

Wie reliabel ist die Identifikation der somatischen Dysfunktion in der Osteopathie?

Ein systematisches Review

MASTER – THESIS

zur Erlangung des akademischen Grades

Master of Science

im Universitätslehrgang Osteopathie

vorgelegt von

Sebastian Soika BSc

Matrikelnummer: 11852185

Department für Gesundheitswissenschaften, Medizin und Forschung

an der Donau-Universität Krems

Betreuerin: Katharina Wimmer, MSc



Lienz, 31.Okttober 2020

Eidesstattliche Erklärung

Ich, Sebastian Soika B.Sc., geboren am 16.09.1984 in Eggenfelden, erkläre,

1. dass ich meine Master Thesis selbständig verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubten Hilfen bedient habe,
2. dass ich meine Master Thesis bisher weder im In- noch im Ausland in irgendeiner Form als Prüfungsarbeit vorgelegt habe,
3. dass ich, falls die Master Thesis mein Unternehmen oder einen externen Kooperationspartner betrifft, meinen Arbeitgeber über Titel, Form und Inhalt der Master Thesis unterrichtet und sein Einverständnis eingeholt habe.

Sebastian Soika

Danksagungen / Widmungen

Ich bedanke mich bei meiner Frau Vera für die unendliche Geduld, mit der sie mir begegnet.

Auch gilt mein Dank meiner Tochter Lena, die mir gezeigt hat, was wirklich im Leben zählt.

Danke für die unendliche Inspiration, die ihr für mich darstellt.

Zusammenfassung

Hintergrund: Die somatische Dysfunktion steht im Zentrum der osteopathischen Diagnostik. Um diese erfolgreich zu diagnostizieren, werden vier Kriterien definiert, die unter dem Akronym TART zusammengefasst werden. Dabei wird nach Parametern wie Asymmetrie, Bewegungseinschränkung, Gewebeveränderungen und Gewebeempfindlichkeit gesucht. Die Forschungsfrage dieses Reviews bezieht sich auf die Reliabilität der Identifikation der somatischen Dysfunktion. Weiteres soll geklärt werden, ob die somatische Dysfunktion eine sinnvolle Hypothese oder ein längst überholtes Konzept darstellt.

Methodik: Für die Literatursuche dieses systematischen Reviews wurden die Datenbanken und Journals von Pubmed, International Journal of Osteopathic Medicine, Ostepathic research web, Cochrane Library, PEDro, sowie dem Journal of Bodywork and Movement Therapies durchsucht. Aus den insgesamt 333 Studien wurden 20 Artikel zur Bewertung mit der QAREL-Checkliste ausgewählt.

Ergebnisse: Die Studien konnten durchwegs eine gute bis sehr gute methodologische Qualität aufweisen. Die Ergebnisse der Reliabilitätstestung waren allgemein heterogen. Die Asymmetrie wurde am häufigsten untersucht, wies aber auch die schlechtesten Ergebnisse auf. Die Kriterien Gewebebeschaffenheit, Empfindlichkeit und Bewegungsausmaß zeigten teils eine gute Reliabilität auf. Studien mit vorherigem Übereinstimmigkeitstraining wiesen allgemein bessere Ergebnisse auf.

Konklusion: Aufgrund der Ergebnisse wird eine kombinierte Testung aller TART-Kriterien empfohlen. Von einer isolierten statischen Untersuchung des Beckens kann derzeit abgeraten werden. Für zukünftige Studien wird ein Übereinstimmigkeitstraining empfohlen, da dies durchwegs zu besseren Ergebnissen der Reliabilität führte.

Schlüsselwörter: Somatische Dysfunktion, Asymmetrie, Bewegungseinschränkung, Gewebeveränderung, Gewebeempfindlichkeit

Abstract

Background: The osteopathic dysfunction focuses on the somatic dysfunction. To successfully diagnose it, four categories, which are combined under the acronym TART, are defined. Therefore parameters like asymmetry, restriction of motion, tissue change and tissue sensibility are searched. The research question of the review refers to the reliability of the identification of the somatic dysfunction. Furthermore it has to be clarified if the somatic dysfunction is a useful theory (supposition) or an outdated concept.

Methodology: For the literature search of this systematic review data bases and journals of Pubmed, International Journal of Osteopathic Medicine, Osteopathic research web, Cochrane Library, PEDro, as well as the journal of Bodywork and Movement Therapies were searched through. Out of the overall 333 studies, twenty articles were chosen for evaluation with the QAREL-check list.

Results: The methodological quality of the studies could be described as good to very good throughout. The results of the reliability test in general were heterogeneous. The asymmetry was researched the most but showed the poorest result. The criteria tissue consistency, sensitivity and range of movement partly showed a good reliability. Studies with previous consensus training presented overall better results.

Conclusion: Due to the results a combined test of all the TART-criteria is recommended. For now an isolated static examination of the pelvis is dissuaded. For future studies a consensus training is recommended because it consistently showed better results of reliability.

Keywords: Somatic dysfunction, asymmetry, restriction of motion, tissue texture changes, tenderness

Inhaltsverzeichnis

Eidesstattliche Erklärung.....	I
Danksagungen / Widmungen	II
Zusammenfassung.....	III
Abstract	IV
1. Einleitung	1
1.1. Somatische Dysfunktion: „A little history“	1
1.2. Die somatische Dysfunktion - sinnvolle Hypothese oder überholtes Konzept?	3
2. Theoretischer Hintergrund	5
2.1. Erklärungsmodelle der somatischen Dysfunktion.....	5
2.2. Gewebefaktoren als Erklärungsmodell	6
2.2.1. Verletzungen und Inflammation der Facettengelenke	6
2.2.2. Einklemmung von meniskoiden Strukturen in und um das Facettengelenk	7
2.2.3. Periartikuläres Gewebe wie Muskeln, Ligamente und Gelenkkapseln	7
2.2.4. Derangement oder Degeneration der Bandscheiben	8
2.3. Notwendige neuroanatomische Grundlagen	8
2.3.1. Afferente Rezeptoren und Nervenfasereinteilung.....	8
2.3.2. Faserkategorien peripherer Nervenbahnen.....	10
2.3.3. Organisation im Hinterhorn und Rückenmark	12
2.4. Neurologisches Modell der somatischen Dysfunktion.....	14
2.4.1. Das fazilitierte Segment	15
2.4.2. Zentrale Sensibilisierung.....	16
2.5. Diagnostik der somatischen Dysfunktion	18

3. Methodologie	20
3.1. Forschungsfrage	20
3.2. Vorannahme	20
3.3. Studien- und Forschungsdesign	20
3.4. Einschlusskriterien	20
3.5. Ausschlusskriterien	21
3.6. Literatursuche	21
3.7. Datenanalyse mit der QAREL-Checkliste	26
4. Ergebnisse	34
4.1. Reliability of diagnosis of somatic dysfunction among osteopathic physicians and medical students (Bengaard et al., 2012)	34
4.2. Measurement challenges in physical diagnosis: refining inter-rater palpation, perception and communication (Comeaux et al., 2001).....	37
4.3. Interrater Reliability of Osteopathic Sacral Palpation Diagnostic Tests Among Osteopathy Students (Consorti et al., 2018).....	38
4.4. Interobserver Reliability of Osteopathic Palpation Diagnostic Tests of the Lumbar Spine: Improvements from Consensus Training (Degenhardt et al., 2005).....	40
4.5. Maintenance and Improvement of Interobserver Reliability of Osteopathic Palpation Tests Over a 4-Month Period (Degenhardt et al., 2010).....	42
4.6. Factors affecting the intra-examiner and inter-examiner reliability of palpation for supine medial malleoli asymmetry (Fryer, 2006)	44
4.7. The effect of training on the inter-examiner and intra-examiner reliability of the seated flexion test and assessment of pelvic anatomical landmarks with palpation (Fryer et al., 2005).....	45
4.8. Inter-examiner and intra-examiner agreement for assessing simulated leg length inequality using palpation and observation during a standing assessment (Gibbons et al., 2002) 48	

4.9. Reliabilitätsstudie des Befunds nach dem Mechanical Link: Methodologische Studie über die Befunderhebung der Knochen und Gelenke der Extremitäten nach der Methode des Mechanical Link (Hafen, 2019).....	50
4.10. Interexaminer reliability study of a standardized myofascial diagnostic technique of the superior thoracic inlet (Hutchinson et al., 2017).....	52
4.11. Reliability of physical examination to assess asymmetry of anatomical landmarks indicative of pelvic somatic dysfunction in subjects with and without low back pain (Kmita & Lucas, 2008).....	53
4.12. Accuracy of Anterior Superior Iliac Spine Symmetry Assessment by Routine Structural Examination (Lee et al., 2015)	55
4.13. Counterstrain and traditional osteopathic examination of the cervical spine compared (McPartland & Goodridge, 1997).....	57
4.14. Inter-examiner and intra-examiner agreement for assessing sacroiliac anatomical landmarks using palpation and observation: pilot study (O'Haire & Gibbons, 2000)	59
4.15. Inter-examiner reliability of palpation for tissue texture abnormality in the thoracic paraspinal region (Paulet & Fryer, 2009).....	60
4.16. The assessment of pelvic landmarks using palpation: A reliability study of undergraduate students (Rajendran & Gallagher, 2011)	62
4.17. Intra-examiner and inter-examiner reliability of a positional diagnostic screen for the lumbar spine (Spring et al., 2001)	63
4.18. The effects of experience on the inter-reliability of osteopaths to detect changes in posterior superior iliac spine levels using a hidden heel wedge (Sutton et al., 2013)	65
4.19. Interexaminer reliability of three methods of combining test results to determine side of sacral restriction, sacral base position, and innominate bone position (Tong et al., 2006) 67	
4.20. Inter-examiner and intra-examiner reliability of the standing flexion test (Vincent-Smith & Gibbons, 1999).....	68
4.21. Zusammenfassung der Ergebnisse	70

5. Diskussion	75
Tabellenverzeichnis.....	88
Abbildungsverzeichnis	91
Abkürzungsverzeichnis	92

1. Einleitung

1.1. Somatische Dysfunktion: „A little history“

Schon A.T. Still sprach bei der Auffindung von Funktionsstörungen im muskuloskeletalen System von einer sogenannten „osteopathischen Läsion“. Bereits im Jahre 1874 stellte er die Theorie auf, dass ein Krankheitsprozess in Gang kommen kann, wenn der Lebensfluss unterbrochen ist. Bei seiner damaligen Vorstellung einer Läsion, kommt es zur Obstruktion von Körperflüssigkeiten wie Blut, Lymph- und Cerebrospinalflüssigkeit (Still, 1908). Er glaubte, dass eine Obstruktion dieser Flüssigkeiten, oder ihrer neuronalen Kontrollzentren, den Fluss derer behindern könnte (McCole, 1935). Die ursprüngliche Bezeichnung „osteopathische Läsion“, wird oft auch als Synonym verwendet mit Termini wie chiropraktische Subluxation, Gelenkblockierung oder osteopathische Dysfunktion (Meert, 2009), intervertebrale Dysfunktion, intervertebrale Läsion oder segmentale Dysfunktion (Fryer, 2016).

Erste neurophysiologische Erklärungen der osteopathischen Läsion lieferten J. Stedman Denslow, D.O. und Irvin Korr, Ph.D. in den späten 1940er Jahren (Denslow, 1953; Korr, 1947). Korr postulierte dabei den Begriff des fazilitierten Rückenmarksegments. Seine Erklärung war, dass die Belastung eines Rückenmarksegments reduziert sei, und es dadurch schneller zu einer osteopathischen Läsion kommen könne. Die Belastbarkeit des Segments wird dabei durch die Fazilitierung herabgesetzt. Außerdem kommt noch eine neurovegetative Dysbalance zwischen Sympathikus und Parasympathikus als zusätzlicher Reiz hinzu. Ein weiterer Störfaktor wie ein Kältereziz oder ein Bagatelltrauma in das bereits geschwächte Segment, kann zur gesteigerten Reizantwort in alle segmental verbundenen Körperstrukturen wie Myotom, Sklerotom, Dermatom oder Viscerotom führen. Die Folge daraus ist eine neurologische Überlastung des Segments, die als osteopathische Läsion bezeichnet wird (Meert, 2009). Demnach ist die osteopathische Läsion weit mehr als eine rein strukturelle Gelenkblockierung. Durch die Weiterentwicklungen des Konzepts der osteopathischen Läsion wurde die internationale Terminologie im Jahre 1975 auf den Begriff somatische Dysfunktion (SD) geändert (Comeaux, 2005). Diese definiert sich laut der American Association of Colleges of Osteopathic Medicine als eine veränderte oder eingeschränkte Funktion des muskuloskeletalen Systems und der damit in Verbindung stehenden neurovaskulären Strukturen (Principles & Medicine, 1982).

Greenman gibt in seinem Lehrbuch der osteopathischen Medizin folgende Kriterien zur Diagnostik der SD an (DeStefano, 2011):

- **A** steht für asymmetry. Beispiele für Asymmetrie sind z.B. die Schulterhöhe, die Höhe der Beckenkämme, sowie Form und Funktion des Thorax.
- **R** steht für restriction of motion. Darunter versteht man Abweichungen im Bewegungsausmaß eines Gelenks. Dies ist für Greenman die bedeutendste Komponente der SD. Die Osteopathin oder der Osteopath interessiert sich hierbei für die Quantität, die Qualität und das Endgefühl einer Bewegung.
- **T** steht für tissue texture abnormality/changes, daher für Veränderungen der Gewebebeschaffenheit und Veränderungen in den Weichteilen (Haut, Faszien, Muskeln, Bändern).

Im Leitfaden Osteopathie (Liem & Dobler, 2017) wird von den Autoren ein weiteres „T“ angegeben.

- T als Abkürzung für Tenderness, also für eine erhöhte Empfindlichkeit des betroffenen Gewebes, wird zu den bereits beschriebenen Kürzeln hinzugefügt. Dadurch entsteht das Akronym **TART**.

Das Konzept wurde im Laufe der Zeit immer weiter erforscht und durch Untersuchungen von Wilbure Cole, D.O., Michael M. Patterson, Ph. D., Richard L. van Buskirk, D.O., Ph. D., Gary Fryer Ph. D., B.Sc., und Paolo Tozzi, M.Sc. Ost., D.O. P.T. ergänzt und erweitert (Liem, 2016). Lederman (2000) übt in seiner Übersichtsarbeit über das fazilitierte Segment Kritik an dem Konzept und gibt an, wie wichtig ein besseres Verständnis der Neurophysiologie für die osteopathische Praxis sei.

Durch die zunehmende Forderung der Wissenschaft, eine evidenzbasierte Praxis anzustreben, gerät jedoch die Theorie der somatischen Dysfunktion zunehmend unter Kritik (Fryer, 1999, 2016). Damit stellt sich die Frage, ob die Theorie der SD eine sinnvolle Hypothese ist, oder ein längst überholtes Konzept darstellt. Auch Tozzi (2015) versucht in seiner Arbeit, diese Forschungsfrage aufzugreifen und spricht von einem mehr neurologisch-faszialen Modell der SD.

Die Forschungsfrage der vorliegenden Arbeit bezieht sich auf die Reliabilität der TART-Kriterien und wie diese zur Identifikation einer SD eingesetzt werden können. Es soll erklärt werden, ob und welche der Kriterien sich dafür eignen und wie sich dies auf die tägliche Arbeit in der osteopathischen Praxis auswirken könnte.

In Ausbildungsinstituten für Osteopathinnen und Osteopathen, sowie in einigen Schulen für manuelle Therapie und manuelle Medizin, werden die TART-Kriterien zur Erkennung der SD gelehrt, ohne die Reliabilität dieser Untersuchungsmerkmale zu hinterfragen. Die persönliche Motivation des Autors richtet sich daher auf die Plausibilität des Konzepts und inwieweit dies eine Relevanz für die osteopathische Arbeit und Forschung darstellt. Die Darlegung der Reliabilität in der Diagnostik der SD stellt für die weitere Evidenz der osteopathischen Praxis eine zentrale Bedeutung dar.

1.2. Die somatische Dysfunktion - sinnvolle Hypothese oder überholtes Konzept?

Im ersten Ausbildungsjahr des Studiums der Osteopathie wird die Studentin oder der Student schon bald mit dem Begriff der SD und seiner diagnostischen sowie therapeutischen Relevanz konfrontiert. Das Akronym TART wird dabei als diagnostisches Kriterium angegeben. Wie genau die einzelnen Kriterien allerdings die SD widerspiegeln oder ob diese überhaupt der Physiologie unseres Körpers entsprechen, wird selten hinterfragt. Zum kleinen Denkanstoß einige Beispiele:

- Kann ein Bewegungsverlust nicht auch eine Kompensationsreaktion des Körpers darstellen? Nach einem Schleudertrauma z.B. macht es wenig Sinn, an Restriktionen der Halswirbelsäule zu arbeiten, da diese reflektorisch als Schutzreaktion dienen und somit nicht primär behandelt werden sollten (Kia, 2013).
- Hypermobilität findet in dem Konzept der SD und den TART Kriterien so gut wie keine Erwähnung, jedoch haben wir täglich mit solchen klinischen Bildern zu tun (Kia, 2013).
- Asymmetrie ist das Kriterium, dem wohl am kritischsten zu begegnen ist. Asymmetrie kann völlig normal sein, denn kein Mensch hat gleich lange Beine oder symmetrische Gesichtshälften. Auch im Sport wie z.B. beim Fußball wäre es fatal, eine Symmetrie anzustreben, da ein Standbein und ein Schussbein eine sportartspezifische Asymmetrie darstellen können. Wer entscheidet also, ob eine „gute“ oder eine „schlechte“ Asymmetrie vorliegt (Robert, 2019).

Dies sind nur einige Beispiele von vielen, die klarmachen sollen, wie wichtig es ist, gelehrt Konzepte auch kritisch zu hinterfragen und auf den Stand der aktuellen Wissenschaft zu bringen. Da die Osteopathie anstrebt, ein Teil der evidenzbasierten Medizin zu sein, muss sie

sich auch kritisch hinterfragen lassen, auch wenn dies bedeutet, lieb gewonnene Gewohnheiten abzuschütteln.

Der Kern der Forschungsfrage in der vorliegenden Arbeit beschäftigt sich damit, inwieweit die SD eine sinnvolle Hypothese ist, die weiterhin zur Diagnostik benutzt werden sollte. Auf der anderen Seite könnte die SD auch ein überholtes Konzept darstellen, das einer Aktualisierung oder einer Überarbeitung bedarf.

Eine fundierte wissenschaftliche Basis der osteopathischen Grundprinzipien bekommt somit auch für jede Patientin und jeden Patienten Relevanz, der sich für eine osteopathische Behandlung entscheidet. Eine einheitliche diagnostische Anwendung und Reproduzierbarkeit könnte ein wichtiger Schritt in Richtung weiterer wissenschaftlicher Anerkennung sein. Die Reproduzierbarkeit und Reliabilität von osteopathischen Tests zur Auffindung der SD könnten somit einen entscheidenden Weg für eine bessere evidenzbasierte Praxis leisten und den Grundstein für weitere Forschung bilden.

2. Theoretischer Hintergrund

2.1. Erklärungsmodelle der somatischen Dysfunktion

In diesem Kapitel werden die theoretischen Erklärungsmodelle der SD näher betrachtet. Beschrieben werden Faktoren, die möglicherweise zu einer SD führen können und wie sich diese als klinische TART Zeichen widerspiegeln könnten.

Aus didaktischen Gründen werden die Faktoren, die eine SD hervorrufen, in mechanische (die folgend als Gewebefaktoren angeführt werden) und in neurologische Faktoren eingeteilt (Fryer, 2016). Dem Autor ist dabei sehr wohl bewusst, dass diese Einteilung eine rein theoretische ist und diese nur zum besseren Verständnis dient.

Gewebefaktoren, die zu klinischen Symptomen der SD beitragen können (Fryer, 2003):

- Verletzungen und Inflammation der Facettengelenke
- Einklemmung von meniskoiden Strukturen in und um das Facettengelenk
- Periartikuläres Gewebe wie Muskeln, Ligamente und Gelenkkapseln
- Derangement oder Degeneration der Bandscheiben

Neurologische Faktoren, auch als nozizeptiv-gesteuerte Veränderungen bezeichnet (Fryer, 2016):

- Fazilitiertes Segment
- Zentrale Sensibilisierung

Die oben angeführten Faktoren sind in der **Abbildung 1** nochmals zusammengefasst und werden im Anschluss detailliert beschrieben.

