezeichneter Gesundheit	1	2	3	4	5

Fragebogen FFb-H-R

Manualmedizinische Untersuchung

Fragebogen zum Stand der Aktivitäten (FFb-H-R)

Datum ()

Patientennummer ()

In den folgenden Fragen geht es um Tätigkeiten aus dem täglichen Leben.

Bitte beantworten Sie jede Frage so, wie es für Sie **im Moment** (wir meinen in Bezug auf die letzten 7 Tage) zutrifft.

Sie haben **drei** Antwortmöglichkeiten:

- | | | |
|-----|---------------------------------|--|
| [1] | Ja | d.h. Sie können die Tätigkeit ohne Schwierigkeit ausführen. |
| [2] | Ja, aber mit Mühe | d.h. Sie haben dabei Schwierigkeiten, z.B. Schmerzen, es dauert länger als früher, oder Sie müssen sich dabei abstützen. |
| [3] | Nein oder nur mit fremder Hilfe | d.h. Sie können es gar nicht oder nur, wenn eine andere Person Ihnen dabei hilft. |

- | | | | |
|--|-----|-----|-----|
| Können Sie sich strecken, um z.B. ein Buch von einem hohen Schrank oder Regal zu holen ? | [1] | [2] | [3] |
| Können Sie einen mindestens 10 kg schweren Gegenstand (z.B. vollen Wassereimer oder Koffer) hochheben und 10 Meter weit tragen ? | [1] | [2] | [3] |
| Können Sie sich von Kopf bis Fuß waschen und abtrocknen ? | [1] | [2] | [3] |
| Können Sie sich bücken und einen leichten Gegenstand (z.B. Geldstück oder zerknülltes Papier) vom Fußboden aufheben ? | [1] | [2] | [3] |
| Können Sie sich über einem Waschbecken die Haare waschen ? | [1] | [2] | [3] |
| Können Sie 1 Stunde auf einem ungepolsterten Stuhl sitzen ? | [1] | [2] | [3] |
| Können Sie 30 Minuten ohne Unterbrechung stehen (z.B. in einer Warteschlange) ? | [1] | [2] | [3] |
| Können Sie sich im Bett aus der Rückenlage aufsetzen ? | [1] | [2] | [3] |
| Können Sie Strümpfe an- und ausziehen ? | [1] | [2] | [3] |
| Können Sie im Sitzen einen kleinen heruntergefallenen Gegenstand (z.B. eine Münze) neben Ihrem Stuhl aufheben ? | [1] | [2] | [3] |
| Können Sie einen schweren Gegenstand (z.B. einen gefüllten Kasten Mineralwasser) vom Boden auf den Tisch stellen ? | [1] | [2] | [3] |
| Können Sie 100 Meter schnell laufen (nicht gehen), etwa um einen Bus noch zu erreichen ? | [1] | [2] | [3] |
-

Welche Berufsausbildung haben Sie abgeschlossen?					
<input type="checkbox"/>	Lehre (berufliche- betriebliche Ausbildung)	<input type="checkbox"/>	Fachhochschule, Ingenieurschule	<input type="checkbox"/>	andere Berufsausbildung
<input type="checkbox"/>	Fachschule (Meister-, Technikerschule, Berufsfachakademie)	<input type="checkbox"/>	Universität, Hochschule	<input type="checkbox"/>	keine Berufsausbildung

Sind Sie zur Zeit erwerbstätig?					
<input type="checkbox"/>	ja, ganztags	<input type="checkbox"/>	nein, Hausfrau / Hausmann	<input type="checkbox"/>	nein, Erwerbs-, BU-Rente
<input type="checkbox"/>	ja, mindestens halbtags	<input type="checkbox"/>	nein, in Ausbildung	<input type="checkbox"/>	nein, Altersrente
<input type="checkbox"/>	ja, weniger als halbtags	<input type="checkbox"/>	nein, arbeitslos / erwerbslos	<input type="checkbox"/>	nein, anderes

In welcher beruflichen Stellung sind Sie hauptsächlich derzeit beschäftigt bzw. (falls nicht mehr berufstätig) waren Sie zuletzt beschäftigt? *₁					
<input type="checkbox"/>	Arbeiter	<input type="checkbox"/>	Beamter	<input type="checkbox"/>	Sonstiges
<input type="checkbox"/>	Angestellter	<input type="checkbox"/>	Selbständiger		

Wie hoch ist das monatliche Nettoeinkommen*₂ Ihres Haushaltes insgesamt?					
<input type="checkbox"/>	bis unter 500 Euro	<input type="checkbox"/>	500 bis unter 1000 Euro	<input type="checkbox"/>	1000 bis unter 1500 Euro
<input type="checkbox"/>	1500 bis unter 2000 Euro	<input type="checkbox"/>	2000 bis unter 2500 Euro	<input type="checkbox"/>	2500 bis unter 3000 Euro
<input type="checkbox"/>	3000 bis unter 3500 Euro	<input type="checkbox"/>	3500 Euro und mehr		