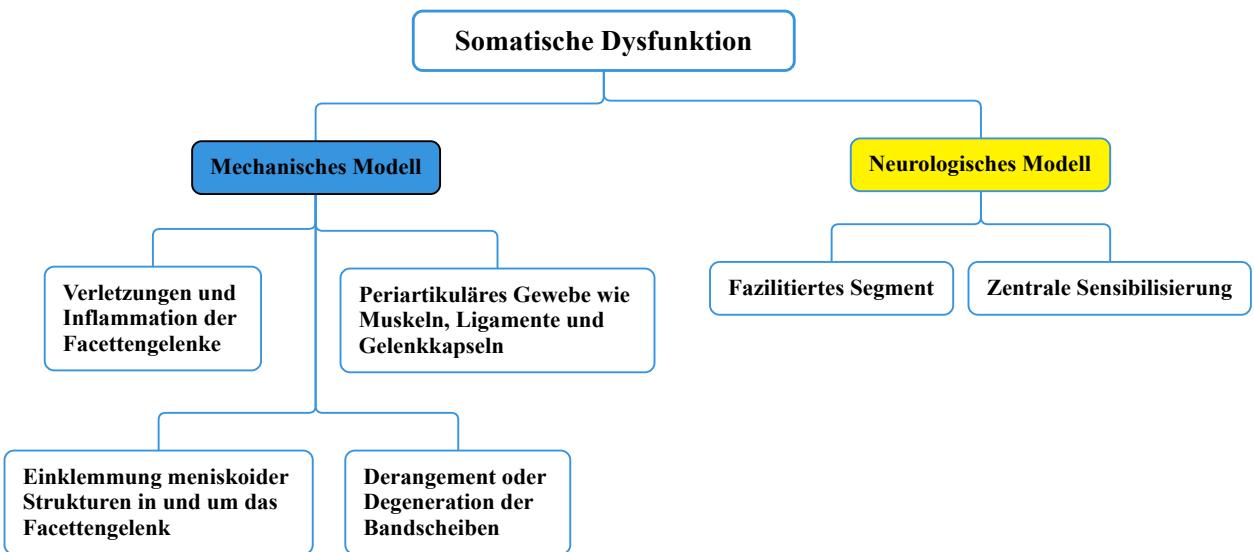


Abbildung 1: Faktoren der somatischen Dysfunktion (eigene Darstellung)

2.2. Gewebefaktoren als Erklärungsmodell

2.2.1. Verletzungen und Inflammation der Facettengelenke

Nach einem Trauma jeglicher Art kann es zu einer Verstauchung der Gelenkkapsel vom Facettengelenk kommen, eine mögliche Folge wäre eine intervertebrale Dysfunktion mit Schmerzen (Bogduk, 2005). Eine posttraumatische Entzündungsreaktion mit anschließendem Gelenkerguss kann sich in den klinischen Zeichen einer SD bemerkbar machen. Es können dadurch die Kardinalsymptome Schmerz und veränderte Sensibilität der kapsuloligamentären Strukturen, eingeschränkter Bewegungsumfang und veränderte Gewebebeschaffenheit durch die vorangegangene Entzündungsreaktion des Gewebes auftreten (Fryer, 2016).

Gegen solche Hypothesen sprechen unter anderem Untersuchungen von Nazarian et al. (1998), Fryer und Adams (2011), in denen keine Korrelation zwischen symptomatischen Patienten und Inflammation hergestellt werden konnte. Bei der Untersuchung von Nazarian wurden Facettengelenke von symptomatischen Halswirbelsäulen- (HWS) und Lendenwirbelsäulenpatienten (LWS) auf Entzündungszeichen mittels Ultraschalls untersucht. Es konnten keine Inflammationszeichen in oder um die Facettengelenke gefunden werden. In der Studie von Fryer und Adams wurden fünf akute HWS Patienten mit Schmerz auf einer Seite im Nacken mittels Magnetresonanztomographie auf Entzündungszeichen hin untersucht. Auch dort wurden keine Entzündung- oder Schwellungszeichen erkannt. Natürlich besteht die

Möglichkeit, dass die Untersuchungsmethoden für eine verdeckte oder sehr subtile Entzündung nicht sensitiv genug sind (Fryer, 2016). Andere Autoren geben wiederum an, dass es durch externe Krafteinwirkung zur Verletzung des Gelenkknorpels mit anschließender Entzündungsreaktion kommen kann (Lewit, 2012).

2.2.2. Einklemmung von meniskoiden Strukturen in und um das Facettengelenk

Die Einklemmung von meniskoiden Strukturen oder auch Kapselfalten wird zum Teil als Auslöser einer Entzündungsreaktion mit folgenden akuten Schmerzen diskutiert (Webb, Collins, Rassoulian, & Mitchell, 2011).

Ein in Flexion blockiertes Facettengelenk mit Einklemmung innerhalb oder außerhalb einer Kapselfalte wäre eine mögliche Erklärung für eine akute Rückenschmerzproblematik. Diese Hypothese scheint auf den ersten Blick eine schlüssige Möglichkeit zu sein, der es aber an notwendiger Evidenz mangelt (Bogduk, 2005; Webb et al., 2011).

2.2.3. Periartikuläres Gewebe wie Muskeln, Ligamente und Gelenkkapseln

Die Gründe für eine mögliche veränderte segmentale Beweglichkeit scheinen vielfältig zu sein. Fibrosierungen der kapsuloligamentären Strukturen nach Verletzungen wären nur eine der denkbaren Ursachen (Frank, Hart, & Shrive, 1999; Jung, Fisher, & Woo, 2009).

Eine andere Möglichkeit wäre eine segmentale Hypermobilität (Steilen, Hauser, Woldin, & Sawyer, 2014), die in der klassischen Theorie der SD nur von wenigen Autoren bedacht wird (DeStefano, 2011). In den üblichen Erklärungen zu den TART-Kriterien und ihrer Diagnose wird eine Hypermobilität überhaupt nicht in Erwägung gezogen. Durch viskoelastische Phänomene oder Verletzungen kann es zu Langzeitveränderungen am Kapsel-Bandapparat der Facettengelenke und somit zu einer vermehrten Beweglichkeit kommen (Steilen et al., 2014).

Sowohl für segmentale Hypermobilität nach einem Cervicalen-Beschleunigungstrauma (Steilen et al., 2014) als auch für Mindermobilitäten nach Immobilisation oder Verletzungen mit anschließender intraartikulärer Adhäsion gibt es ausreichende Evidenz in der Literatur (Cramer, Henderson, Little, Daley, & Grieve, 2010).

2.2.4. Derangement oder Degeneration der Bandscheiben

Woher kommt der Schmerz, dies war die Frage in einem Artikel von Mooney (1987). Jahre später war die Antwort, dass die Ursache von chronischen Schmerzen unter anderem in den intervertebralen Bandscheiben liegen kann (Bogduk, 2005; Bogduk, Aprill, & Derby, 2013).

Aktuellere Studien hingegen zeigen, dass degenerative Veränderungen etwas ganz Natürliches sind und auch bei Personen ohne Schmerzen gefunden werden können (Brinjikji et al., 2015; Kushchayev et al., 2018). Die Bandscheibendegeneration tritt häufig als Teil des Alterungsprozesses und bei asymptomatischen Menschen auf, auch wenn die vermeintlich „geschädigten“ Gewebe in der Bildgebung dargestellt werden konnten (Brinjikji et al., 2015).

Eine Bandscheibendegeneration kann sich als klinisches Zeichen in einer eingeschränkten segmentalen Beweglichkeit in allen Richtungen zeigen (Colloca, Keller, Moore, Gunzburg, & Harrison, 2007; Ellingson, Shaw, Giambini, & An, 2016). Eine palpable Veränderung der paarvertebralen Gewebsbeschaffenheit, ausgelöst durch eine reflektorische Kontraktion der stabilisierenden Muskulatur, wäre eine durchaus denkbare Theorie (Indahl, Kaigle, Reikerås, & Holm, 1995).

Eine Bandscheibendegeneration hätte dadurch also das Potential, die Kardinalsymptome der SD auszulösen. Vor allem wären die eingeschränkte segmentale Beweglichkeit und damit die „range of motion“ betroffen (Fryer, 2016).

2.3. Notwendige neuroanatomische Grundlagen

Da das Nervensystem ein sehr komplexes Gebilde darstellt, sollen im folgenden Kapitel die notwendigsten neuroanatomisch- und physiologischen Grundlagen beschrieben werden. Um den Rahmen der Arbeit nicht zu sprengen, beschränkt sich der Autor nur auf das, was zum besseren Verständnis des neurologischen Modells in der Osteopathie nötig ist.

2.3.1. Afferente Rezeptoren und Nervenfasereinteilung

Gelenkrezeptoren wurden von Freeman & Wyke (1967) in Typ I - IV eingeteilt. Diese dienen der Informationsaufnahme und werden daher als afferent bezeichnet. Die meisten Rezeptoren befinden sich im Weichteilmantel eines Gelenks, an den anliegenden Muskeln und an den Insertionen der Gelenkkapseln (Wolff, 2013). Laut Wolff sind Propriozeptoren vom Typ I - III

Mechanorezeptoren und haben Stellungs- und Bewegungsmeldefunktion über physiologische Abläufe im Gelenk. Die Typ IV-Rezeptoren hingegen sind Nozizeptoren und wachen über die Intaktheit des Systems.

Die Nervenfasereinteilung erfolgt laut dem Standardlehrbuch der Physiologie (Pape, Kurtz, & Silbernagl, 2018) in zwei Kategorien und wird in den **Tabellen 1 und 2** dargestellt. Sie unterscheiden sich hauptsächlich in Leitungsgeschwindigkeit, Faserdurchmesser und Funktion. Die Einteilung von Lloyd und Hunt wird zur Beschreibung von Afferenzen aus Muskulatur, Gelenken, Faszien usw. benutzt, bei der von Erlanger und Gasser sind neben Afferenzen auch noch efferente Fasern enthalten (Böhni, Lauper, & Locher, 2015).

Hervorzuheben sind sogenannte nozizeptive Fasern, die von klinisch wichtiger Bedeutung sind:

- **A δ -Fasern** (dünn myelinisiert, schnell leitend 15 m/s) (Böhni et al., 2015)
- **C - Fasern** (nicht myelinisiert, langsam leitend 1 m/s) (Böhni et al., 2015)

Tabelle 1: Nervenfasereinteilung nach Lloyd und Hunt (Pape et al., 2018)

Nervenfasereinteilung nach Lloyd und Hunt

Fasertyp	Faserdicke μm	Geschwindigkeit m/s	Funktion
Ia/IIb	13	75	Ia: Muskelspindelafferzen IIb: Golgi-Sehnenorgan (Propriozeption)
II	9	55	Mechanorezeptoren der Haut, Spindelafferzen, Lagesinn, Gelenkkapsel
III	3	11	Fasern für Drucksensibilität
IV	1	1	Marklose nozizeptive Fasern

Tabelle 2: Nervenfasereinteilung nach Erlanger und Gasser (Pape et al., 2018)

Nervenfasereinteilung nach Erlanger und Gasser			
Fasertyp	Faserdicke μm	Geschwindigkeit m/s	Funktion
Aα	15	100	motorisch zur Skelettmuskulatur, Muskelspindelafferzenzen
Aβ	8	50	Hautafferzenzen für Berührung und Druck, Propriozeption Gelenke
Aγ	5	20	Motorische Efferenz zu den Muskelspindeln
Aδ	3	15	Hautafferenz für Thermo- und Nozizeption (heller Schmerz)
B	3	7	sympathische präganglionäre Fasern
C	1	1	Dumpfe Nozizeption und Thermozeption der Haut sowie sympathische postganglionäre Fasern

2.3.2. Faserkategorien peripherer Nervenbahnen

Im peripheren Nervensystem (PNS) leiten afferente-sensible Bahnen Informationen zum zentralen Nervensystem (ZNS) und efferent-motorische Bahnen Impulse zu den Erfolgsorganen (Trepel & Dalkowski, 2011). Man kann sieben Faserkategorien unterteilen, wobei der Autor in der vorliegenden Arbeit die allgemeinen und speziellen Faserkategorien zur Vereinfachung zusammenfasst hat. Somit entstehen vier große Gruppen von Faserkategorien die von Trepel and Dalkowski (2011) in ihrem Standardwerk zur Neuroanatomie wie folgt beschrieben werden:

- **Somatomotorische (=somatoefferente) Fasern**

Versorgen ausschließlich die Skelettmuskulatur.

- **Somatosensible (=allgemein und speziell somatoafferente) Fasern**

Vermitteln Impulse aus der Haut, den Schleimhäuten, den Muskelspindeln und den Rezeptoren in Gelenkkapseln und Sehnen.

- **Visceromotorische** (allgemein und speziell visceroefferente) **Fasern**

Versorgen die glatte Gefäß- und Eingeweidemuskulatur, das Herz und die Drüsen. Enthalten sympathische und parasympathische Nervenfasern.

- **Viscerosensible** (allgemein und speziell visceroafferente) **Fasern**

Leiten Informationen von den inneren Organen und dem Gefäßsystem zum ZNS.

Der Somatosensibilität unterliegt eine weitere Einteilung in Exterozeption (Oberflächensensibilität), wo der Reizort die Haut, die Schleimhäute von Mund und Nase und die Augenoberfläche ist, sowie eine Propriozeption (Tiefensensibilität), wo sich der Reizort in Muskel, Sehne oder Gelenkkapsel befindet. Die Reizart der Exterozeption wird zusätzlich noch in eine epikritische (Tastsinn, Vibration, leichte Berührung und leichter Druck) und eine sogenannte protopathische Wahrnehmung (Schmerz, Temperatur und grobe mechanische Reize) untergliedert. Das erste somatosensible Neuron befindet sich im Spinalganglion und somit im PNS. Die protopathischen Bahnen schalten auf ein zweites Neuron im Hinterhorn um, kreuzen segmental und ziehen im Tractus spinothalamicus nach cranial zum ZNS weiter. Die epikritische Sensibilität wird in der Hinterstrangbahn (Tractus gracilis und Tractus cuneatus) ungekreuzt ins ZNS geleitet und kreuzt erst auf Höhe Hirnstamm zur kontralateralen Seite. Das zweite Neuron der Proprioception befindet sich ebenfalls im Rückenmark und zieht im Tractus spinocerebellaris ungekreuzt zum Kleinhirn, lediglich das Lageempfinden kreuzt im Bereich des Hirnstamms zur kontralateralen Seite. Die somatomotorische Bahn des ZNS hat ein am Vorderhorn des Rückenmarks endendes Motoneuron, dies ist der Signalgeber für die Muskulatur. Dort werden zwei weitere efferente Fasern unterschieden. Zum einen die $A\alpha$ -Fasern, welche die Bewegungen eines Muskels auslösen und die $A\gamma$ -Fasern, die auch unabhängig von einer konkreten Bewegung eine „Grundspannung“ der Muskulatur regulieren (Schulte, Schumacher, & Schünke, 2018).

Die visceromotorischen Fasern des Sympathikus haben drei Möglichkeiten zur Umschaltung auf das zweite Neuron. Das erste Neuron befindet sich im Seitenhorn des Rückenmarks von C8-L2.

1. Möglichkeit: Sympathische Fasern für die Blutgefäße der Extremitäten und Rumpfwand, sowie der Haut werden im Grenzstrangganglion umgeschaltet und ziehen im Spinalnerv weiter in ihr Innervationsgebiet.

2. Möglichkeit: Efferente sympathische Fasern der Viscera ziehen durch Grenzstrangganglien hindurch, werden in Ganglien weiter umgeschaltet, um dann ein Organ direkt zu innervieren.

3. Möglichkeit: Sympathische Fasern werden im Organ selbst umgeschaltet.

Parasympathische Zentren liegen in den Hirnnervenkernen und im Sakralmark und werden in organnahen Ganglien oder in den Organen selbst direkt umgeschaltet. Visceroafferente Fasern laufen auch immer in sympathischen und parasympathischen Nervenfasern mit (Schulte et al., 2018).

2.3.3. Organisation im Hinterhorn und Rückenmark

Der vordere Flügel der schmetterlingsähnlichen grauen Substanz im Rückenmark wird auch als Vorderhorn bezeichnet. Dieser Teil enthält vorwiegend motorische Kerne, auch Motoneurone genannt. Der hintere schmalere Teil des Schmetterlingsflügels wird als Hinterhorn bezeichnet und steht mehr im Dienst der Sensibilität. Wie weiter oben schon beschrieben, ist im Bereich von C8-L3 noch ein Seitenhorn zu finden, das Neuronengruppen des vegetativen Nervensystems enthält. Zwischen den Neuronen liegen sogenannte Interneurone, kleine zwischengeschaltete Neurone inmitten zweier anderer Neuronen (Trepel & Dalkowski, 2011). Die graue Substanz lässt sich weiter in unterschiedlich aufgebaute Zellschichten einteilen, diese werden Laminae (I-X) nach Rexed genannt. Eine weitere Einteilungsmöglichkeit wäre nach Form und Lage in Zellgruppen oder Kerngebiete. Die Schichteinteilung ist im Hinterhorn besonders ausgeprägt (Schulte et al., 2018). Die beiden Einteilungen werden in **Abbildung 2** dargestellt.

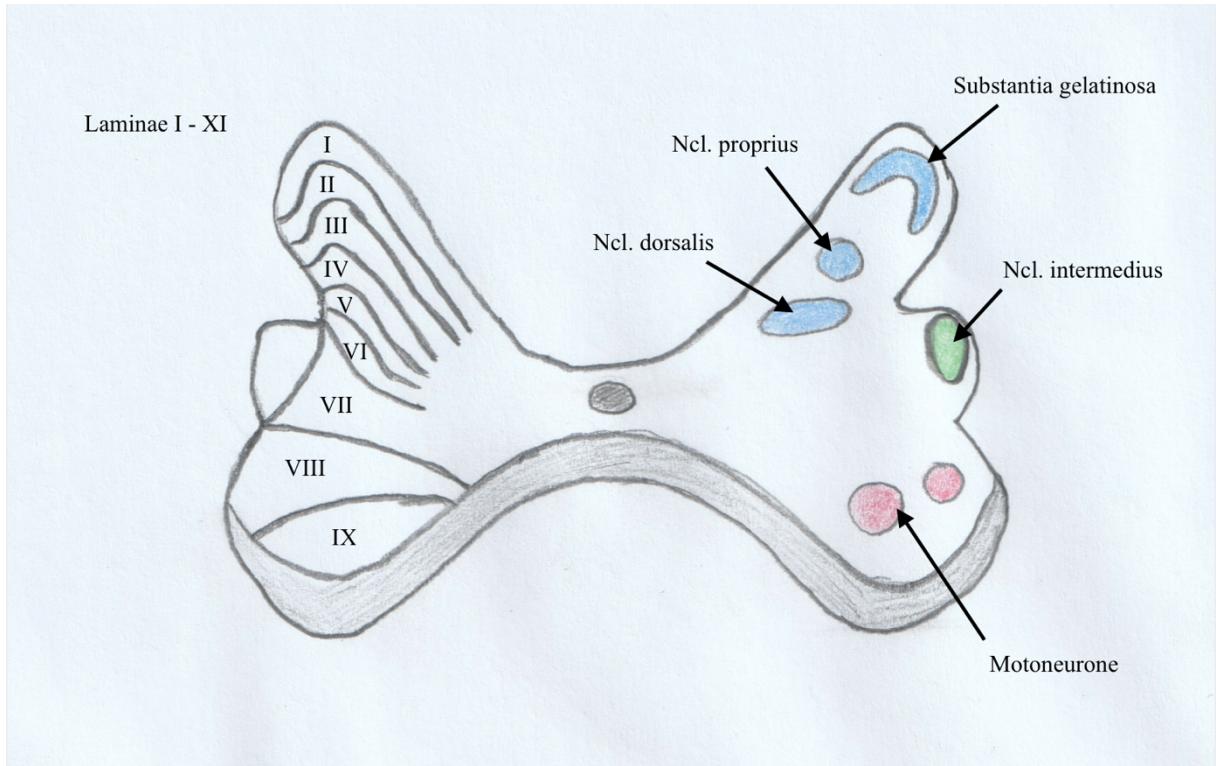


Abbildung 2: Graue Substanz im Rückenmark (eigene Darstellung)

Das Hinterhorn besteht aus Laminae I-VII, wo ein Großteil der sensiblen Fasern umgeschaltet werden. In der Substantia gelatinosa, die in der Lamina I-II zu finden ist, sowie in der Lamina VII enden nozizeptive Afferenzen aus A_δ und C Fasern, die nach Umschaltung auf das 2. Neuron über den Tractus spinothalamicus zum Thalamus geleitet werden. Hier enden auch absteigende Bahnen aus den Raphekernen und dem Ncl. caeruleus (beides Teile der Formatio reticularis). Diese können durch Neurotransmitter wie Serotonin, Noradrenalin und GABA die Übertragung von nozizeptiven Impulsen hemmen. Dies geschieht zum Teil mithilfe von inhibitorischen Interneuronen (Trepel & Dalkowski, 2011).

In Lamina III und IV liegt der Ncl. proprius und bezieht propriozeptive Afferenzen aus dem Bewegungsapparat. Das wichtigste Kerngebiet von Lamina V und VI ist der Ncl. Dorsalis, der nur thorakolumbal zu finden ist. Er empfängt ebenfalls Informationen zur Tiefensensibilität (Trepel & Dalkowski, 2011).

Neben den oben genannten Kerngebieten sind in der Lamina V noch sogenannte Wide-Dynamic-Range-Neurone zu finden (WDR-Neurone). Diese multirezeptiven Neurone sammeln Reize aller Art wie Nozizeption, Somatoafferzen, Visceroafferzen, aber auch Informationen aus deszendierenden Bahnsystemen. Dies wird auch als multirezeptive Konvergenz des WDR-Neurons bezeichnet (Böhni et al., 2015).

Wenn diese WDR-Neurone durch Nozizeptoren aus der Peripherie wie Haut, Muskeln, Gelenke, Gefäße, Meningen oder der Viscera gereizt werden, kann es zum Phänomen der Konvergenz kommen. Die Erregbarkeit der WDR-Neurone kann sich dadurch erheblich steigern, wodurch Signale, die vorher nur als unterschwellig galten, ein Aktionspotential hervorrufen und somit zur Erregung des Neurons führen. Dies kann über Axonkollateralen, also Verbindungen zu Motoneuronen im Vorderhorn und zu Ursprungsneuronen des Sympathikus im Seitenhorn, zur motorischen und vegetativen Systemaktivierung führen (Böhni et al., 2015). Die **Abbildung 3** zeigt das WDR-Neuron mit seinen vielzähligen Afferenzen.

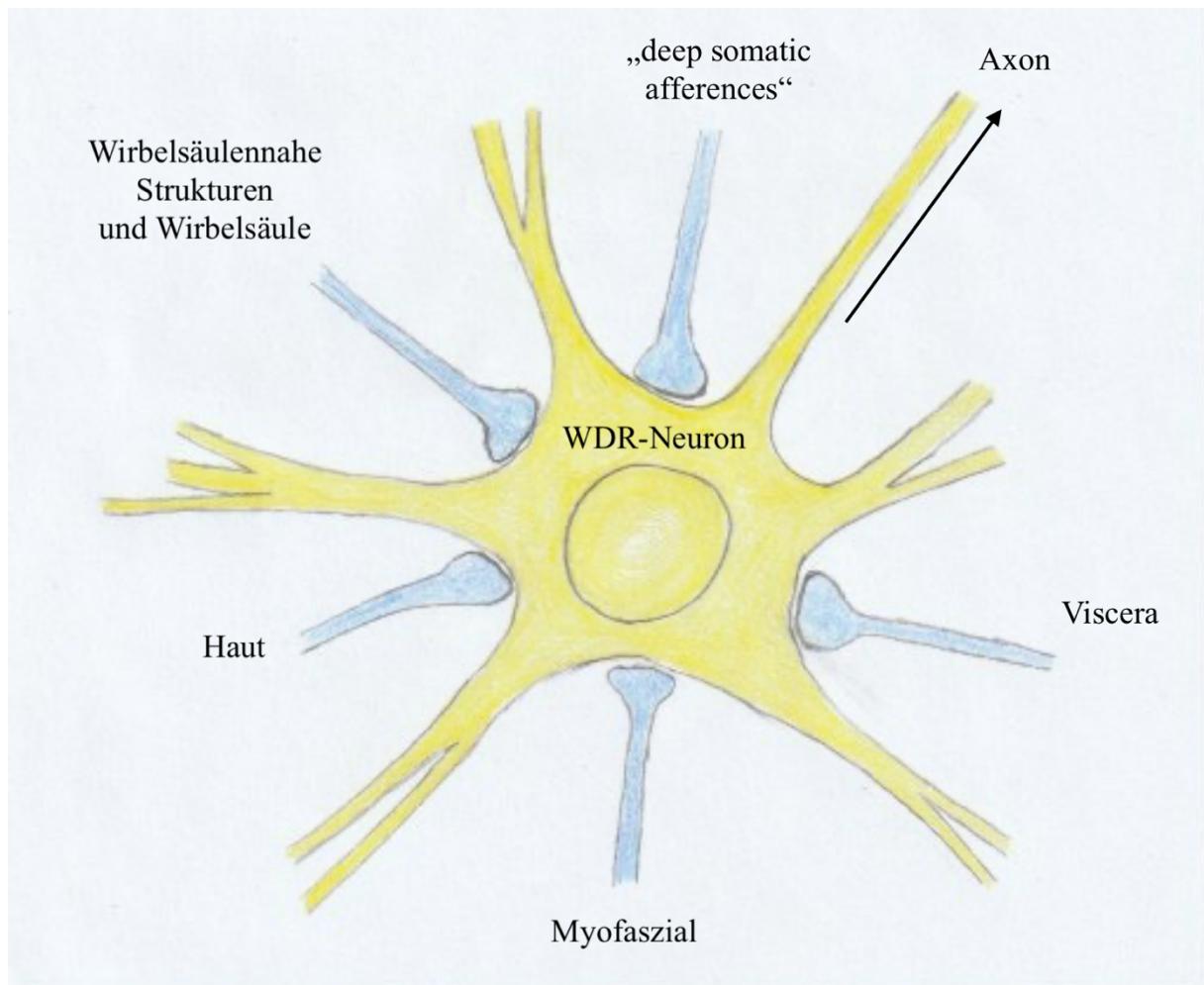


Abbildung 3: WDR-Neuron (eigene Darstellung)

2.4. Neurologisches Modell der somatischen Dysfunktion

In der Anfangszeit betrachtete man eine osteopathische Läsion eher als ein mechanisches Problem. Erst später wurden von Korr Irvin, Ph.D. und Denslow, D.O. neurophysiologische Erklärungsmodelle in Betracht gezogen. So entstand das Modell des fazilitierten Segments, das in der täglichen osteopathischen Praxis seine Anwendung findet (Liem, 2018).

2.4.1. Das fazilitierte Segment

Vereinfacht dargestellt kommt es beim Modell des fazilitierten Segments zu einer übertriebenen Reaktion auf segmentaler Ebene im Rückenmark. Es gibt zwei Input-Routen in das Hinterhorn des Rückenmarks. Dies wären zum einen somatosensible-afferente Informationen aus dem muskuloskeletalalen System (Gelenke, Muskeln, Sehnen, Haut, Faszien), zum anderen viscerosensible-afferente Informationen aus den inneren Organen. Zwei Output-Routen werden ebenfalls beschrieben. Zum einen die Somatomotorik zur Skelettmuskulatur als efferente Bahn, zum anderen Visceroefferenzen zu den inneren Organen, zu den Blutgefäßen, den Schweißdrüsen und den pilomotorisch arbeitenden Muskeln der Haut (Lederman, 2000). Die Input- und Output Routen werden in der **Abbildung 4** zum besseren Verständnis bildlich dargestellt. Korr und Denslow waren der Annahme, dass durch eine erhöhte Aktivität aus einem der Input-Signale, wie z.B. einer myofaszialen Verletzung, eine übertriebene Reaktion in den Output-Signalen entstehen könnte. Dies könnte sich in einer erhöhten motorischen und sympathischen Reaktion zeigen, die als Fazilitation auf segmentalerer Ebene bezeichnet wird (Denslow, 1953; Fryer, 2016; Korr, 1947; Lederman, 2000). Eine Fazilitation der Interneurone verringert deren Reizschwelle, was zu einer übertriebenen Reizreaktion auf segmentaler Ebene führen könnte. Die Folge aus der erhöhten Aktivität der efferenten Bahnen, wäre ein erhöhter Tonus der segmental innervierten Muskulatur und Veränderungen in der Vaso-, Sudo- und Pilomotorik der sympathisch innervierten Strukturen (Lederman, 2000).

Es erfolgte 1990 durch Van Buskirk (1990) eine Erweiterung des ursprünglichen Konzepts durch neuere Erkenntnisse in Bezug auf nozizeptive Reize und deren Reflexantwort. Die Nozizeptoren nehmen dadurch eine größere Rolle in der Entstehung einer somatischen Dysfunktion ein und führen zu motorischen und sensiblen Reflexantworten. Dem Konzept des fazilitierten Segments stehen jedoch auch einige Kritiken im Wege. So gibt Lederman (2000) in einem Artikel einige Probleme des Modells wie folgt an:

- Absteigende Bahnen und somit deszendierende Einflüsse aus höheren Zentren finden keine Erwähnung in dem Konzept. Dies ist unrealistisch, da die segmentalen Zentren nicht isoliert arbeiten, sondern unter Einfluss von höheren Zentren stehen. Motorische sowohl auch autonome Funktionen werden vom Gehirn gesteuert und unterliegen somit auch dessen Einflüssen.
- Muskeln beziehen ihre Innervation aus verschiedenen segmentalen Höhen, dabei sind sehr viele Motoneurone und Interneurone daran beteiligt. Somit entsteht auch keine segmentale Antwort, sondern eher eine plurisegmentale. Dies widerspricht im Grunde dem Konzept des

fazilitierten Segments. Auch Carreiro (2011) gibt an, dass eine Segmentierung nur außerhalb des Rückenmarks stattfindet. Innerhalb des Rückenmarks geht die exakte anatomische Trennung verloren.

- Laut dem Modell müsste jede Verletzung eines Muskels oder einer Muskelgruppe auch zu Dysfunktionen in den inneren Organen führen. Dafür gibt es derzeit keine Hinweise, weder wissenschaftliche noch empirische.

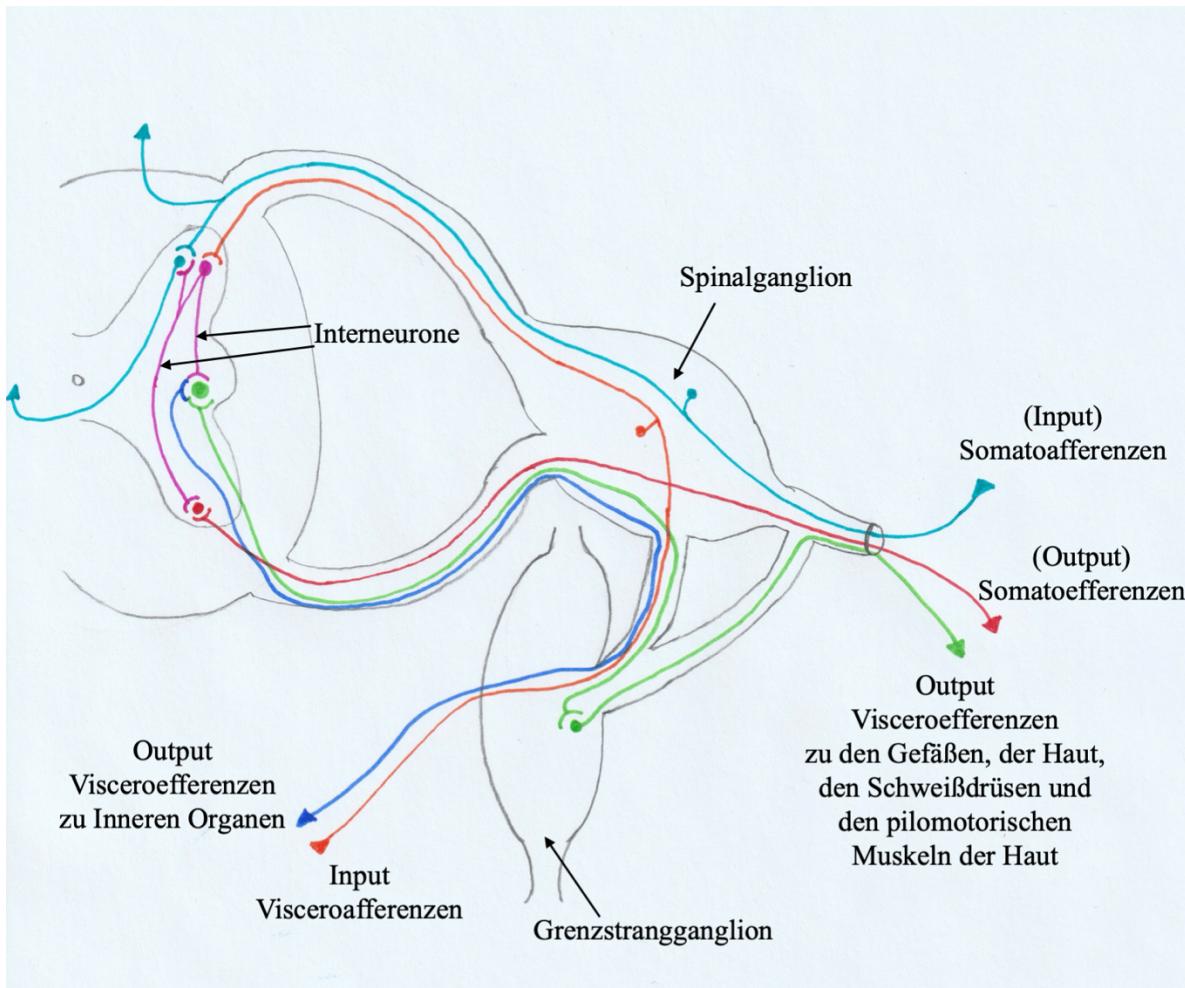


Abbildung 4: Somatoafferzen, Visceroafferzen, Somatoefferzen und Visceroefferzen (eigene Darstellung)

2.4.2. Zentrale Sensibilisierung

Das Konzept des fazilitierten Rückenmarksegments teilt sich einige Gemeinsamkeiten mit dem aktuelleren Konzept der zentralen Sensibilisierung und ergänzt sich damit. Vermehrte afferente Reize, die Neurone im Hinterhorn beeinflussen, sowie die Bedeutung von nozizeptiven Signalen sind Teil des Modells (Fryer, 2016). Die individuelle Schmerzwahrnehmung eines Patienten und der Einfluss von höheren Zentren finden in diesem Konzept ebenfalls ihre Beachtung (Wolff, 2013).

Durch anhaltende nozizeptive Signale aus der Peripherie kommt es zu Veränderungen der Informationsverarbeitung in den spinalen nozizeptiven Systemen. Es sind sensibilisierte afferente Neurone, erregende und hemmende Interneurone, wie auch erregende und hemmende supraspinale Kontrollsysteme des Hirnstamms daran beteiligt (Böhni et al., 2015). Eine Sensibilisierung von Nozizeptoren, ausgelöst durch zum Beispiel eine Inflammation, erzeugt eine verstärkte Erregung zentraler Neurone im Hinterhorn und eine plastische Veränderung zentraler Informationsverarbeitungen (Latremoliere & Woolf, 2009).

Die zentralen Sensibilisierungsvorgänge im Hinterhorn (Laminae I, II, und V) weisen laut Böhni et al. (2015) folgende Merkmale auf:

- Erhöhte Spontanaktivität
- Erniedrigte Erregungsschwelle für mechanische und/oder thermische Reize (Hyperalgesie)
- Erhöhte Intensität und verlängerte Dauer von Entladungen als gesteigerte Antwort auf noxische Reize
- Vergrößerung und Neubildung von rezeptiven Feldern
- Auslösung von Schmerzen durch Reize, die eigentlich nicht als schmerhaft empfunden werden (Allodynie)

Die Hyperalgesie, bei der schmerzhafte Reize als viel deutlicher empfunden werden, sowie die Allodynie, wo nicht schmerzhafte Reize als Schmerz empfunden werden (wie leichte Berührung oder Bewegungen), könnten also den klinischen Befund von „tenderness“ bei einer somatischen Dysfunktion erklären. Auch dann, wenn die Verletzung oder Störung in der Peripherie schon ausgeheilt oder nicht mehr vorhanden ist (Fryer, 2016).

Durch eine Reizung von afferenten Nozizeptoren kommt es zur lokalen Inflammation des Gewebes, ausgelöst durch eine neurogene Entzündungsreaktion. Dabei werden proinflammatorische Neuropeptide wie Substanz P oder calcitonin-gene-related-peptide (CGRP) aus peripheren Nervenendigungen freigesetzt (Herbert & Holzer, 2002). Diese Neuropeptide kommen ursprünglich aus dem Spinalganglion und werden über Nervenfasern in die Peripherie entsandt, um dort bei Bedarf eine Entzündungskaskade in Gang zu setzen (Willis, 1999). Dieses Phänomen, welches auch als Hinterhorn-Reflex bezeichnet wird, lässt sich nicht nur durch lokale Faktoren in der Peripherie auslösen, sondern auch durch Stimulationen im periaquäduktalem Grau, einem Teil des Mittelhirns. Somit haben höhere Zentren über

absteigende Bahnen auch die Möglichkeit, diesen Hinterhorn-Reflex auszulösen (Peng, Kenshalo, & Gracely, 2003). Durch die aktivierten Nozizeptoren und die dadurch ausgelöste neurogene Entzündung, könnte die in der SD wichtige Gewebeveränderung („tissue texture changes“) eine ursächliche Erklärung finden (Fryer, 1999).

Neuropeptide wie Substanz P und CGRP werden jedoch nicht nur in der Peripherie freigesetzt, sondern auch auf Hinterhornebene. Dort führen sie zur Schwellenänderung der WDR-Neurone, die ihre Selektivität verlieren und ihre rezeptiven Felder vergrößern können (Böhni et al., 2015). Eine langfristige Folge aus diesen Prozessen wäre eine Sensibilisierung auf Hinterhornebene.

Die zentrale Sensibilisierung und die neurogene Entzündung als Teil dieser wären somit eine mögliche Erklärung für „tenderness“ und „tissue texture changes“ als diagnostisches Kriterium der SD (Fryer, 2016).

Es sollte jedoch nicht vergessen werden, dass die nozizeptive Verarbeitung im Hinterhorn unter erregender und hemmender supraspinaler Kontrolle liegt. Absteigende Kontrolle über Transmitter wie Noradrenalin, Serotonin, Glutamat und Opioide kann eine Hemmung oder Förderung der nozizeptiven Impulsübertragung beeinflussen (Böhni et al., 2015).

2.5. Diagnostik der somatischen Dysfunktion

Bei der Identifikation der somatischen Dysfunktion, orientieren sich OsteopathInnen an den diagnostischen TART Kriterien (Liem & Dobler, 2017). Zur Erkennung einer Asymmetrie werden eher inspektorische Tests und Stellungstests von anatomischen Landmarken benutzt. Auch werden kombinierte universelle Muster wie eine Pronationsstellung des rechten Fußes, Iliumrotation nach anterior rechts, Iliumrotation nach posterior links, Pubis inferior rechts, Pubis superior links, Sakrumtorsion nach anterior über die linke diagonale Achse, Skoliose der unteren BWS mit Konvexität nach rechts, Skoliose der LWS mit Konvexität nach links, Schultertiefstand rechts und eine Protraktion der rechten Schulter mit skoliotischer Stellung der HWS mit Konvexität links beschrieben (DeStefano, 2011). Greenman gibt in seinem Lehrbuch der osteopathischen Medizin an, dass ein Abweichen bzw. das Fehlen dieser Asymmetrie häufiger klinische Symptome macht wie ein ausgeglichenes Asymmetrie-Muster (DeStefano, 2011). Der amerikanische Osteopath Gordon Zink beschrieb ein Torsionsmuster der Übergangszonen zwischen Cranium und Halswirbelsäule (= occipitoatlantal Übergang), zwischen HWS und Brustwirbelsäule (= cervikothorakaler Übergang), zwischen BWS und LWS (= thorakolumbaler Übergang) und zwischen LWS und Sakrum (= lumbosakraler

Übergang). Er beschrieb ein allgemeines Kompensationsmuster, das anzustreben sei und eine Abweichung dessen was bei symptomatischen Patienten zu finden ist (Robert, 2019). Auch hier findet die Asymmetrie ihren Platz, wird aber als normal angesehen. Die Diagnostik erfolgt eher durch Palpation und Bewegungsuntersuchung.