*1) Minimalabfrage zur Berufsklassifizierung, zusammen mit Schulbildung, Berufsbildung und Einkommen

*2) Nettoeinkommen: Die Summe aus Lohn / Gehalt / Einkommen usw., jeweils nach Abzug der Steuern und Sozialabgaben

Outcome-Measurements Rohdaten

SF 36				
Gesamt				
	MV	min	max	n
PF	62,1	0,0	240,0	144
RP	44,5	0,0	100,0	142
BP	32,6	0,0	100,0	145
GH	51,2	5,0	92,0	143
VT	41,8	6,3	81,3	141
SF	61,6	12,5	100,0	145
RE	61,8	0,0	100,0	142
MH	58,5	10,0	100,0	142
PCS	36,9	18,1	82,2	138
MCS	43,8	12,5	63,4	138
VAS Gesundheitszustand	60,6	3,0	100,0	133
männlich				
PF	64,6	10,0	100,0	71
RP	46,5	0,0	100,0	69
BP	33,2	0,0	74,0	71
GH	50,7	20,0	82,0	70
VT	41,8	12,5	81,3	70
SF	63,6	12,5	100,0	71
RE	64,4	8,3	100,0	70
MH	59,7	20,0	100,0	70
PCS	37,1	23,0	53,8	68
MCS	44,2	20,9	61,2	68
VAS Gesundheitszustand	62,8	20,0	95,0	67
weiblich				
PF	59,7	0,0	240,0	73
RP	42,6	0,0	100,0	73
BP	32,1	0,0	100,0	74
GH	51,7	5,0	92,0	73
VT	41,9	6,3	75,0	71
SF	59,6	12,5	100,0	74
RE	59,3	0,0	100,0	72
MH	57,4	10,0	90,0	72
PCS	36,6	18,1	82,2	70
MCS	43,5	12,5	63,4	70
VAS Gesundheitszustand	58,3	3,0	100,0	66

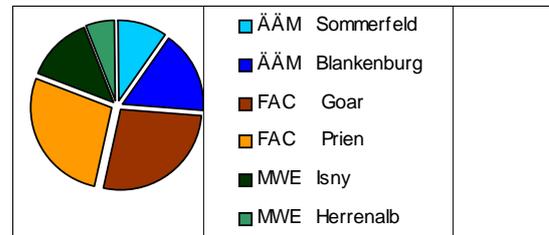
Fragebogen zum Stand der Aktivitäten (FFb-H-R)						
	Gesamt		männlich		weiblich	
	Anzahl	%	Anzahl	%	Anzahl	%
nach oben strecken						
ja	86	59	45	63	41	55
mit mühe	52	36	24	34	28	38
nur mit fremder Hilfe	7	5	2	3	5	7
10 kg 10m tragen						
ja	52	36	37	52	15	21
mit mühe	61	42	28	39	33	45
nur mit fremder Hilfe	31	22	6	8	25	34
selbst waschen						
ja	110	75	54	76	56	75
mit mühe	35	24	17	24	18	24
nur mit fremder Hilfe	1	1			1	1
kleine Gegenstände vom Boden aufheben						
ja	102	70	50	70	52	69
mit mühe	40	27	20	28	20	27
nur mit fremder Hilfe	4	3	1	1	3	4
am Waschbecken die Haare waschen						
ja	71	49	39	55	32	43
mit mühe	58	40	29	41	29	39
nur mit fremder Hilfe	16	11	3	4	13	18
1h auf ungepolsterten Stuhl						
ja	36	25	23	33	13	17
mit mühe	87	60	40	57	47	63
nur mit fremder Hilfe	22	15	7	10	15	20
30 min stehen						
ja	34	23	20	28	14	19
mit mühe	88	60	42	59	46	61

nur mit fremder Hilfe	24	16	9	13	15	20
aus Rückenlage aufsetzen						
ja	72	49	37	52	35	47
mit mühe	58	40	30	42	28	37
nur mit fremder Hilfe	16	11	4	6	12	16
Stümpfe anziehen						
ja	97	66	48	68	49	65
mit mühe	48	33	23	32	25	33
nur mit fremder Hilfe	1	1			1	1
im Sitzen kleine Gegenstände vom Boden aufheben						
ja	97	67	48	69	49	65
mit mühe	42	29	21	30	21	28
nur mit fremder Hilfe	6	4	1	1	5	7
schweren Gegenstand vom Boden auf den Tisch heben						
ja	43	29	30	42	13	17
mit mühe	65	45	33	46	32	43
nur mit fremder Hilfe	38	26	8	11	30	40
100 m schnell laufen						
ja	53	36	25	35	28	37
mit mühe	63	43	33	46	30	40
nur mit fremder Hilfe	30	21	13	18	17	23