Die veränderte Gewebeempfindlichkeit oder „tenderness“ erfolgt durch Palpation und Provokationstests, wobei die individuelle Schmerzwahrnehmung des Patienten dabei eine große Rolle spielt (Fryer, 2016). Eine Veränderung der Gewebebeschaffenheit von Haut, Faszien, Muskeln, Ligamenten etc. wird ebenfalls durch Palpation und Observation, aber auch durch das Endgefühl bei Bewegungstests evaluiert. Der Bewegungsumfang, ob hypo- oder hypermobil, wird durch Observation von aktiven Bewegungen und durch Palpation im Sinne eines Mobilitätstests untersucht (DeStefano, 2011). Die Palpation scheint daher eines der wichtigsten Instrumente für die Diagnostik in der Osteopathie zu sein. Auch Chaitow (2001) gibt in seinem Werk über Palpationstechniken und Diagnostik an, wie wichtig die Palpation als Grundlage der Befunderhebung ist.

3. Methodologie

3.1. Forschungsfrage

Diese Übersichtsarbeit, will die aktuelle Evidenzlage zur diagnostischen Reliabilität der SD überprüfen. Daraus ergibt sich folgende Forschungsfrage: Wie reliabel ist die Identifikation der somatischen Dysfunktion in der Osteopathie?

3.2. Vorannahme

Die Vorannahme dieser Arbeit besteht darin, dass eine Reliabilität für die Identifikation der SD besteht. Dies wird in dem vorliegenden Review auf ihre Richtigkeit überprüft.

3.3. Studien- und Forschungsdesign

Das Forschungsdesign der vorliegenden Studie entspricht einem systematischen Review und wird in **Abbildung 5** als Mind-Map dargestellt.

3.4. Einschlusskriterien

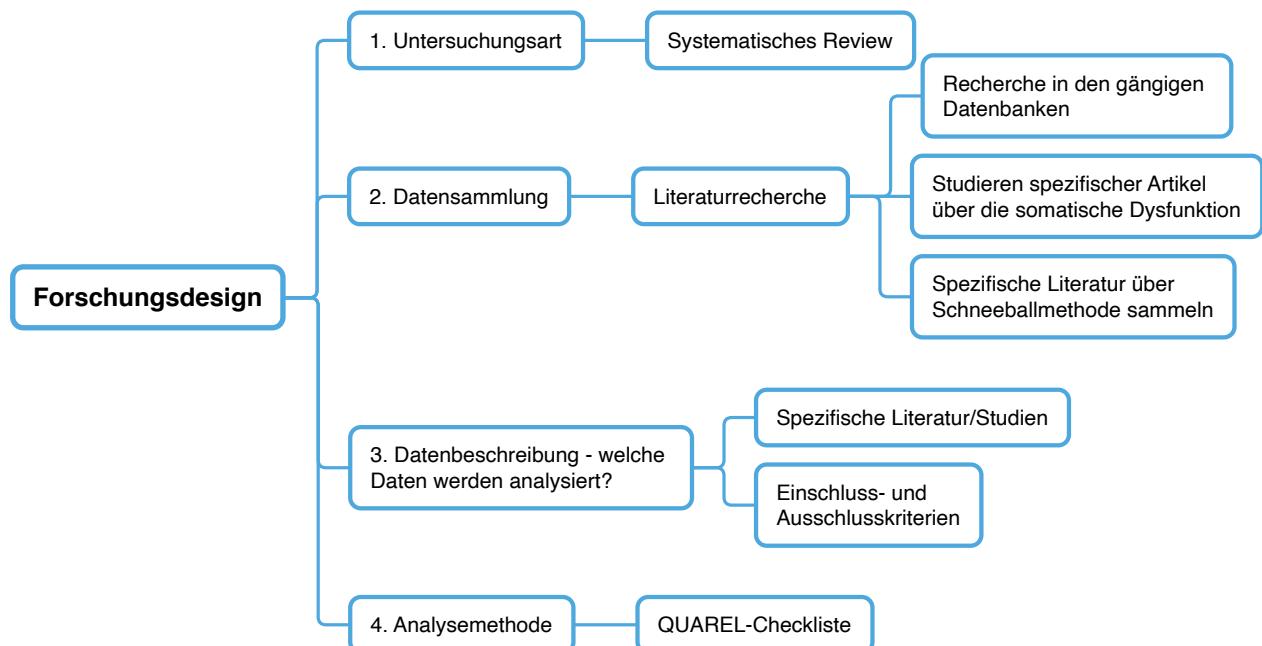


Abbildung 5: Mind-Map zum Forschungsdesign (eigene Darstellung)

Eine systematische Literatursuche wurde vom Autor von Januar 2020 bis Februar 2020 durchgeführt. Gesucht wurde in den gängigen Literaturdatenbanken, Journals sowie in aktueller Fachliteratur. Die Einschlusskriterien werden wie folgt aufgelistet:

- Artikel, die sich mit der Reliabilität der osteopathischen Untersuchung und Palpation sowie mit osteopathischen Tests auseinandersetzen, die zur Identifikation einer somatischen Dysfunktion zur Anwendung kommen.
- Die Art der Artikel beschränkte sich dabei auf intra- und interrater Reliabilitätsstudien.
- Artikel die einen entsprechenden wissenschaftlichen Aufbau aufweisen und daher auch mit einem Bewertungssystem beurteilt werden können.
- Die Sprache begrenzte sich bei der Recherche auf Studien in Englisch und Deutsch.

3.5.Ausschlusskriterien

- Arbeiten, die nicht den definierten Keywords entsprachen und daher für die vorliegende Arbeit nicht relevant sind.
- Studien, in denen die Untersuchung nicht von OsteopathiestudentInnen oder fertig ausgebildeten OsteopathInnen durchgeführt wurden.
- Studien, welche die Reliabilität von Tests überprüfen, die nicht den TART Kriterien entsprechen bzw. nicht zur Identifikation der somatischen Dysfunktion dienen.

3.6.Literatursuche

Die Literaturrecherche fand von Januar 2020 bis Februar 2020 statt. Gesucht wurde nach den vorher definierten Keywords: Reliabilität der osteopathischen Palpation und Untersuchung. Da die meisten Datenbanken und Journals in englischer Sprache vorlagen, wurden die Suchbegriffe auch auf Englisch verwendet. Um die Suche einzuschränken, wurden Boolesche Operatoren benutzt, die als wahrscheinlich wichtigstes Hilfsmittel bei der Literaturrecherche angesehen werden. Der Begriff OR als Operator zwischen "osteopathy" und „OMT“ und zwischen „examination“ und „palpation“ sollte Artikel ausfindig machen, die mindestens einen der beiden Begriffe beinhalten. Weiters wurde der Operator AND benutzt, um die Begriffe „osteopathy OR OMT“ und „examination OR palpation“ untereinander und mit dem Begriff

„reliability“ zu kombinieren. Der Operator AND soll nur zu Publikationen führen, in denen beide Begriffe vorkommen (Scherfer & Hempel, 2006).

Die Suchanfrage lautete daher wie folgt: Reliability AND (osteopathy OR OMT) AND (examination OR palpation).

Mit diesen Suchanfragen wurden folgende Datenbanken und Journals durchsucht:

- Pubmed
- IJOM (International Journal of Osteopathic Medicine)
- Ostepathic research web
- Cochrane Library
- PEDro
- JBMT (Journal of Bodywork and Movement Therapies)

Bei der Suche in Pubmed wurde zusätzlich der Filter „Clinical Trial“ eingesetzt, um ausschließlich Studien auf experimenteller Basis ausfindig zu machen. Damit konnten die Suchtreffer von 1944 auf 127 reduziert werden. Wenn möglich wurden auch Suchfilter in den anderen Datenbanken oder Journals verwendet, um ungewünschte Treffer auszugrenzen. Es wurde auch die hauseigene Bibliothek, bestehend aus osteopathischer und manualtherapeutischer Fachliteratur, sowie die Fachzeitschriften „Osteopathische Medizin“, „Manuelle Therapie“ und „D.O. Deutsche Zeitschrift für Osteopathie“ durchsucht. Des Weiteren wurden Artikel mit der Schneeballmethode ausfindig gemacht. Da die Suche in den Journalen mit den Booleschen Operatoren nicht konform geht und auch PEDro dieses Tool nicht erkennt, wurde eine erneute Suche mit einem modifizierten Suchbegriff auf diesen Internetseiten durchgeführt: Reliability of osteopathic diagnostic tests. Die Ergebnisse der Suche werden nun folgend in den **Tabellen 3-4** dargestellt:

Tabelle 3: Ergebnisse der Literatursuche

Suchbegriff: Reliabilität AND (osteopathy OR OMT) AND (examination OR palpation)	
Datenbank/Journal	Gefundene Artikel
Pubmed	127
IJOM (International Journal of Osteopathic Medicine)	7
Osteopathic research web	0
JBMT (Journal of Bodywork and Movement Therapies)	2
PEDro	0
Cochrane Library	0
Gesamt	136

Tabelle 4: Literatursuche

Suchbegriff: Reliability of osteopathic diagnostic tests	
Datenbank/Journal	Gefundene Artikel
IJOM (International Journal of Osteopathic Medicine)	96
JAOA (The Journal of the American Osteopathic Association)	61
Osteopathic research web	0
JBMT (Journal of Bodywork and Movement Therapies)	37
PEDro	0
Cochrane Library	0
Gesamt	194

Die Gesamtsuche ergab 330 Treffer sowie 3 zusätzliche Suchergebnisse über andere Suchmethoden wie die Schneeballmethode. Bei der Überprüfung nach Duplikaturen, konnten keine Artikel doppelt gefunden werden. Es wurden 298 Artikel nach dem Screening der Titel und der Zusammenfassungen von der Studie ausgeschlossen. Die übrigen 35 wurden nach dem Screening des Volltextes um weitere 15 Artikel reduziert. Somit blieben für die geplante Studie eine Gesamtzahl von 20 Artikeln übrig. Der Vorgang ist in einem Flow-Chart in **Abbildung 6**

dargestellt. Im Anschluss folgt eine Auflistung der 20 inkludierten Studien in **Tabelle 5**. Dort werden tabellarisch die Autoren mit Erscheinungsjahr und der Titel der Studien angegeben, um einen besseren Überblick zu wahren.

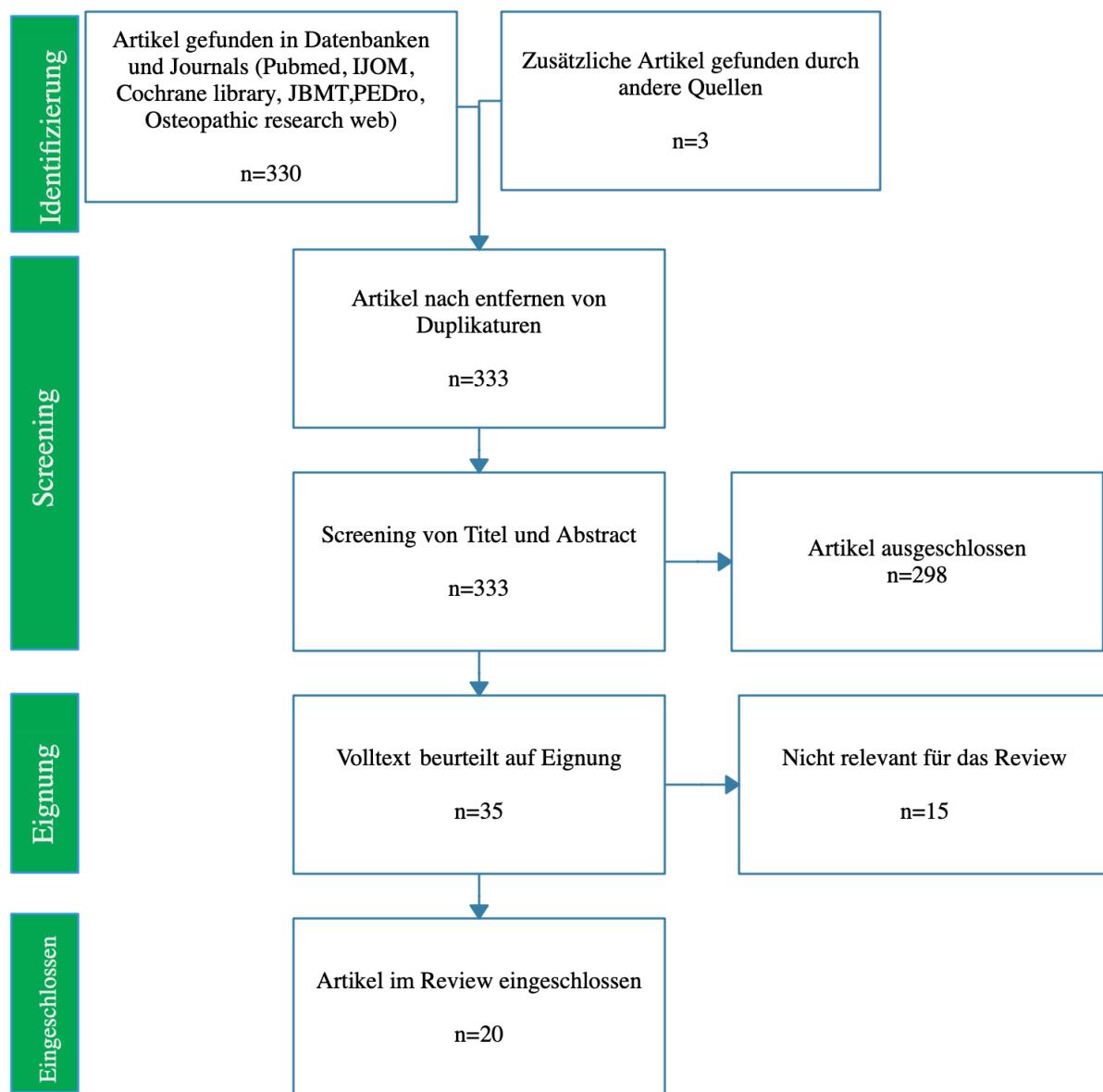


Abbildung 6: Flow-Chart (eigene Darstellung)

Tabelle 5: Inkludierte Studien

Studie	AutorInnen	Titel
1	Bengaard, Bogue, and Crow (2012)	Reliability of diagnosis of somatic dysfunction among osteopathic physicians and medical students
2	Comeaux, Eland, Chila, Pheley, and Tate (2001)	Measurement challenges in physical diagnosis: refining inter-rater palpation, perception and communication
3	Consorti, Basile, Pugliese, and Petracca (2018)	Interrater Reliability of Osteopathic Sacral Palpatory Diagnostic Tests Among Osteopathy Students
4	Degenhardt, Snider, Snider, and Johnson (2005)	Interobserver Reliability of Osteopathic Palpatory Diagnostic Tests of the Lumbar Spine: Improvements from Consensus Training
5	Degenhardt, Johnson, Snider, and Snider (2010)	Maintenance and Improvement of Interobserver Reliability of Osteopathic Palpatory Tests Over a 4-Month Period
6	Fryer (2006)	Factors affecting the intra-examiner and inter-examiner reliability of palpation for supine medial malleoli asymmetry
7	Fryer, McPherson, and O'Keefe (2005)	The effect of training on the inter-examiner and intra-examiner reliability of the seated flexion test and assessment of pelvic anatomical landmarks with palpation
8	Gibbons, Dumper, and Gosling (2002)	Inter-examiner and intra-examiner agreement for assessing simulated leg length inequality using palpation and observation during a standing assessment
9	Hafen (2019)	Reliabilitätsstudie des Befunds nach dem Mechanical Link: Methodologische Studie über die Befunderhebung der Knochen und Gelenke der Extremitäten nach der Methode des Mechanical Link
10	Hutchinson et al. (2017)	Interexaminer reliability study of a standardized myofascial diagnostic technique of the superior thoracic inlet
11	Kmita and Lucas (2008)	Reliability of physical examination to assess asymmetry of anatomical landmarks indicative of pelvic somatic dysfunction in subjects with and without low back pain
12	Lee, Pyle, and Redding (2015)	Accuracy of Anterior Superior Iliac Spine Symmetry Assessment by Routine Structural Examination
13	McPartland and Goodridge (1997)	Counterstrain and traditional osteopathic examination of the cervical spine compared
14	O'Haire and Gibbons (2000)	Inter-examiner and intra-examiner agreement for assessing sacroiliac anatomical landmarks using palpation and observation: pilot study
15	Paulet and Fryer (2009)	Inter-examiner reliability of palpation for tissue texture abnormality in the thoracic paraspinal region
16	Rajendran and Gallagher (2011)	The assessment of pelvic landmarks using palpation: A reliability study of undergraduate students
17	Spring, Gibbons, and Tehan (2001)	Intra-examiner and inter-examiner reliability of a positional diagnostic screen for the lumbar spine

18	Sutton, Nono, Johnston, and Thomson (2013)	The effects of experience on the inter-reliability of osteopaths to detect changes in posterior superior iliac spine levels using a hidden heel wedge
19	Tong, Heyman, Lado, and Isser (2006)	Interexaminer reliability of three methods of combining test results to determine side of sacral restriction, sacral base position, and innominate bone position
20	Vincent-Smith and Gibbons (1999)	Inter-examiner and intra-examiner reliability of the standing flexion test

3.7. Datenanalyse mit der QAREL-Checkliste

In der vorliegenden Arbeit, die einem systematischen Review entspricht, soll die diagnostische Genauigkeit zur Erkennung der somatischen Dysfunktion eingehender untersucht werden. Zur Bewertung der gefundenen Artikel benutzt der Autor die Quality Appraisal of Reliability Studies-Checkliste (QAREL-Checkliste), eine 11-Punkte Checkliste zur qualitativen Beurteilung der diagnostischen Reliabilität von Studien (Lucas, Macaskill, Irwig, & Bogduk, 2010). Die im Jahr 2009 im Journal of Clinical Epidemiology veröffentlichte QAREL-Checkliste wurde in ihrer Entwicklung an bereits vorhandene Checklisten wie der Quality Assessment of Diagnostic Accuracy Studies (Whiting, Rutjes, Reitsma, Bossuyt, & Kleijnen, 2003) und der Standards for Reporting of Diagnostic Accuracy-Checkliste (STARD) angelehnt (Bossuyt et al., 2011). Checklisten wie die STARD werden zur Überprüfung von Studien für die diagnostische Genauigkeit verwendet. Dabei werden ein oder mehrere Tests mit dem Ergebnis eines Referenzstandards verglichen (Bossuyt et al., 2011). Es sind hier jedoch oft Punkte enthalten, die für Studien mit einem Studiendesign zur Überprüfung der diagnostischen Reliabilität nicht relevant sind (Lucas et al., 2010).

Die QAREL-Checkliste benutzt keine Numerische-Rating-Skala, da jeder Punkt separat für sich betrachtet eine Aussage zur Qualität einer Studie hat. Jedes Item kann entweder mit einem klaren „Ja“ oder „Nein“ beantwortet werden. Alternativ ist noch die Aussage „unklar“ möglich. Bei manchen Punkten, steht eine zusätzliche Option zur Auswahl, die als „nicht anwendbar“ (N/A) bezeichnet wird. Ein „Ja“ als Antwort bedeutet eine gute Qualität des jeweiligen Punktes, ein „Nein“ eine mangelnde Qualität. Die Punkte 1 und 2 beziehen sich auf ein mögliches Bias der bewerteten Studien, sowie die Repräsentativität der UntersucherInnen und der untersuchten Personen. Die Punkte 3 - 7 beziehen sich auf die Verblindung der UntersucherInnen, Punkt 8 auf die Reihenfolge der Untersuchungen. Punkt 9 beschreibt, ob das Zeitintervall zwischen den einzelnen Messungen angemessen war. Unter Punkt 10 wird beurteilt, ob der Test korrekt angewandt und interpretiert wurde und Punkt 11, ob angemessene statistische Methoden zur

Analyse der Reliabilität zur Anwendung kamen. Da die QAREL-Checkliste nur im englischen Original vorliegt und eine Übersetzung wie bei der STARD-Checkliste eine Abstimmung mit den OriginalautorInnen erfordert und von einem Expertenteam übernommen werden sollte, werden die einzelnen Punkte weiterhin auf Englisch beibehalten (Breuer et al., 2012). In der folgenden **Tabelle 6** wird die QAREL-Checkliste dargestellt:

Tabelle 6: QAREL-Checkliste (Lucas et al., 2010)

Quality Appraisal of Diagnostic Reliability (QAREL) Checklist		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?				
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?				
3.	Were raters blinded to the findings of other raters during the study?				
4.	Were raters blinded to their own prior findings of the test under evaluation?				
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?				
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?				
7.	Were raters blinded to additional cues that were not part of the test?				
8.	Was the order of examination varied?				
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?				
10.	Was the test applied correctly and interpreted appropriately?				
11.	Were appropriate statistical measures of agreement used?				
		Gesamt			

Die einzelnen Punkte werden nun, wie in dem Ursprungsartikel von Lucas et al. (2010) beschrieben, näher erläutert:

Punkt 1: Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?

Das Ziel einer Reliabilitätsstudie sollte sein, einen Test auf seine Zuverlässigkeit hin zu untersuchen. Die Personen, an denen der Test geprüft wird, sollten daher auch repräsentativ zu jener Population sein, bei denen der Test später auch zur Anwendung kommt. Es macht keinen Sinn, einen Test an einer Personengruppe durchzuführen, bei denen dieser später nicht appliziert wird. Ein Beispiel wäre, wenn ein Test an Erwachsenen angewandt wird, der später nur bei pädiatrischen Patienten zur Anwendung kommt. Die Qualität einer Studie muss darunter nicht unbedingt leiden, eine Repräsentativität auf das typische Patientengut wäre allerdings nicht mehr gegeben.

Um die Frage korrekt beantworten zu können, werden einige Informationen zu den Teilnehmern der Studie benötigt:

- Was ist die Forschungsfrage der vorliegenden Studie und an wem will die Autorin oder der Autor die Ergebnisse anwenden?
- Was sind die kennzeichnenden Merkmale der untersuchten Population? (Alter, Gewicht, Größe, Diagnose usw.) Dies ist wichtig, um mögliche Verzerrungen zu vermeiden und zu wissen, an wem die Untersuchung angewendet wurde.
- In welcher Umgebung wurde die Studie ausgeführt? (Forschungsabteilung einer Universität, Praxis, Klinik usw.) Dies ist ebenfalls wichtig, um Verzerrungen zu vermindern, um zu wissen, an wem die Ergebnisse der Studie zukünftig zur Anwendung kommen sollen.
- Wie wurden die TeilnehmerInnen für die Studie angeworben? Fand eine Randomisierung statt?
- Wie waren die Auswahlkriterien für die StudienteilnehmerInnen? Repräsentieren die TeilnehmerInnen auch die typische Patientengruppe, bei der der Test später auch zur Anwendung kommt?
- Gab es Dropouts oder fehlende Daten? Diese können ein Bias erzeugen, das in der Folge die Repräsentativität der TeilnehmerInnen reduziert.

Die Frage 1 kann mit einem „Ja“ beantwortet werden, wenn die TeilnehmerInnen repräsentativ für die Gruppe sind, an der der Test später zur Anwendung kommt. Wenn die TeilnehmerInnen nicht repräsentativ sind, wird die Frage mit einem „Nein“ beantwortet. Bei Unsicherheiten und fehlenden Informationen, um die Frage korrekt zu beantworten, wird „unklar“ angekreuzt.

Punkt 2: Was the test performed by raters who were representative of those whom the authors intended the results to be applied?

Die klinische Anwendbarkeit einer Studie kann sich reduzieren, wenn der Test von UntersucherInnen durchgeführt worden ist, die nicht repräsentativ genug sind. Die TesterInnen sollten daher auch die Gruppe widerspiegeln, bei der der Test in der täglichen Praxis zum Einsatz kommt. Wenn die Untersuchergruppe nicht der späteren Anwendergruppe entspricht, wird dadurch nicht die Qualität einer Studie vermindert, sondern die Anwendbarkeit in der Praxis.

Um die Frage richtig zu beantworten, sind Informationen nötig, die klären, welche Untersuchergruppe die Studie repräsentieren soll:

- Was war die Forschungsfrage und an wem sollen die Studienergebnisse zur Anwendung kommen?
- Welche Erkennungsmerkmale haben die UntersucherInnen? (Ausbildungsstand und Qualifikation, Erfahrungswerte der Testdurchführung sowie spezielles Training für die Studie)
- In welcher Umgebung wurde die Studie ausgeführt? Die ist wichtig, um zu ermitteln, für welche UntersucherInnen die Studienergebnisse später zur Anwendung kommen können.
- Wie waren die Auswahlkriterien für die StudienteilnehmerInnen? Die ist ebenfalls wichtig, um die Zielgruppe zu ermitteln, die einen praktischen Nutzen von den Studienergebnissen haben könnte.
- Gab es Dropouts oder fehlende Daten? Auch hier ist eine Wichtigkeit gegeben, um ein mögliches Bias zu reduzieren.

Wenn der Test von UntersucherInnen ausgeführt wurde, die repräsentativ für spätere Nutzer sind, kann Frage 2 mit einem „Ja“ beantwortet werden. Wenn die UntersucherInnen nicht repräsentativ genug sind, wird die Frage mit einem „Nein“ beantwortet. Bei Unsicherheiten zur Beantwortung oder bei unzureichenden Informationen wird die Frage mit einem „Unklar“ beantwortet.

Punkt 3: Were raters blinded to the findings of other raters during the study?

Die Punkte 3-7 beziehen sich auf die Verblindung der TesterInnen und deren Einfluss auf die Testergebnisse. Bei fehlender Verblindung der UntersucherInnen, hat dies einen Einfluss auf die Qualität einer Studie und kann die Ergebnisse und deren Anwendbarkeit unbrauchbar machen.

Punkt 3 ist für Studien relevant, in denen mehr als eine Testerin oder ein Tester die Untersuchung durchführt. Wenn die UntersucherInnen zueinander nicht verblindet wurden, kann dies ebenfalls die Testergebnisse beeinflussen.

Wenn zwei oder mehr TesterInnen an der Studie involviert sind und diese zu den Testergebnissen der anderen UntersucherInnen verblindet wurden, kann die Frage mit einem „Ja“ beantwortet werden. Erfolgte keine Verblindung, wird sie mit einem „nein“ beantwortet. Bei Unsicherheiten oder einem Mangel an Information erfolgt das Kreuz bei „unklar“ und bei nur einer Testerin oder einem Tester bei „nicht anwendbar“.

Punkt 4: Were raters blinded to their own prior findings of the test under evaluation?

Das Ergebnis einer Reliabilitätsstudie kann durch eine mangelnde Verblindung der Messwerte unbrauchbar werden. Wenn ein Messergebnis mehrmals wiederholt wird und der Testerin oder dem Tester das frühere Ergebnis bewusst ist, könnte dies die weiteren Ergebnisse beeinflussen. Damit könnte sich die Qualität der Studie reduzieren. Diese Frage ist nur für Studien relevant, in denen ein Tester mehrere Testwiederholungen macht, wie in einer intrarater Reliabilitätsstudie.

Wenn die UntersucherInnen korrekt verblindet wurden, ist die Frage mit einem „Ja“ zu beantworten, mit einem „Nein“, wenn keine Verblindung stattgefunden hat. Ein „Unklar“ ist bei Unsicherheiten oder unzureichenden Informationen auszufüllen. Wenn die UntersucherInnen die ProbandInnen nur einer Testung unterzogen oder nur eine Messung vornahmen, wie bei einigen interrater Reliabilitätsstudien, wird die Frage mit „nicht anwendbar“ beantwortet.

Punkt 5: Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?

Ein Referenzstandard stellt die momentan beste Methode dar, eine bestimmte Variabel zu messen. Wenn die Ergebnisse des derzeitigen Referenzstandards bekannt sind, könnte die

Testerin oder der Tester dadurch Rückschlüsse auf die Grunderkrankung oder die Dysfunktion des betroffenen Individuums ziehen. Es wäre denkbar, dass dies die Anwendbarkeit und die Qualität einer Studie gefährdet. Der Referenzstandard wird hier in Punkt 5 separat von anderen klinischen Informationen, die nicht Teil der Untersuchung sind, aufgeführt. Diese werden in Punkt 6 näher betrachtet.

Wenn die UntersucherInnen zu den Ergebnissen des Referenzstandards verblindet wurden oder ihnen die Grunderkrankung der zu testenden Personen nicht bekannt ist, wird die Frage mit einem „Ja“ beantwortet. Erfolgte keine Verblindung, wird sie mit einem „Nein“ beantwortet. Bei unzureichenden Informationen oder Unsicherheit bei der Beantwortung der Frage wird diese mit einem „Unklar“ vermerkt. Wenn kein allgemeiner Referenzstandard zur Verfügung steht oder die Testpersonen nicht mit dem Referenzstandard untersucht wurden, kann die Frage mit „nicht anwendbar“ beantwortet werden.

Punkt 6: Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?

Jegliches klinische Wissen, wie die Patientengeschichte, frühere Behandlungsmaßnahmen oder Nebenerkrankungen kann ein Rückschluss auf den Gesundheitszustand der Testperson sein. Das Wissen über Erkrankungen, Nichterkrankung oder Dysfunktionen einer Testperson kann die Untersucherin oder der Untersucher in seinem Testergebnis beeinflussen. Nur die Informationen, die für die Untersuchung nötig sind oder die von dem Studiendesign vorgegeben werden, sollten den TeilnehmerInnen zur Verfügung stehen.

Waren die TesterInnen über klinische Informationen, die nicht Teil der Testung oder der Studie waren, verblindet, dann wird die Frage mit einem „Ja“ beantwortet. Wenn dies nicht zutrifft und keine Verblindung stattgefunden hat, ist die Antwort auf die Frage ein „Nein“. Wenn nicht genug Informationen zur Verfügung stehen oder Unsicherheit besteht, wird „Unklar“ angekreuzt und bei fehlender Anwendbarkeit der Frage „nicht anwendbar“ vermerkt.

Punkt 7: Were raters blinded to additional cues that were not part of the test?

Wie in den vorangegangenen Fragen, geht es hier um Informationen, die das Testresultat beeinflussen können. Dies wären auch zusätzliche Hinweise wie Tattoos, Narben, Akzente in der Sprache und andere einmalige Merkmale mit Wiedererkennungswert. Wenn UntersucherInnen auf solche Hinweise nicht verblindet wurden, könnte dies ebenfalls Einfluss auf die Untersuchungsergebnisse nehmen.

Bei vorliegender Verblindung wird die Frage mit einem „Ja“ beantwortet, bei fehlender Blindierung mit einem „Nein“ und bei Unsicherheit oder mangelnden Informationen mit einem „Unsicher“.

Punkt 8: Was the order of examination varied?

Die Untersuchungsabfolge, in denen die TesterInnen einen Test an den ProbandInnen durchführen, kann unter Umständen zu einem Bias oder zu Qualitätseinbußen der Studie führen. Vor allem bei interrater Studien ist es sinnvoll, die Reihenfolge der TesterInnen zu variieren.

Ein Beispiel wäre, wenn Person A immer als Erste mit dem Testablauf beginnt und Person B somit immer als Zweite an der Reihe ist. Wenn Person A an einem Probanden eine Hautrötung durch manuellen Druck hinterließe, kann dies Einfluss auf das Testergebnis von Person B nehmen. Eine Randomisierung der Untersuchungsreihenfolge, wäre eine Möglichkeit Bias zu reduzieren.

Erfolgte eine Beachtung der Untersuchungsreihenfolge, ist die Antwort auf die Frage ein „Ja“, erfolgte keine Beachtung, ein „Nein“. Wenn ein Artikel nicht genug Informationen aufweist oder Unsicherheit zur klaren Beantwortung der Frage besteht, wird die Antwort ein „Unklar“ sein. Falls es für die Studie nicht von Bedeutung ist, wie z.B. bei der Beurteilung von Röntgenbildern, ist die Antwort ein „nicht anwendbar“.

Punkt 9: Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?

Die Stabilität einer Testvariablen ist entscheidend für Studien, die sich mit der Reliabilität von Tests auseinandersetzen. Nur wenn die Stabilität einer Variablen gegeben ist, kann die Übereinstimmung von Tests zwischen mehreren TesterInnen gegeben werden. Ob dies zutrifft, kann nur bestimmt werden, wenn der Referenzstandard bekannt ist. Wenn dies nicht der Fall ist, sollten Studien so geplant werden, dass es eine theoretische Stabilität der Testvariablen gibt. Als Beispiel wird die Blutdruckmessung genannt, die nur Nutzen hat, wenn auch die natürliche Blutdruckabweichung in der Studie ihre Beachtung findet. Echte oder theoretische Veränderungen der Variablen sollten daher immer in Betracht gezogen werden.

Wenn dies berücksichtigt wurde, kann die Frage mit einem „Ja“ beantwortet werden, wenn nicht, mit einem „Nein“. Bei einem Mangel an Informationen oder bei Unsicherheiten wird die Frage mit einem „Unklar“ beantwortet.

Punkt 10: Was the test applied correctly and interpreted appropriately?

Eine exakte Beschreibung des Testprozesses einer Studie ist für die Beantwortung dieser Frage notwendig. Auch die richtige Anwendung und Interpretation der Testergebnisse spielen eine entscheidende Rolle für die Aussagekraft einer Studie. Wenn die Interpretation einer Untersuchung nicht notwendig oder möglich ist, wie bei einer Messung der Körpergröße oder des Gewichts, wird nur die Anwendung und Beschreibung der Ausführung bewertet. Weitere Informationen sind notwendig für die Beantwortung von Punkt 10:

- Wurden die Tests korrekt angewandt? Korrekt bedeutet, so wie er standardisiert im Theorienteil der Studie beschrieben wurde. Eine abweichende Ausführung kann eine ansonsten gut ausgeführte Studie in ihrer Qualität vermindern. Veränderungen von in der Literatur standardisierten Tests müssen in der Studie entsprechend detailliert beschrieben werden.
- Sind die Kriterien zur Interpretation eines Tests ausreichend definiert? Wenn das Testergebnis einer Interpretation unterzogen werden muss und die Kriterien unzureichend definiert sind, kann dies Einfluss auf die Studienqualität nehmen. Wenn zwei oder mehr TestInnen eine abweichende Meinung zur Interpretation eines Tests haben, könnte dies die anschließenden Testergebnisse verfälschen.

Punkt 10 wird mit einem „Ja“ beantwortet, wenn der Test korrekt angewandt und die Kriterien zur Interpretation adäquat definiert wurden. Wenn dies nicht der Fall war, wird die Frage mit einem „Nein“ beantwortet und „Unklar“ bei nicht klaren Angaben in der Studie.

Punkt 11: Were appropriate statistical measures of agreement used?

Je nach Studiendesign sollten unterschiedliche statistische Verfahren zum Einsatz kommen. So kommt die Interklassenkorrelation zur Abschätzung der Reliabilität von Messwertreihen zur Anwendung und der Kappa-Wert bei der Einschätzung von inter- und intrarater Reliabilität.

Wenn die statistische Analyse korrekt angewendet wurde, kann die Frage mit einem „Ja“ beantwortet werden. Bei unpassender Statistik wird der letzte Punkt mit einem „Nein“ beantwortet und bei mangelnden Informationen oder Unsicherheit mit einem „Unklar“.

Die QAREL Checkliste stellt somit ein einfaches und schnelles Instrument zur Beurteilung von Reliabilitätsstudien dar (Lucas et al., 2010). Die Gutachterin oder der Gutachter, der dieses Tool benutzt, sollte sich jedoch intensiv mit den oben beschriebenen Punkten vertraut machen.

4. Ergebnisse

In diesem Kapitel werden die in der Literaturrecherche gefundenen Studien einzeln dargestellt und mit der QAREL-Checkliste bewertet. Da bei dem QAREL Bewertungssystem keine Numerische-Rating-Skala zur Anwendung kommt, entscheidet nicht die Gesamtzahl an positiven Punkten über die Qualität der Studie. Daher werden in dieser Übersichtsarbeit die Ergebnisse für jede der inkludierten Studien tabellarisch dargestellt. Durch die vorherige Zusammenfassung der jeweiligen Artikel sollte die Punktevergabe anhand der QAREL-Checkliste klar hervorgehen. Übereinstimmungswerte des Cohen's Kappa werden nach der Einteilung von Landis und Koch (Landis & Koch, 1977) interpretiert. Dabei entspricht ein Wert von 0 – 0.20 einer schwachen Übereinstimmung, 0.21 – 0.40 einer ausreichenden, 0.41 – 0.60 einer moderaten, 0.61 – 0.80 einer starken und 0,81 – 0.99 einer nahezu perfekten Übereinstimmung. Diese Interpretation wird in **Tabelle 7** zusammenfassend dargestellt.

Tabelle 7: Interpretation des Kappa Koeffizienten nach Landis and Koch (1977)

Wert	Übereinstimmungswert
0 – 0.20	Schwach
0.21 – 0.40	Ausreichend
0.41 – 0.60	Moderat
0.61 – 0.80	Stark
0.81 – 0.99	Nahezu perfekt

4.1. Reliability of diagnosis of somatic dysfunction among osteopathic physicians and medical students (Bengaard et al., 2012)

Die von Bengaard und Team durchgeführte Studie von 2012 beschäftigte sich mit der Diagnose der SD über das Untersuchungskriterium der Asymmetrie. Ein simpler palpatorischer Test zur Stellungsüberprüfung der Spina iliaca anterior superior (SIAS) kam dabei zur Anwendung. Alle 139 teilnehmenden UntersucherInnen waren entweder fertig ausgebildete OsteopathInnen oder Studierende. Allesamt waren BesucherInnen der Versammlung der American Academy of Osteopathy, bei der auch die Studie selber über mehrere Abende verteilt stattfand.

Aus den Teilnehmern der Konferenz wurden freiwillige Personen rekrutiert, wobei mehr StudentInnen als postgraduierte OsteopathInnen bei der Studie teilnahmen. Zwölf der ebenfalls an der Konferenz teilnehmenden Personen, stellten sich als freiwillige ProbandInnen zur Verfügung. Außerdem kam ein fixiertes Plastikskelett als Testobjekt zum Einsatz. Die TeilnehmerInnen wurden im Voraus über den Testablauf und die Testausführung informiert. Die TesterInnen sollten innerhalb einer Minute die Stellung der SIAS mit einer leichten Palpation ohne zusätzlichen Mobilitätstest ausfindig machen. Eine Abweichung von mindestens 1cm auf einer Seite nach superior entspreche einem positiven Befund. Auf einem vorgefertigten Zettel mit Informationen zum Ablauf konnten die TesterInnen ihre Ergebnisse nach dem Test und ihren Ausbildungsstand eintragen. Die verschiedenen Antwortmöglichkeiten waren entweder SIAS links superior, SIAS rechts superior oder SIAS bilateral auf gleicher Höhe befindend. Es untersuchten pro Schicht (insgesamt 6 Schichten) 25 TesterInnen jeweils 2 lebende ProbandInnen und das Skeletmodell. Die ProbandInnen wurden gebeten, sich möglichst nicht zu bewegen oder ihre Position nicht zu verändern. Das Skeletmodell wurde auf der Bank fixiert, mit beiden SIAS auf gleicher Höhe eingestellt.

Die Analyse der Palpationsdaten an den lebenden ProbandInnen wurde mittels dem Fleiss' Kappa Koeffizienten durchgeführt. Die Übereinstimmung zwischen den unterschiedlichen TesterInnen ergab dabei einen Koeffizienten von 0.028 was einer schwachen Übereinstimmung entspricht. Der Fleiss' Kappa Wert innerhalb der Studentengruppen war mit 0,038 leicht höher, jedoch trotzdem noch weit unterhalb eines guten Ergebnisses.

Bei der Auswertung der Tests an dem fixierten Plastikmodell ist ein gutes Ergebnis erwartet worden. Im Gegensatz zu den lebenden ProbandInnen war den StudienleiterInnen das richtige Ergebnis hier bekannt. Es wurde angenommen, dass ein Großteil der 139 TesterInnen die SIAS auf gleicher Höhe vorfinden. Es konnten jedoch nur 15 aller 139 UntersucherInnen dieses Ergebnis richtig positiv identifizieren. In diesem Teil der Studie konnten die TesterInnen mit mehr Erfahrung jedoch etwas besser abschließen als die Studierenden.

Die Palpation der SIAS als asymmetrisches Zeichen zur Identifikation einer SD konnte in der vorliegenden Studie also keine Reliabilität aufweisen.

Als mögliche Fehlerquelle der Studie wird unter anderem die „Lebendigkeit“ der getesteten Personen diskutiert. Diese könnten in der Zeit der Testung durchaus ihre Position verändert haben. Auch ein möglicher Einfluss durch die wiederholte Testung von 25 UntersucherInnen auf die Stellung der SIAS wäre denkbar. Die Getesteten gaben auch an, dass einige TesterInnen

eine Kompression auf die SIAS ausgeübt haben, welche nicht Teil des Tests war und die ebenfalls die Position des Beckens verändert haben könnte.