Charakterisierung der Stichprobe

Gesamt					
	MV	min	max	Männlich (48%)	Weiblich (52%)
Größe	171	152	196	178	165
Gewicht	76	49	106	86	67
BMI	26	18	34	27	25
Alter	48	20	80	47	48
VAS	60	3	100	62	58
Gesundheitszustand					

Verteilung auf die Schulen			
		Anzahl	%
ÄÄM Sommerfeld		15	9,93
ÄÄM Blankenburg		25	16,56
FAC Goar		40	26,49
FAC Prien		42	27,81
MWE Isny		20	13,25
MWE Herrenalb		9	5,96



soziodemographische Daten

		Personen im Haushalt			
				%	
	Anzahl	%	1	2	3
			22	15,17	
männlich	73	48,34	2	51	35,17
weiblich	78	51,66	3	29	20,00
			4	26	17,93
deutsch	143	97,28	5	10	6,90
nicht-deutsch	4	2,72	6	5	3,45
			7	1	0,69
ledig	24	16,33	8	1	0,69
verheiratet	99	67,35			
geschieden/getrennt	17	11,56	davon sind älter als 18		
verwitwet	7	4,76	1	23	16,20
			2	89	62,68
fester Partner ja	116	78,91	3	22	15,49
fester Partner nein	31	21,09	4	5	3,52
			5	1	0,70
Haupt-	52	35,37	6		

Volksschule					
POS-Real-Mittlere Reife	60	40,82	7	1	0,70
Fachhochschulreife	10	6,80	8	1	0,70
Abitur	19	12,93			
anderen Schulabschluß	2	1,36			
kein Schulabschluß	4	2,72			
Lehre	75	52,82	Nettoeinkommen (Euro)		
Fachschule Meister usw	26	18,31	unter 500	5	3,82
Fachhochschule Ingenieur	13	9,15	500 - 1000	19	14,50
Uni Hochschule andere	10	7,04	1000-1500	29	22,14
Berufsausbildung keine Berufsausbildung	13	9,15	1500-2000	27	20,61
ganztags Arbeit mindestens halbtags weniger als halbtags	5	3,52	2000-2500	19	14,50
Hausfrau (mann)	9	6,25	2500-3000	13	9,92
in Ausbildung	1	0,69	3000-3500	7	5,34
arbeitslos	13	9,03	3500 und mehr	12	9,16
Erwerbs-EU-Rente	1	0,69			
Altersrentner	9	6,25	Arbeiter	41	28,67
anderes	4	2,78	Angestellter	72	50,35
			Beamter	21	14,69
			Selbständiger	5	3,50
			Sonstiges	4	2,80

Einverständniserklärung



Universitätsklinikum Jena

Universitätsklinikum Jena · Postfach · 07740 Jena

Medizinische Fakultät

Institut
für Physiotherapie

Prof.
Dr. med. U. Smolenski

Kollegiengasse 9
D-07743 Jena

Telefon: 0 36 41 · 93 78 30
Telefax: 0 36 41 · 93 78 32

E-Mail: smolensk@mti-n.uni-
jena.de

Einverständniserklärung

Hiermit erkläre ich

Name: Vorname: geb. am

wohnhaft in:

mich mit der Teilnahme an der Untersuchung zur Reliabilität der manualmedizinischen Tests bei Low back pain (LBP)-Patienten einverstanden.

Über Ziele und Inhalte der Untersuchung sowie den geplanten Ablauf bin ich durch den unterzeichnenden Arzt informiert worden.

Ich wurde informiert, dass meine Daten in anonymisierter Form gespeichert und ausgewertet werden.

Ich habe jederzeit die Möglichkeit und das Recht, die Teilnahme an der Untersuchung abzubrechen.