Da auch ein Plastikmodell zur Anwendung kam, wurde Frage 1 mit „Unklar“ beantwortet. Frage 7 musste mit „nein“ beantwortet werden, da Narben oder Ähnliches die TesterInnen zu einem falschen Ergebnis geführt haben könnten. Eine korrekte Ausführung des Tests konnte nicht nachgewiesen werden, da einige Probanden von einem Kompressionstest anstatt einer sanften Palpation berichteten. Die Auswertung dieser Studie anhand der QAREL-Checkliste wird in **Tabelle 8** dargestellt.

Tabelle 8: Reliability of diagnosis of somatic dysfunction among osteopathic physicians and medical students (QAREL-Checkliste)

QAREL Checkliste für die Studie 1		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?		X		
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?		X		
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?		X		
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X			
7.	Were raters blinded to additional cues that were not part of the test?	X			
8.	Was the order of examination varied?	X			
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?		X		
10.	Was the test applied correctly and interpreted appropriately?	X			
11.	Were appropriate statistical measures of agreement used?	X			
		Gesamt	4	3	2

4.2. Measurement challenges in physical diagnosis: refining inter-rater palpation, perception and communication (Comeaux et al., 2001)

In der vorliegenden Studie wurde die interrater Reliabilität anhand eines vorher festgelegten Diagnoseprotokolls untersucht. Die UntersucherInnen waren alle mindestens 10 Jahre im Beruf und seit mindestens 3 Jahren im Unterrichtsbereich der funktionellen Techniken nach Johnston tätig. Nach dieser Methode werden Parameter wie Gewebebeschaffenheit und passive Bewegungsüberprüfung eingesetzt, um ein spezifisches Segment, welches als Schlüsselläsion definiert wird, ausfindig zu machen. In dieser Studie beschränkte sich der Untersuchungsbereich auf die Segmente von C2 bis Th8. Die 54 freiwilligen Testpersonen wurden durch eine öffentliche Anzeige ausfindig gemacht. Dabei wurden klare Ein- und Ausschlusskriterien definiert. Die UntersucherInnen waren verblindet zu jeglichen Informationen über die zu testenden Personen und zu den Untersuchungsergebnissen der anderen UntersucherInnen.

Die Testpersonen wurden jeweils in einen Raum geführt und erhielten keine weiteren Informationen über den Testablauf. Die erste Testerin oder der erste Tester betrat den Raum und markierte Th1 und Th8 zur besseren Orientierung mit jeweils einem roten Punkt. Dieser blieb bis zum Ende der Untersuchungen dort haften. Eine Verblindung zu zusätzlichen Informationen wie regionalen Hautveränderungen oder Narben fand nicht statt.

Nach der Testung wurde die gefundene Zone mit einem weiteren Punkt markiert und ein Foto gemacht. Der zusätzliche Punkt wurde wieder entfernt. Die Untersuchungen fanden innerhalb von vier Halbtagen statt. Die Reihenfolge der TesterInnen wurde nicht randomisiert, dies soll aber laut der Studiendiskussion in zukünftigen Untersuchungen geändert werden. Die statistische Auswertung der Daten fand mittels eines Statistikprogramms statt. Die Ergebnisse des Cohen's Kappa Wertes betrugen dabei 0.12 bis 0.56.

Wie eine Stabilität der Messwerte erreicht werden kann, wird im Diskussionsteil angegeben und soll in zukünftigen Arbeiten zum Einsatz kommen. Da diese Arbeit eine Pilotstudie war, konnte sie einige Probleme aufdecken, die folgende Arbeiten inspirieren soll, dies besser zu machen. Die Auswertung mit der QAREL-Checkliste wird in **Tabelle 9** dargestellt.

Tabelle 9: Measurement challenges in physical diagnosis: refining inter-rater palpation, perception and communication (QAREL-Checkliste)

QAREL Checkliste der Studie 2		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?			X	
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?			X	
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X			
7.	Were raters blinded to additional cues that were not part of the test?	X			
8.	Was the order of examination varied?	X			
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?		X		
10.	Was the test applied correctly and interpreted appropriately?	X			
11.	Were appropriate statistical measures of agreement used?	X			
		Gesamt	6	2	1
					2

4.3. Interrater Reliability of Osteopathic Sacral Palpatory Diagnostic

Tests Among Osteopathy Students (Consorti et al., 2018)

Die von Consorti et al. durchgeführte Studie aus dem Jahr 2018 war bis dato die erste Reliabilitätsstudie, die sich mit der Anwendung aller diagnostischen TART Kriterien am Sakrum auseinandersetzte. Die Untersuchung selber wurde von drei freiwilligen Studierenden durchgeführt, die sich im letzten Jahr des Osteopathie Vollzeitstudiums befanden. Die StudentInnen wurden einem speziellen Training unterzogen, das die Ergebnisse der Interrater-Reliabilität erhöhen sollte. Da ein solches Übereinstimmkeitstraining schon in früheren

Studien zur Anwendung kam und dies nachweislich die Reliabilität verbesserte, entschieden sich die Autoren für ein solches Vorgehen. Das Training wurde über drei Tage mit jeweils dreistündigen Einheiten durchgeführt. Es wurden dabei Übungen zum Erkennen von Gewebeveränderungen, Asymmetrien, Bewegungseinschränkungen und Sensibilitätsveränderungen des Gewebes durchgeführt. Ebenfalls wurde die Interpretation der Testergebnisse untereinander abgeglichen. Ein Osteopath mit einem Jahr Berufserfahrung leitete die Trainingseinheiten.

Die Studie wurde an 52 Osteopathie-StudentInnen durchgeführt, die auf freiwilliger Basis rekrutiert wurden. Ein- und Ausschlusskriterien wurden dabei klar definiert und schlossen nur Probandinnen und Probanden aus, die vor kurzem ein Trauma erlitten, eine Operation im Lumbalbereich oder andere akute Erkrankungen hatten, welche die Studienergebnisse verfälschen könnten. Auch wurden Personen aus der Studie ausgeschlossen, die nicht länger als elf Minuten auf dem Bauch liegen konnten, da dies für die Untersuchung eine nötige Voraussetzung war.

Die Untersuchung selbst wurde in einem Raum mit drei identischen Behandlungsstühlen, mit jeweils drei Metern Abstand und einem Sichtschutz dazwischen, durchgeführt. Die UntersucherInnen wurden per Zufall den Stühlen zugewiesen und führten ihre Tests am bis auf die Unterwäsche entkleideten Modell durch. Der Wechsel der TesterInnen wurde ebenfalls randomisiert und von einer Aufsichtsperson geregelt. Die Patientengeschichte, Erkrankungen und andere Informationen über die Testpersonen, waren den TesterInnen nicht bekannt. Es fand ebenfalls eine Verblindung zu den Testergebnissen der anderen TesterInnen statt. Auf eine Verblindung von zusätzlichen Hinweisen wie Narben oder Hautveränderungen wurde verzichtet. Beim Test der Gewebeempfindlichkeit wurden die Probanden angeleitet, ihre Antwort lediglich zu flüstern, um ein Bias zu reduzieren. Der derzeitige Referenzstandard für die Untersuchung einer SD besagt, dass mindestens zwei der vier Parameter positiv sein sollen und die Anamnese und das ganze Patientenbild auch einen Teil der Diagnose ausmachen. Auf diese Information hin waren die TesterInnen jedoch verblindet.

Die Ergebnisse der Studie gaben einen allgemeinen Fleiss'Kappa Koeffizienten für die Erkennung einer somatischen Dysfunktion von 0.17 an. Die genauen Ergebnisse betragen für Gewebeveränderungen 0.28, für Asymmetrie 0.29, für Bewegungseinschränkungen 0.32 und 0.34 für die Gewebeempfindlichkeit. In **Tabelle 10** findet sich die Auswertung mit der QAREL-Checkliste.

Tabelle 10: Interrater Reliability of Osteopathic Sacral Palpatory Diagnostic Tests Among Osteopathy Students (QAREL-Checkliste)

QAREL Checkliste der Studie 3		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?		X		
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?				X
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?				X
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X			
7.	Were raters blinded to additional cues that were not part of the test?		X		
8.	Was the order of examination varied?	X			
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?			X	
10.	Was the test applied correctly and interpreted appropriately?	X			
11.	Were appropriate statistical measures of agreement used?	X			
		Gesamt	7	0	2
					2

4.4. Interobserver Reliability of Osteopathic Palpatory Diagnostic Tests of the Lumbar Spine: Improvements from Consensus Training (Degenhardt et al., 2005)

Ziel dieser Übersichtsarbeit war es zu untersuchen, ob ein vorheriges Übereinstimmungstraining die intertester Reliabilität von osteopathischen Tests verbessern könnte. Die Untersuchungen wurden von drei TesterInnen durchgeführt, die auch Teil der Autorengruppe der Studie waren. Diese wiesen alle mehr als zehn Jahre Berufserfahrung auf und waren Mitglieder des Kirksville College of Osteopathic Medicine (KCOM), wo auch die

Untersuchung stattfand. Die getesteten Personen waren Freiwillige, die im Umfeld des Colleges rekrutiert wurden.

Die Studie selber fand in drei Phasen aufgeteilt statt. Phase 1 entsprach einer Untersuchung vor dem Training, Phase 2 beschreibt das Training selber und Phase 3 eine weitere Untersuchung die nach dem Training durchgeführt wurde. In den Phasen eins und drei, wo auch Daten erhoben wurden, fand eine doppelte Verblindung statt. Die Untersuchung bezog sich auf die Lumbalwirbel eins bis vier. Der fünfte Lumbalwirbel wurde von der Testung ausgeschlossen, da dieser sehr häufig anatomische Variablen aufweist, die zur Verzerrung der Testergebnisse führen könnten. In Phase 1 wurden 42 Personen in 14 Untergruppen zu jeweils drei Subjekten aufgeteilt, die über vier Tage lang einem genau definierten Testprotokoll unterzogen wurden. Die Testung enthielt Verfahren zur Bestimmung von Gewebeempfindlichkeit, Veränderungen der Gewebebeschaffenheit und Tests zur Bestimmung von Asymmetrien und des Bewegungsausmaßes. In Phase 2, die der Trainingsphase entsprach, wurde über vier Monate lang jede Woche ein zweistündiges Training durchgeführt. Dabei trafen sich alle UntersucherInnen und stimmten ihre Tests und Interpretationen aufeinander ab. In Phase 3 wurde erneut eine Untersuchung wie in Phase 1 durchgeführt, diesmal jedoch an 77 freiwilligen TeilnehmerInnen. Die Untersuchung fand diesmal über sechs Tage statt.

Eine statistische Auswertung wurde mittels dem Kappa Koeffizienten errechnet. Die Ergebnisse der Testungen vor und nach dem Training konnten durchwegs eine Verbesserung der Reliabilität aufweisen. Der Test für die Gewebeempfindlichkeit verbesserte sich von einem Kappa Wert 0.32 auf 0.68, der für die Gewebebeschaffenheit von 0.12 auf 0.45. Die Tests für das Bewegungsausmaß verbesserten sich von 0.10 auf 0.20 und die für Asymmetrie von 0.17 auf 0.34. Die Beurteilung der Studie anhand der QAREL-Checkliste wird in **Tabelle 11** dargestellt

Tabelle 11: Interobserver Reliability of Osteopathic Palpatory Diagnostic Tests of the Lumbar Spine: Improvements from Consensus Training (QAREL-Checkliste)

QAREL Checkliste der Studie 4		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?		X		
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?		X		
3.	Were raters blinded to the findings of other raters during the study?		X		

4.	Were raters blinded to their own prior findings of the test under evaluation?	X
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?	X
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X
7.	Were raters blinded to additional cues that were not part of the test?	X
8.	Was the order of examination varied?	X
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?	X
10.	Was the test applied correctly and interpreted appropriately?	X
11.	Were appropriate statistical measures of agreement used?	X
Gesamt		6 2 1 2

4.5. Maintenance and Improvement of Interobserver Reliability of Osteopathic Palpatory Tests Over a 4-Month Period (Degenhardt et al., 2010)

Im Jahr 2010 wurde eine weitere Arbeit von Degenhardt und seinem Team veröffentlicht. Ziel war es erneut, die Reliabilität von osteopathischen Palpationstests zur Identifikation der somatischen Dysfunktion an der LWS zu untersuchen. Es kam wieder ein Übereinstimmigkeitstraining über einen Zeitraum von vier Monaten zum Einsatz mit der Forschungsfrage, ob dieses die Reliabilität der Testungen verbessern könne. Es wurden 64 ProbandInnen mit und ohne Low Back Pain durch ein Zufallsverfahren an der KCOM rekrutiert. Die Ein- und Ausschlusskriterien wurden genau definiert. Die zwei UntersucherInnen, beide Mitglieder der American Osteopathic Association, hatten zwischen drei bis zehn Jahre Berufserfahrung. Die OsteopathInnen wurden vor der eigentlichen Untersuchung einem Übereinstimmigkeitstraining unterzogen, das an zusätzlich rekrutierten Testpersonen durchgeführt wurde. Dabei wurden Messungen vor und nach dem Training durchgeführt. Der Ablauf war dem der vorherigen Studie ident. Es gab ein genaues Protokoll der Testdurchführung und Interpretation, das während der Trainingseinheiten nochmals

angepasst wurde. Nach einer zweimonatigen Pause startete die eigentliche Testphase. Die 64 ProbandInnen wurden dabei auf die TART-Kriterien hin untersucht. Die TesterInnen waren voll verblindet und auch die Reihenfolge der Testungen wurde variiert. Wieder war unklar, wie mit der Teststabilität verfahren wurde. Die statistische Analyse erfolgte erneut mit dem Kappa Koeffizienten. Innerhalb des viermonatigen Intervalls verbesserte sich die Untersuchung der Asymmetrie von 0.56 auf 0.72, der Gewebebeschaffenheit von 0.23 auf 0.55. Bei der Testung auf Beweglichkeit wurde eine Verbesserung von 0.30 auf 0.50 erzielt und bei der Schmerzprovokation (Tenderness) von 0.61 auf 0.88, was einer sehr guten Reliabilität entspricht. Alle vier Tests wiesen bei der letzten Testung eine moderate bis starke interrater Reliabilität auf, diese verbesserte sich zunehmend über den Faktor Zeit. In **Tabelle 12** wird die Beurteilung der Studie anhand der QAREL-Checkliste dargestellt.

Tabelle 12: Maintenance and Improvement of Interobserver Reliability of Osteopathic Palpatory Tests Over a 4-Month Period (QAREL-Checkliste)

QAREL Checkliste der Studie 5		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?			X	
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?			X	
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X			
7.	Were raters blinded to additional cues that were not part of the test?	X			
8.	Was the order of examination varied?	X			
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?			X	
10.	Was the test applied correctly and interpreted appropriately?	X			
11.	Were appropriate statistical measures of agreement used?	X			

4.6.Factors affecting the intra-examiner and inter-examiner reliability of palpation for supine medial malleoli asymmetry (Fryer, 2006)

Der Autor dieser Arbeit überprüfte die Palpation des medialen Malleolus in Rückenlage als asymmetrisches Zeichen einer SD des Iliosakralgelenks. Es handelte sich dabei um eine intra- und interrater Reliabilitätsstudie. Fünf Osteopathie-StudentInnen im letzten Jahr der Ausbildung führten die Testungen durch. Die Untersuchung wurde an zwei Gruppen mit je fünf asymptomatischen weiblichen Studentinnen durchgeführt. Die Testdurchläufe fanden mit einer Woche Abstand statt. Sichtbare Merkmale wie Tattoos, Hautveränderungen, Narben oder Schmerzen in Rückenlage waren die definierten Ausschlusskriterien. Die UntersucherInnen führten vor der Studie ein Training über 20 Minuten durch. Dabei wurden die Tester aufeinander abgeglichen, die genaue Testdurchführung definiert und die Interpretation des Tests besprochen.

Die fünf Probandinnen wurden bis zu den Knien abgedeckt und wechselten die Tische zwischen den Untersuchungsrunden. Damit wurde die Untersuchungsreihenfolge variiert. Jede TesterIn führte die Untersuchung an jeder Probandin dreimal in drei unterschiedlichen Runden durch. Durch dieses Prozedere wurden die TesterInnen zu ihren vorher gefundenen Ergebnissen verblindet. Das Feedback der Untersucherinnen nach der ersten Gruppe veranlasste den Autor nach Probandinnen zu suchen, die eine tatsächliche Abweichung der Malleolenposition von vier bis zehn Millimetern hatten, die er aus 15 weiteren Freiwilligen selber aussuchte. Damit wurden Probandinnen ausgeschlossen, die keine oder nur eine Abweichung unter drei Millimetern hatten.

Die Ergebnisse wurden mit dem Kappa-Wert statistisch angegeben. Bei dem ersten Testdurchlauf wurde ein interrater Wert von 0.22 und beim zweiten Durchlauf ein Wert von 0.94 ermittelt. Das Ergebnis der intrarater Testung brachte Werte von 0.31 bei der ersten Gruppe und erneut 0.94 bei der zweiten Gruppe hervor. Somit besteht laut dieser Studie eine sehr gute Wahrscheinlichkeit, eine Asymmetrie der knöchernen Punkte ab einem Wert von 4mm zu identifizieren. Die Ergebnisse der Beurteilung der Studie mit der QAREL-Checkliste werden in **Tabelle 13** dargestellt.

Tabelle 13: Factors affecting the intra-examiner and inter-examiner reliability of palpation for supine medial malleoli asymmetry (QAREL-Checkliste)

QAREL Checkliste der Studie 6		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?		X		
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?	X			
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?			X	
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X			
7.	Were raters blinded to additional cues that were not part of the test?	X			
8.	Was the order of examination varied?	X			
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?			X	
10.	Was the test applied correctly and interpreted appropriately?	X			
11.	Were appropriate statistical measures of agreement used?	X			
		Gesamt	9	0	1
					1

4.7. The effect of training on the inter-examiner and intra-examiner reliability of the seated flexion test and assessment of pelvic anatomical landmarks with palpation (Fryer et al., 2005)

Zur Erkennung einer SD im Bereich des Beckens wird eine Vielzahl an Tests und Untersuchungsverfahren beschrieben. Sowohl der sogenannte Flexionstest im Stand als auch die statische Untersuchung von knöchernen Referenzpunkten wurden in dieser Arbeit auf ihre inter- und intrarater Reliabilität hin genauer untersucht. Ob ein vorheriges

Übereinstimmigkeitstraining sich auf die Testergebnisse auswirkt, war ein weiteres Ziel dieser Forschungsarbeit. Insgesamt wurden zehn StudentInnen im letzten Jahr des Osteopathiestudiums als freiwillige Tester rekrutiert. Zehn weibliche Studierende wurden als freiwillige Probandinnen ausgewählt. Um Wiedererkennungsmerkmale zu reduzieren, sollten alle Probandinnen bei der Untersuchung die gleiche Unterwäsche tragen. Auch wurden Studentinnen ausgeschlossen, die andere Wiedererkennungsmerkmale wie Tattoos oder Narben aufwiesen oder eine längere Rücken- oder Bauchlage nicht tolerierten.

Die TesterInnen wurden in zwei Gruppen aufgeteilt. Eine der Gruppen erhielt zweimal eine Stunde ein Übereinstimmigkeitstraining, bei dem die TesterInnen ihre Tests untereinander ab sprachen. Die andere Gruppe erhielt kein Training. Kurz vor dem Beginn der eigentlichen Untersuchung erhielten alle UntersucherInnen eine Einführung zur Testausführung und Interpretation. Die abgedeckten Probandinnen wurden nacheinander in Bauchlage und Rückenlage einer positionellen Untersuchung der vorgegebenen Knochenpunkte unterzogen. Im Anschluss folgte noch der Flexionstest im Sitzen. Nach der Untersuchungsrunde mussten die UntersucherInnen einen visuellen Gedächtnistest ausführen und wurden dann für den nächsten Untersuchungsgang einer anderen Startposition zugewiesen. Dieses Prozedere sollte dazu dienen, die vorherigen Testergebnisse aus dem Kurzzeitgedächtnis zu löschen. Die Ergebnisse wurden mittels Kappa Koeffizienten dargestellt und werden zur besseren Übersicht in den folgenden **Tabellen 14 und 15** aufgelistet.

Tabelle 14: Intrarater Reliabilität

Intrarater Reliabilität		
Test	Gruppe	Kappa Koeffizient
Malleolus medialis	Trainiert	0.68
	Untrainiert	0.59
Spina iliaca anterior superior (SIAS)	Trainiert	0.65
	Untrainiert	0.19
Spina iliaca posterior superior (SIPS)	Trainiert	0.54
	Untrainiert	0.49
Angulus inferior lateralis (AIL)	Trainiert	0.20
	Untrainiert	0.03

Angulus inferior lateralis (AIL)	Trainiert	0.07
	Untrainiert	0.20
Flexionstest im Sitzen	Trainiert	0.41
	Untrainiert	0.02

Tabelle 15: Interrater Reliabilität

Interrater Reliabilität		
Test	Gruppe	Kappa Koeffizient
Malleolus medialis	Trainiert	0.31
	Untrainiert	0.28
Spina iliaca anterior superior (SIAS)	Trainiert	0.24
	Untrainiert	-0.01
Spina iliaca posterior superior (SIPS)	Trainiert	0.08
	Untrainiert	0.15
Angulus inferior lateralis (AIL)	Trainiert	0.04
	Untrainiert	-0.01
Angulus inferior lateralis (AIL)	Trainiert	0.04
	Untrainiert	-0.01
Flexionstest im Sitzen	Trainiert	0.14
	Untrainiert	0.07

In **Tabelle 16** findet sich die Bewertung der Arbeit mit der QAREL-Checkliste.

Tabelle 16: The effect of training on the inter-examiner and intra-examiner reliability of the seated flexion test and assessment of pelvic anatomical landmarks with palpation (QAREL-Checkliste)

QAREL Checkliste der Studie 7		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?	X			
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?			X	
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X			
7.	Were raters blinded to additional cues that were not part of the test?		X		
8.	Was the order of examination varied?	X			
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?			X	
10.	Was the test applied correctly and interpreted appropriately?			X	
11.	Were appropriate statistical measures of agreement used?		X		
		Gesamt	7	0	3
					1

4.8. Inter-examiner and intra-examiner agreement for assessing simulated leg length inequality using palpation and observation during a standing assessment (Gibbons et al., 2002)

Die Untersuchung der Beinlänge als asymmetrisches Zeichen der SD wird in dieser Arbeit von Gibbons und seinem Kollegen untersucht. Acht StudentInnen im letzten Jahr der Osteopathie-

ausbildung wurden durch Zufallsverfahren aus einer Gruppe Freiwilliger ausgewählt. Die StudentInnen hatten die Möglichkeit, sich vor der Untersuchung untereinander abzusprechen. Dies sollte den standardisierten Ablauf der Untersuchung verbessern. Als Probandinnen, wurden 27 freiwillige Studentinnen der Victoria University of Melbourne rekrutiert. Ausschlusskriterien wie Schmerzen im unteren Rücken, Schmerzen in der unteren Extremität oder Probleme, länger als fünf Minuten still stehen zu können, wurden dabei angegeben. Die Datenerhebung fand über drei Tage verteilt statt. Es wurden je neun TeilnehmerInnen kreisförmig in einem großen Raum verteilt aufgestellt. Die acht UntersucherInnen wurden den ProbandInnen randomisiert zugeteilt. Durch eine statische Untersuchung der knöchernen Referenzpunkte Crista iliaca, SIPS, Trochanter major und der Glutealfalte im Seitenvergleich, sollte angegeben werden, ob beide Beine gleich lang waren oder rechts oder links ein längeres Bein vorhanden war. Ob oder wie eine Verblindung zu den Befunden der anderen TesterInnen stattgefunden hat, war unklar. Nachdem alle TesterInnen jeweils neun Probandinnen untersucht hatten, verließen diese den Raum. Es wurden entweder eine 0.5 cm hohe, eine 1cm hohe oder keine Erhöhung in die Schuhe der Probandinnen gelegt. Die TesterInnen wussten nichts über die Anwesenheit oder Abwesenheit einer Erhöhung, könnten jedoch die Probandinnen an Merkmalen wie Hautveränderungen oder der Körperhaltung wiedererkannt haben. Dies wurde dann solange wiederholt, bis alle 27 Probandinnen insgesamt sechsmal von jeder UntersucherIn getestet wurden.

Bei der interrater Reliabilität wurde ein Kappa Koeffizient von 0.27 bei keiner Erhöhung errechnet. Bei einer Erhöhung von 0,5cm war der Wert bei 0.28 und bei einer 1cm Erhöhung bei 0.19. Die intrarater Reliabilität brachte Werte von 0.35 bei keiner Erhöhung, 0.34 bei einer 0,5cm Erhöhung und 0.14 bei einer 1cm Erhöhung hervor. Die Auswertung mit der QAREL-Checkliste wird in **Tabelle 17** dargestellt.

Tabelle 17: Inter-examiner and intra-examiner agreement for assessing simulated leg length inequality using palpation and observation during a standing assessment (QAREL-Checkliste)

QAREL Checkliste der Studie 8		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?		X		
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?		X		
3.	Were raters blinded to the findings of other raters during the study?			X	

4.	Were raters blinded to their own prior findings of the test under evaluation?	X
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?	X
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X
7.	Were raters blinded to additional cues that were not part of the test?	X
8.	Was the order of examination varied?	X
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?	X
10.	Was the test applied correctly and interpreted appropriately?	X
11.	Were appropriate statistical measures of agreement used?	X
Gesamt		6 1 3 1

4.9. Reliabilitätsstudie des Befunds nach dem Mechanical Link:
Methodologische Studie über die Befunderhebung der Knochen und Gelenke der Extremitäten nach der Methode des Mechanical Link (Hafen, 2019)

Bei dem Konzept des Mechanical Link wird durch einen sogenannten Spannungstest, die Qualität des Gewebes überprüft. Dabei wird nach einer osteopathischen Läsion gesucht, die sich in einer Gewebeveränderung, wie zum Beispiel einer Fibrosierung, präsentiert. Die Untersuchung ist nach den TART Kriterien der SD eher eine Testung auf Gewebeveränderungen. Durch einen sanften Druck mit den Fingern in das Gewebe wird nach Spannungsveränderungen gesucht, die sich mit einer veränderten Dichte bemerkbar machen. In der Studie wurde die interrater- wie auch die intrarater Reliabilität überprüft. Vier TesterInnen, alle mit der Methode des Mechanical Link vertraut und in einer Vorstudie geschult, wurden rekrutiert. Diese führten 36 Tests an den Knochen und Gelenken der Extremitäten bei 40 freiwillig rekrutierten Probandinnen und Probanden durch. Hochgerechnet wurden daher insgesamt 11.520 Tests durchgeführt. Es wurden je vier TeilnehmerInnen, in unterschiedlichen Räumen von je einer TesterIn untersucht. Dann rotierten die TesterInnen, bis jede zu

untersuchende Person zweimal getestet wurde. Die UntersucherInnen hatten keine Einsicht in die Patientenakten oder in die Ergebnisse der anderen TesterInnen. Ob es eine Verblindung zu den eigenen Befunden gab, kam aus der Studie nicht hervor. Auch zusätzliche Hinweise mit Wiedererkennungswert, wie eine Hautrötung nach mehrmaligem Testen oder Narben und Tattoos, wurden nicht beachtet. Die Ergebnisse wurden mittels Kappa Koeffizienten berechnet und ergaben Werte von 0.61 für die interrater Reliabilität und 0.72 für die intrarater Reliabilität. Die Auswertung anhand der QAREL-Checkliste wird in der nachfolgenden **Tabelle 18** näher dargestellt.

Tabelle 18: Reliabilitätsstudie des Befunds nach dem Mechanical Link: Methodologische Studie über die Befunderhebung der Knochen und Gelenke der Extremitäten nach der Methode des Mechanical Link (QAREL-Checkliste)

QAREL Checkliste der Studie 9		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?			X	
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?				X
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X			
7.	Were raters blinded to additional cues that were not part of the test?		X		
8.	Was the order of examination varied?		X		
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?			X	
10.	Was the test applied correctly and interpreted appropriately?	X			
11.	Were appropriate statistical measures of agreement used?	X			
		Gesamt	6	2	2
					1

4.10. Interexaminer reliability study of a standardized myofascial diagnostic technique of the superior thoracic inlet (Hutchinson et al., 2017)

In dieser Arbeit wird eine myofasziale Untersuchung der oberen Thoraxapertur auf ihre interrater Reliabilität hin überprüft. Die hier untersuchte Körperregion wird auch als superior thoracic inlet bezeichnet und beschreibt den Bereich zwischen oberem Sternum, den Rippen eins bis zwei und den Brustwirbeln eins bis vier. Es kam ein diagnostisches Protokoll zur Anwendung, das Bewegungsmuster in drei Ebenen evaluieren sollte. Zwei TesterInnen, die sich im dritten Jahr der Osteopathie Ausbildung befanden, wurden zufällig rekrutiert. Diese waren zu den myofaszialen Mustern der ProbandInnen und zu den Befunden der anderen Testperson verblindet.

In der ersten Phase der Studie wurden 52 freiwillige StudentInnen rekrutiert. Diese Phase diente lediglich als Testphase zur Überprüfung der Übereinstimmung der Testergebnisse. Die Ein- und Ausschlusskriterien wurden spezifisch angegeben. Die Untersuchung selbst fand standardisiert und in randomisierter Reihenfolge statt. Mit einer vorher eingübten Grifftechnik wurde die obere Thoraxöffnung in drei Raumebenen auf myofasziale Spannung überprüft. Die Testrichtungen waren eine Rotation, eine Translation und eine Sagittalbewegung. Die Richtung, in die diese Bewegung leichter und besser geht, auch als side of ease angegeben, wurde als die Dysfunktionsrichtung definiert. In der ersten Phase war das Ziel, eine Übereinstimmung von über 80% zu erreichen und wenn nötig Anpassungen vorzunehmen. Da diese Übereinstimmung erreicht wurde, konnte mit der Phase zwei begonnen werden. Nun wurden insgesamt 82 neue ProbandInnen rekrutiert. Diese wurden erneut wie in Phase eins auf myofasziale Spannungsmuster hin überprüft. Die Ergebnisse der zweiten Phase, alle mit dem Kappa Koeffizienten angegeben, waren in der Rotationsrichtung 0.65, in der Translationstestung 0.59 und in der sagittalen Richtung 0.70. In der **Tabelle 19** wird die Studie mit der QAREL-Checkliste bewertet.

Tabelle 19: Tabelle 19: Interexaminer reliability study of a standardized myofascial diagnostic technique of the superior thoracic inlet (QAREL-Checkliste)

QAREL Checkliste der Studie 10		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?			X	
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?			X	
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X			
7.	Were raters blinded to additional cues that were not part of the test?	X			
8.	Was the order of examination varied?	X			
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?		X		
10.	Was the test applied correctly and interpreted appropriately?	X			
11.	Were appropriate statistical measures of agreement used?	X			
		Gesamt	7	1	1
					2

4.11. Reliability of physical examination to assess asymmetry of anatomical landmarks indicative of pelvic somatic dysfunction in subjects with and without low back pain (Kmita & Lucas, 2008)

Die Autoren dieser Arbeit untersuchten asymmetrische Befunde am Becken bei Subjekten mit und ohne low back pain. Es handelte sich um eine intrarater und interrater Reliabilitätsstudie. Die vier TesterInnen wurden an der University of Western Sydney rekrutiert. Zwei der UntersucherInnen waren erfahren und für mindestens 10 Jahre im Beruf, die beiden anderen

waren noch studierende OsteopathInnen im letzten Jahr des Studiums. Die 12 rekrutierten ProbandInnen wurden genauen Ein- und Ausschlusskriterien unterzogen. Dabei wurden asymptomatische und symptomatische Personen in die Studie inkludiert. Es wurde darauf geachtet, dass keine Tattoos oder Narben die Befunde der TesterInnen beeinflussen konnten. Die UntersucherInnen wurden vorher einem Training und einer Einweisung zur genauen Testdurchführung und zur besseren Übereinstimmung unterzogen. Der Untersuchungsablauf wurde unter Verblindung der TesterInnen zu den eigenen Befunden und den Befunden der anderen durchgeführt.

Die getesteten Personen wurden oberhalb und unterhalb des Beckens abgedeckt, sodass nur der Beckenbereich für die TesterInnen sichtbar war. Am Tag der Untersuchung waren nur neun der vorher zwölf rekrutierten ProbandInnen anwesend. Die Positionsuntersuchung der SIPS, des Iliosacralen sulcus und des AIL wurde in Bauchlage durchgeführt. Die Reihenfolge der TesterInnen erfolgte randomisiert. Nachdem die Testpersonen die Position in der Rückenlage einnahmen, erfolgte in einer weiteren Runde die Testung der SIAS und des Malleolus medialis. Die Dauer jeder Testung betrug zwischen einer und zwei Minuten. Zwischen den Untersuchungen fand eine Pause von fünf Minuten statt. Damit wurde versucht, eine theoretische Teststabilität zu generieren. Insgesamt wurden zwei komplett Untersuchungsrunden durchgeführt. Die Datenanalyse erfolgte mittels der Berechnung von Cohens Kappa und des Fleiss' Kappa Koeffizienten.

Die Ergebnisse der intrarater- und interrater Reliabilität werden in **Tabelle 20** dargestellt. Die Auswertung der Studie mit der QAREL-Checkliste wird in **Tabelle 21** präsentiert.

Tabelle 20: Ergebnisse der intrarater- und interrater Reliabilitätstestung

	Intrarater Reliabilität	Interrater Reliabilität
Test	Cohens Kappa	Fleiss' Kappa
SIPS	0.29 – 0.39	0.04
Sulcus sacralis	-0.28 – 0.83	-0.4
AIL (anterior-posterior)	0.29 – 0.44	-0.3
AIL (inferior-superior)	0.29 – 0.34	-0.01
SIAS	0.25 – 0.63	0.13
Malleolus medialis	0.20 – 1.0	0.17

Tabelle 21: Reliability of physical examination to assess asymmetry of anatomical landmarks indicative of pelvic somatic dysfunction in subjects with and without low back pain (QAREL-Checkliste)

QAREL Checkliste der Studie 11		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?	X			
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?				X
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X			
7.	Were raters blinded to additional cues that were not part of the test?	X			
8.	Was the order of examination varied?	X			
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?	X			
10.	Was the test applied correctly and interpreted appropriately?	X			
11.	Were appropriate statistical measures of agreement used?	X			
		Gesamt	10	0	0
					1

4.12. Accuracy of Anterior Superior Iliac Spine Symmetry Assessment by Routine Structural Examination (Lee et al., 2015)

Das Hauptziel, das diese Studie verfolgte, war zu evaluieren, wie hoch die Sensitivität und die Spezifität einer palpatorischen Testung der SIAS auf Asymmetrie ist. Weitere Punkte, die geklärt werden sollten, bezogen sich auf den Einfluss des dominanten Auges bei der Testung sowie das Trainingslevel, anhand des Erfahrungswertes der TesterInnen gemessen. Die TeilnehmerInnen waren per Email rekrutierte studierende und fertige Osteopathen. Die

insgesamt 147 StudienteilnehmerInnen führten die Untersuchung an drei befestigten Plastikmodellen durch. Eines der Beckenskelette wurde mit beiden SIAS auf gleicher Höhe eingestellt, die beiden anderen einmal mit der linken SIAS 5mm superior und einmal mit SIAS links 10mm superior. Die Skelettteile wurden mit einem Tuch abgedeckt, damit eine mögliche Asymmetrie nicht sofort erkennbar war.

Vor der Testung wurde das dominante Auge der TesterInnen mittels eines Eigentests evaluiert. Auch wurde der Untersuchungsablauf und die Testung selber besprochen, um einen möglichst standardisierten Ablauf zu erreichen.

Begonnen wurde mit der Testung auf der rechten Seite mit dem rechten Auge in der Mitte ausgerichtet. Danach sollten alle TesterInnen von Links untersuchen, mit dem linken Auge in der Mittellinie ausgerichtet. Die Ergebnisse wurden mittels einer Statistiksoftware ausgearbeitet und ergaben eine Sensitivität von 82,5% für die Abweichung von 5mm und 91,7% für die eine Abweichung von 10mm. Die Spezifität betrug dabei nur 31,0%. Es wurden keine signifikanten Unterschiede zwischen rechts und links dominanten Augen gefunden. Nur bei der 5mm Abweichung konnten die erfahreneren TesterInnen besser abschließen als die noch Studierenden. Die Ergebnisse der Auswertung mit der QAREL-Checkliste werden in der **Tabelle 22** dargestellt.

Tabelle 22: Accuracy of Anterior Superior Iliac Spine Symmetry Assessment by Routine Structural Examination (QAREL-Checkliste)

QAREL Checkliste der Studie 12		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?	X			
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?			X	
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?			X	
7.	Were raters blinded to additional cues that were not part of the test?	X			

8.	Was the order of examination varied?	X			
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?	X			
10.	Was the test applied correctly and interpreted appropriately?	X			
11.	Were appropriate statistical measures of agreement used?	X			
	Gesamt	4	3	2	2

4.13. Counterstrain and traditional osteopathic examination of the cervical spine compared (McPartland & Goodridge, 1997)

In dieser interrater Reliabilitätsstudie wird die Untersuchung nach dem Konzept des Strain-counterstrain (SCS) mit der Untersuchung nach den klassischen TART Kriterien verglichen. SCS ist eine von Dr. Lawrence Jones entwickelte Methode, deren Diagnose auf der Palpation von myofaszialen Spannungen beruht. Spezifisch definierte Punkte werden dabei Gelenksdysfunktionen zugeordnet, diese werden als sogenannte Tenderpoints bezeichnet. Die Therapie nach der SCS Methode beruht auf dem Konzept der indirekten Behandlung. Diese kleinen Zonen, meist druckempfindlich für den Patienten, sind bei der Palpation als Festigkeit zu fühlen. In dieser Studie beschränkte sich der zu untersuchende Bereich auf die obere cervicale Region C0 – C3. Es wurden 18 ProbandInnen rekrutiert, sieben davon mit chronischen Nackenbeschwerden und elf asymptatisch. Es wurden genaue Ausschlusskriterien definiert wie akute Entzündungen, Nervenwurzelreizungen, operative Eingriffe an der HWS, Schwindel und andere Erkrankungen, die das Ergebnis der Studie verzerren könnten. Die UntersucherInnen, beide mit Erfahrungswerten zwischen 10-40 Jahren osteopathischer Praxis, wurden von den Befunden der jeweils anderen Person verblindet. Die Reihenfolge der Untersuchung unterlag einer Randomisierung.

Für den Untersuchungsgang wurde je eine Testperson in einem Raum in Rückenlage positioniert. Eine TesterIn betrat den Raum und untersuchte die Testperson einmal nach der SCS Methode und anschließend nach den TART Kriterien. Die Testung auf Asymmetrie wurde in dieser Studie jedoch nicht miteinbezogen, sondern lediglich die Gewebebeschaffenheit, das Bewegungsausmaß und die Gewebeempfindlichkeit. Um herauszufinden, ob erfahrene Osteopathen bessere Resultate erzielen als Studierende, wurde eine zweite Phase mit 18

StudentInnen, die sich im zweiten Jahr des Osteopathiestudiums befanden durchgeführt. Die ProbandInnen dieser Phase waren erneut StudentInnen, die wieder zufällig rekrutiert wurden.

Die Ergebnisse der Studie wurden mittels Kappa Koeffizienten berechnet. Der Wert bei den erfahreneren TesterInnen mit der SCS Methode bei der symptomatischen Gruppe betrug 0.45 und bei der asymptomatischen Gruppe 0.19, was einen totalen Kappa Koeffizienten von 0.29 entspricht. Bei den TART Kriterien wurde bei der symptomatischen Gruppe ein Kappa Wert von 0.35 errechnet, bei der asymptomatischen Gruppe ein Wert von 0.34, was einem totalen Wert von 0.39 entspricht. Die Methode nach SCS hat demnach bei den echten PatientInnen mit Symptomen etwas besser abgeschlossen wie die TART Untersuchung. Bei den asymptomatischen ProbandInnen war dies genau umgekehrt. Die Studentengruppe hat insgesamt weniger gut abgeschlossen als die erfahreneren TesterInnen, ein direkter Vergleich war jedoch nicht möglich, da die Testung hier nur an asymptomatischen Testpersonen durchgeführt wurde. Die Ergebnisse der Auswertung mit der QAREL-Checkliste werden in **Tabelle 23** dargestellt.