.....
Unterschrift des Probanden

.....
Unterschrift des Arztes

....., den

Lebenslauf

Vorname Name: Steffen Derlien, geb. Conradi

Geburtsdatum: 29.06.1969

Geburtsort: Homberg / Efze

Familienstand: verheiratet

Schulbesuch:

Grundschule	1975 – 77 Grundschule Werkel
	1977 – 79 Osterbachgrundschule Homberg/Efze
	1979 – 80 Realschule Homberg/Efze
Gymnasium von	1980 – 86 Ursulinenschule Fritzlar
	1986 – 89 Goetheschule Kassel
Abitur:	17.05.1989

Studium: Studium der Sportwissenschaft und Biologie für das Lehramt an Gymnasien
Ergänzungsrichtung „Rehabilitations- und Alterssport“
Ergänzungsrichtung „ Förderung sportlicher Begabungen“
Friedrich-Schiller-Universität Jena
WS 1990 – 26. März 1996

Abschluss: 1. Staatsexamen für das Lehramt an Gymnasien
Ergänzungsrichtung : Rehabilitations- und Alterssport“
Ergänzungsrichtung „ Förderung sportlicher Begabungen“

Berufliche Tätigkeit: 06.05.1996 – 31.08.2003 Sportwissenschaftler / Sporttherapeut Hufeland-Kliniken Weimar
01.02.2002 – 31.08.2003 Sportwissenschaftler / Sporttherapeut Forschungsprojektstelle Stelle (50%) am Klinikum der FSU Jena - Institut für Physiotherapie
01.09.2003 – 31.06.2007 befristete Vollbeschäftigung als Sportwissenschaftler / Sporttherapeut am Klinikum der FSU Jena – Institut für Physiotherapie
seit 01.07.2007 unbefristet

Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass mir die geltende Promotionsordnung der Fakultät für Sozial- und Verhaltenswissenschaften bekannt ist. Ich habe die Dissertation selbst angefertigt und alle von mir benutzten Hilfsmittel, persönliche Mitteilungen und Quellen in meiner Arbeit angegeben habe. Weiterhin erkläre ich, dass mich keine Personen bei der Auswahl und Auswertung des Materials unterstützt haben.

Ferner erkläre ich, dass ich nicht die Hilfe einer Promotionsberaters in Anspruch genommen habe und dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Dissertation wurde noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht. Zudem wurde keine gleiche, keine in wesentlichen Teilen ähnliche und auch keine andere Abhandlung bei einer anderen Hochschule bzw. anderen Fakultät als Dissertation eingereicht.

Ich versichere, nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen zu haben.

Steffen Derlien

Jena, den 21. Januar 2009

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mich in der langen Zeit der Anfertigung dieser Arbeit unterstützt haben.

Mein besonderer Dank gilt Herrn Prof. Dr. med. U. C. Smolenski. Er hat es als Chef immer verstanden, die richtigen Worte des Zuspruchs und des Antriebs zu finden. Zudem fehlte es mir nie an der nötigen konstruktiven und freundlichen Unterstützung. Nicht zuletzt ist es seiner Initiative zu verdanken, dass die Deutsche Gesellschaft für Manuelle Medizin einen Teil dieser Arbeit als initiales Forschungsprojekt finanziell unterstützte. Mein Dank gilt auch der Deutschen Gesellschaft dafür.

Selbstverständlich möchte ich meinen Dank auch Herrn Prof. Dr. med. H. Gabriel aussprechen, der sich bereit erklärt hat, diese Arbeit als Betreuer zu übernehmen. Seine wertvollen Hinweise haben es ermöglicht, diese Arbeit so fertig zu stellen.

Ein ganz herzlichen Dank möchte ich PD Dr. G. Kirchner aussprechen. Durch ihn habe ich den Weg zum wissenschaftlichen Arbeiten, das wachsame Auge gegenüber Daten und den kritischen Umgang mit medizinischen Lehrmeinungen gelehrt bekommen. Vielen Dank für die vielen Korrekturen und die mehr als freundschaftliche Unterstützung, nicht nur im Bezug auf die wissenschaftliche Arbeit

Allen beteiligten Ärzten möchte ich für die Unterstützung in Form von Untersuchungsdurchführungen in den Untersuchungskliniken und für die Einhaltung der wissenschaftlich geforderten Kriterien danken. Namentlich sind das: Fr. Dr Kluge, Fr. Dr. Schroeder, Hr. Dr. Sturm, Hr. Dr. Nagel, Hr. Dr. Psczolla, Hr. Dr. Galeazzi, Hr. Dr. Wilde, Hr. Dr. Wagner, Hr. Dr. Biskupek, Hr. Dr. Müller, Hr. Dr. Frey, Hr. Dr. Klaas.

Fr. Loth, ehemalige Mitarbeiterin am Institut für Physiotherapie, möchte ich für die Unterstützung bei der Dateneingabe und –bearbeitung danken.

Fr. Dipl Psych. Kristin Neumer möchte ich für den kritischen Umgang mit den Inhalten der Arbeit und für das Korrekturlesen danken.

Zum Schluss möchte ich meiner Frau Anja danken, die viel Mühe hatte, mir die korrekte Benutzung von „s“, „ss“ und „ß“ beizubringen. Ich hoffe ihre Schüler sind gelehriger. Vielen und vor allem herzlichen Dank nicht nur dafür...