Tabelle 23: Counterstrain and traditional osteopathic examination of the cervical spine compared (QAREL-Checkliste)

QAREL Checkliste der Studie 13		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?		X		
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?		X		
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X			
7.	Were raters blinded to additional cues that were not part of the test?	X			
8.	Was the order of examination varied?	X			
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when		X		

determining the suitability of the time-interval between repeated measures?				
10.	Was the test applied correctly and interpreted appropriately?	X		
11.	Were appropriate statistical measures of agreement used?	X		
	Gesamt	7	1	1 2

4.14. Inter-examiner and intra-examiner agreement for assessing sacroiliac anatomical landmarks using palpation and observation: pilot study (O'Haire & Gibbons, 2000)

In dieser Pilotstudie wurde die interrater- und intrarater Reliabilität für die Palpation von knöchernen Referenzpunkten am Becken untersucht. Es wurden zehn asymptomatische Studentinnen per Zufallsverfahren ausgewählt. Ausgeschlossen wurden Probandinnen, die nicht für längere Zeit auf dem Bauch liegen konnten oder Wiedererkennungsmerkmale wie Tattoos, Piercings oder Narben aufwiesen. Diese sollten zur Untersuchung mit identischer Unterwäsche erscheinen. Die zehn TesterInnen, alles Studierende im letzten Jahr der Osteopathie Ausbildung, erhielten eine Stunde Training zum Standardisieren des Untersuchungsablaufs. Zehn Tische wurden in einem großen Raum mit jeweils einem Meter Abstand aufgestellt. Die Probandinnen wurden zum Teil abgedeckt, um den Wiedererkennungswert weiter zu reduzieren. Diese mussten insgesamt über 2,5 Stunden auf der Stelle liegen bleiben, mit der Anweisung, sich möglichst wenig zu bewegen. Die UntersucherInnen wurden per Randomisierung den Tischen zugewiesen. Dort wurde ihnen mit einer Karte zum Ankreuzen ein Referenzpunkt zum Untersuchen angegeben. Nach jedem Test rotierten die TesterInnen im Uhrzeigersinn zur nächsten Probandin. Sobald alle zehn TesterInnen alle Probandinnen untersucht hatten, verließen sie den Raum und wurden einem Gedächtnistest unterzogen, um die Erinnerung an die vorher gefundenen Ergebnisse zu reduzieren. Die Referenzpunkte SIPS, sulcus sacralis und AIL wurden damit nacheinander auf Stellung überprüft. Bei jeder Probandin war nun eine neue Karte, die den Referenzpunkt vorgab, der nun zu untersuchen war. Dieses Prozedere wurde solange wiederholt, bis alle Referenzpunkte an allen Probandinnen von allen TesterInnen viermal untersucht wurden. Der Kappa Koeffizient für die intrarater Reliabilität betrug für die Palpation der SIPS 0.33, für den sulcus sacralis 0.02 und für den AIL 0.21. Die Ergebnisse der interrater Reliabilität waren für

die SIPS 0.04, für den sulcus sacralis 0.07 und für den AIL 0.08. In **Tabelle 24** wird die Beurteilung der Studie anhand der QAREL-Checkliste dargestellt.

Tabelle 24: Inter-examiner and intra-examiner agreement for assessing sacroiliac anatomical landmarks using palpation and observation: pilot study (QAREL-Checkliste)

QAREL Checkliste der Studie 14		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?	X			
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?			X	
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X			
7.	Were raters blinded to additional cues that were not part of the test?		X		
8.	Was the order of examination varied?	X			
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?			X	
10.	Was the test applied correctly and interpreted appropriately?	X			
11.	Were appropriate statistical measures of agreement used?	X			
		Gesamt	8	1	1
					1

4.15. Inter-examiner reliability of palpation for tissue texture abnormality in the thoracic paraspinal region (Paulet & Fryer, 2009)

Die Autoren dieser Arbeit gehen davon aus, dass die Gewebebeschaffenheit ein wichtiges Zeichen zur Bestimmung einer somatischen Dysfunktion ist. In dieser Studie wurde die interrater Reliabilität der Palpation von Gewebsveränderungen an der thorakalen paraspinalen

Region untersucht. Es wurden zehn ProbandInnen von der Victoria University of Melbourne, Australien rekrutiert. Alle TeilnehmerInnen hatten eine frühere milde Episode von Schmerzen im thorakalen Wirbelsäulenbereich. Die Schmerzen sollten an der numerischen Analogskala nicht mehr als vier von zehn Punkten überschritten haben. Ausschlusskriterien waren momentane Schmerzen in der BWS, Skoliosen, chronische Schmerzsyndrome oder eine Unfähigkeit, länger als 20 Minuten auf dem Bauch liegen zu können. Die UntersucherInnen, alles Studierende im fünften Jahr des Osteopathiestudiums, wurden ebenfalls an der Victoria University rekrutiert. Um die individuellen Variationen der Testungen und Interpretationen zu reduzieren, erhielten alle TesterInnen eine Woche vor dem Beginn der Studie ein einstündiges Training zur Standardisierung.

Die getesteten Personen wurden in Bauchlage liegend auf jeweils einer Behandlungsbank in einem großen Raum verteilt. Die Untersuchungsleiterin markierte vier verschiedene Stellen paravertebral der thorakalen Wirbelsäule. Eine der markierten Stellen wies eine Veränderung der Gewebebeschaffenheit auf, was einem positiven Befund entsprach. Die TesterInnen hatten keinerlei Informationen bezüglich des klinischen Status der Testpersonen. Die Untersuchung dauerte eine Minute, die TesterInnen sollten die Region, die ihnen am auffälligsten erschien, ausfindig machen. Eine vermehrte Hautrötung oder andere Reaktionen am Patienten nach mehreren Testungen wurden nicht weiter in der Studie erwähnt und könnten ein Bias hervorgerufen haben. Die statistische Auswertung ergab für die interrater Reliabilität einen Kappa Koeffizienten von 0.26, was einer schwachen Übereinstimmung der Reliabilität entspricht. Die Bewertung der Arbeit mit der QAREL-Checkliste wird in **Tabelle 25** zusammengefasst.

Tabelle 25: Inter-examiner reliability of palpation for tissue texture abnormality in the thoracic paraspinal region (QAREL-Checkliste)

QAREL Checkliste der Studie 15		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?		X		
4.	Were raters blinded to their own prior findings of the test under evaluation?			X	

5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?						X
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?					X	
7.	Were raters blinded to additional cues that were not part of the test?					X	
8.	Was the order of examination varied?						X
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?					X	
10.	Was the test applied correctly and interpreted appropriately?				X		
11.	Were appropriate statistical measures of agreement used?			X			
		Gesamt	5	1	2	3	

4.16. The assessment of pelvic landmarks using palpation: A reliability study of undergraduate students (Rajendran & Gallagher, 2011)

In dieser Arbeit wurde erneut die Interrater Reliabilität der statischen Untersuchung anatomischer Referenzpunkte des Beckens untersucht. Die TesterInnen waren 20 Studierende der European School of Osteopathy im dritten Jahr der Ausbildung. Vier asymptomatische Testpersonen, ebenfalls StudentInnen im dritten Jahr des Studiums, wurden als Freiwillige rekrutiert. Die TesterInnen waren gegeneinander verblindet und hatten keine näheren Informationen zu den Testpersonen. Da diese aber von derselben Schule waren, könnten sich die TeilnehmerInnen untereinander bereits persönlich bekannt gewesen sein.

Der Testablauf selber wurde nach dem klassischen Modell von Mitchell durchgeführt. Dabei wurde bei einer in Bauchlage befindenden Testperson, eine statische Untersuchung auf Symmetrie oder Asymmetrie durchgeführt. Die überprüften Referenzpunkte waren dabei der sulcus sacralis und der AIL. Die LWS wurde durch einen sogenannten Springing-Test ebenfalls in die Untersuchung miteinbezogen. Die Testreihenfolge wurde dabei randomisiert. Es erfolgte kein Übereinstimmungstraining, da die Testung im Curriculum der Schule enthalten war und die TesterInnen diesen Teil der Ausbildung bereits abgeschlossen hatten. Eine statistische Analyse mit dem Kappa Koeffizienten ergab allgemein einen Wert unter 0.4 und somit eine

schwache Reliabilität. Die Studie wurde mit der QAREL-Checkliste bewertet und in **Tabelle 26** dargestellt.

Tabelle 26: The assessment of pelvic landmarks using palpation: A reliability study of undergraduate students (QAREL-Checkliste)

QAREL Checkliste der Studie 16		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?		X		
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?			X	
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?			X	
7.	Were raters blinded to additional cues that were not part of the test?	X			
8.	Was the order of examination varied?	X			
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?		X		
10.	Was the test applied correctly and interpreted appropriately?	X			
11.	Were appropriate statistical measures of agreement used?	X			
		Gesamt	6	1	2
					2

4.17. Intra-examiner and inter-examiner reliability of a positional diagnostic screen for the lumbar spine (Spring et al., 2001)

Die Untersuchung der LWS spielt in jeder manuellen Methode und somit auch in der Osteopathie eine große Rolle. Verschiedene Untersuchungsverfahren wie die Observation, eine statische Palpation und eine Bewegungstestung kommen dabei zum Einsatz. In dieser Studie

wird die inter- und intrarater Reliabilität einer Stellungsdiagnostik der LWS anhand eines Palpationstests überprüft. Zehn asymptomatische weibliche Freiwillige wurden aus der Studentenpopulation der Victoria University Melbourne rekrutiert. Ein- und Ausschlusskriterien wurden genau definiert und alle Probandinnen wurden aufgefordert, bei der Untersuchung die gleiche Unterwäsche und eine Kapuze zu tragen. Merkmale wie Tattoos, Piercings und Geburtsmale zählten ebenfalls zu den Ausschlusskriterien. Die TesterInnen, insgesamt zehn Studierende, befanden sich alle im letzten Jahr des Osteopathiestudiums. Diese erhielten vor dem Untersuchungsgang eine Trainingsstunde zur korrekten Testausführung und Interpretation. Die Probandinnen sowie auch die TesterInnen wurden randomisiert zu den zehn vorher aufgestellten Behandlungsliegen zugewiesen. Die Probandinnen wurden zuerst im Sitzen, dann in Bauchlage und dann in einer abgewandelten Bauchlage mit abgestützten Ellenbogen untersucht. Getestet wurde der vierte Lendenwirbel, da dieser am zuverlässigsten durch Palpation ausfindig zu machen war. Durch die verschiedenen Positionen konnte jede Testerin und jeder Tester den Wirbel in Flexion (Sitz), in Neutralstellung (Bauchlage), sowie in Extension (Ellenbogenstütz) auf Stellung überprüfen. Nach dem Untersuchungsgang verließen die TesterInnen den Raum für fünf Minuten und wurden einem Gedächtnistraining unterzogen, um die vorherigen Befunde aus dem Kurzzeitgedächtnis zu löschen. Danach erfolgten weitere zwei Durchgänge, wobei die ProbandInnen jeweils randomisiert den Bänken zugewiesen wurden.

Die statistische Berechnung durch den Kappa Koeffizienten ergab einen Wert von 0.04 für die intra- wie auch für die interrater Reliabilität. Es gab daher nur eine schwache Übereinstimmung bei den Befunden der TesterInnen. Die Bewertung der Studie anhand der QAREL-Checkliste wird in **Tabelle 27** dargestellt.

Tabelle 27: Intra-examiner and inter-examiner reliability of a positional diagnostic screen for the lumbar spine (QAREL-Checkliste)

QAREL Checkliste der Studie 17		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?	X			

5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?	X	
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X	
7.	Were raters blinded to additional cues that were not part of the test?	X	
8.	Was the order of examination varied?	X	
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?	X	
10.	Was the test applied correctly and interpreted appropriately?	X	
11.	Were appropriate statistical measures of agreement used?	X	
Gesamt		9	0
		1	1

4.18. The effects of experience on the inter-reliability of osteopaths to detect changes in posterior superior iliac spine levels using a hidden heel wedge (Sutton et al., 2013)

Bei dieser Studie wurde erneut die interrater Reliabilität einer Palpation der SIPS auf Asymmetrie durchgeführt. Die AutorInnen der Arbeit wollten diesmal überprüfen, ob der Erfahrungswert, gemessen an Berufsjahren, Einfluss auf die Testergebnisse habe. Es wurden 40 UntersucherInnen von dem British College of Osteopathic Medicine auf freiwilliger Basis per Email rekrutiert. Eine zusätzliche, auf dem gleichen Weg rekrutierte Studentin wurde als Beckenmodell ausgewählt. Diese durfte eine Woche vor der eigentlichen Studie keine osteopathischen oder anderen manuellen Behandlungen erhalten oder sich intensivem körperlichem Training aussetzen. Die TesterInnen wurden in drei Gruppen eingeteilt. Die erste Gruppe bestand aus 15 StudentInnen im dritten Jahr und die zweite Gruppe aus 15 StudentInnen im vierten Jahr der Osteopathieausbildung. Gruppe drei bestand aus zehn erfahrenen Osteopathen, die zwischen 5-25 Jahren Berufserfahrung vorweisen konnten. Alle TeilnehmerInnen erhielten eine zehnminütige Trainingseinheit vor dem Beginn der Studie, um eine Standardisierung des Tests zu erreichen. Die TesterInnen untersuchten nacheinander die Testperson im Stand und gaben dabei die Höhe der SIPS an. Nach einer vorher festgelegten Reihenfolge wurde eine 5mm Erhöhung unter einen Fuß gelegt. Insgesamt wurde die

Testperson neunmal von jeder Testerin und jedem Tester untersucht. Diesen war aber nicht bewusst, wann und unter welchem Fuß die Erhöhung eingesetzt wurde. Alle Untersuchungsgruppen erreichten keinen Kappa Koeffizienten über 0.4 (0.025 – 0.065), was nur einem schwachen Ergebnis der intrarater Reliabilität entspricht. In **Tabelle 28** werden die Ergebnisse der Auswertung mit der QAREL-Checkliste dargestellt.

Tabelle 28: The effects of experience on the inter-reliability of osteopaths to detect changes in posterior superior iliac spine levels using a hidden heel wedge (QAREL-Checkliste)

QAREL Checkliste der Studie 18		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?			X	
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?			X	
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X			
7.	Were raters blinded to additional cues that were not part of the test?	X			
8.	Was the order of examination varied?		X		
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?			X	
10.	Was the test applied correctly and interpreted appropriately?	X			
11.	Were appropriate statistical measures of agreement used?	X			
		Gesamt	6	1	2
					2

4.19. Interexaminer reliability of three methods of combining test results to determine side of sacral restriction, sacral base position, and innominate bone position (Tong et al., 2006)

Viele Studien wurden bisher veröffentlicht, die sich mit der Reliabilität einer sacroiliacalen Gelenksdysfunktion auseinandergesetzt haben. Die meisten beziehen sich auf einen Test oder eine Methode. In dieser interrater Reliabilitätsstudie wurden drei verschiedene Kombinationen von Tests untersucht. Diese Tests sollten die Seite der Dysfunktion ausfindig machen, die Position der Sakrumbasis und die Stellung des Iliums bestimmen. Innerhalb von vier Monaten, untersuchten zwei TesterInnen 24 ProbandInnen mit LBP. Eine der TesterInnen wusste um die Patientengeschichte, die zweite Person war dagegen verblindet. Insgesamt wurden drei Tests zur Identifikation der Seite einer Dysfunktion benutzt (seated flexion test, standing stork test, standing flexion test), zwei Tests für die Positionsbestimmung der Sakrumbasis (trunk flexion, trunk extension) und zwei Tests für das Ilium (SIAS symmetry, malleolus medialis symmetry). Das Ziel dieser Arbeit war es, drei kombinierte Diagnostikmethoden miteinander zu vergleichen. Methode eins kombinierte nur die Tests mit den besten interrater Koeffizient Ergebnissen. In der Methode zwei musste mindestens ein Test ein positives Ergebnis zeigen, damit von einer Dysfunktion ausgegangen werden kann. In Methode drei müssen alle Testergebnisse positiv sein, um von einer Dysfunktion ausgehen zu können. Die Ergebnisse zeigten das beste Ergebnis bei Methode eins mit einem Kappa Koeffizienten für die Iliumposition von 0.47, für die Seite der Dysfunktion von 0.32 und für die Sakrumbasis von 0.08. Die Seite der Dysfunktion kann somit laut dieser Arbeit am besten mit dem „standing stork test“ gefunden werden, die Sakrumbasisstellung wird mit einem Wert von 0.37 am besten mit dem „trunk flexion“ test beurteilt und die Iliumposition mit der Stellungsdiagnose des malleolus medialis mit einem Kappa Koeffizienten von 0.21. Die Ergebnisse der Auswertung mit der QAREL-Checkliste werden in **Tabelle 29** dargestellt.

Tabelle 29: Interexaminer reliability of three methods of combining test results to determine side of sacral restriction, sacral base position, and innominate bone position (QAREL-Checkliste)

QAREL Checkliste der Studie 19		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			

2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?		X		
3.	Were raters blinded to the findings of other raters during the study?			X	
4.	Were raters blinded to their own prior findings of the test under evaluation?				X
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?				X
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?		X		
7.	Were raters blinded to additional cues that were not part of the test?			X	
8.	Was the order of examination varied?			X	
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?			X	
10.	Was the test applied correctly and interpreted appropriately?		X		
11.	Were appropriate statistical measures of agreement used?			X	
		Gesamt	4	1	4
					2

4.20. Inter-examiner and intra-examiner reliability of the standing flexion test (Vincent-Smith & Gibbons, 1999)

Der Flexionstest im Stand findet zur Beurteilung einer Dysfunktion im Iliosakralgelenk seine Anwendung. Die Reliabilität dieses Testes zu untersuchen war das Ziel dieser Studie. Die neun TesterInnen, alles OsteopathInnen, wurden aus einem Pool von Freiwilligen per Zufallsverfahren ausgesucht. Diese erhielten vor der eigentlichen Testung eine Trainingseinheit zum Standardisieren des Untersuchungsprotokolls. Die 14 Testpersonen, wurden aus einer Liste von Freiwilligen ebenfalls per Randomisierung rekrutiert. Ausschlusskriterien waren LBP, Schmerzen in den unteren Extremitäten oder Operationen an der LWS oder an der Beckenregion. Einschlusskriterien waren eine freie Flexionsbewegung des Rumpfes mit Knieextension. Fünf der TeilnehmerInnen wurden ausgeschlossen, da sie dieses Kriterium nicht erfüllten. Somit blieben noch neun der ursprünglichen 14 TeilnehmerInnen übrig. Alle TeilnehmerInnen sollten ein ähnliches T-Shirt und Shorts tragen, um den

Wiedererkennungswert zu reduzieren. Alle teilnehmenden Personen wurden in verschiedene Räume aufgeteilt und führten dort den Flexionstest im Stand zehnmal hintereinander aus, um falsch-positive Ergebnisse zu reduzieren. Die TesterInnen wurden randomisiert den Räumen zugewiesen und führten jeweils den Test an den Subjekten durch. Die Testpersonen wurden im Anschluss ebenfalls per Zufall in einen anderen Raum gebracht und das Prozedere begann von vorne. Bei der Testausführung wurde auf ein exaktes Protokoll geachtet, von der Stellung der Füße bis zur Position der TesterInnen. Die interrater Reliabilität wies einen Kappa Koeffizienten von 0.052 auf, die intrarater Reliabilität einen deutlich besseren Wert von 0.46. Die Auswertung mit der QAREL-Checkliste ist in der **Tabelle 30** dargestellt.

Tabelle 30: Inter-examiner and intra-examiner reliability of the standing flexion test (QAREL-Checkliste)

QAREL Checkliste der Studie 20		Ja	Nein	Unklar	N/A
Punkt	Frage				
1.	Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied?	X			
2.	Was the test performed by raters who were representative of those whom the authors intended the results to be applied?	X			
3.	Were raters blinded to the findings of other raters during the study?	X			
4.	Were raters blinded to their own prior findings of the test under evaluation?	X			
5.	Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated?				X
6.	Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design?	X			
7.	Were raters blinded to additional cues that were not part of the test?		X		
8.	Was the order of examination varied?	X			
9.	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures?			X	
10.	Was the test applied correctly and interpreted appropriately?	X			
11.	Were appropriate statistical measures of agreement used?	X			
		Gesamt	8	0	2
					1

4.21. Zusammenfassung der Ergebnisse

Um den Überblick zu bewahren, werden nun folgend alle Ergebnisse der analysierten Studien in einer Tabelle zusammengefasst. Als erstes werden die AutorIn oder die AutorInnen der Studie in einer Spalte angegeben. In der zweiten Spalte werden die untersuchten Kriterien mit einer Abkürzung eingetragen, wobei das A für „asymmetry“ steht, das TTC für „tissue texture changes“, das R für „range of motion“ und TART für alle Kriterien, die untersucht wurden. Anschließend werden die jeweilige Körperregion und die Studienart beschrieben. Die Ergebnisse werden, soweit dies möglich ist, als Kappa-Wert angegeben. Nur bei einer Studie wurden die Sensitivität und die Spezifität als Prozentangabe bei den Ergebnissen präsentiert. Zuletzt wird die Übereinstimmung anhand der schon beschriebenen Kappa-Interpretationstabelle von schwach bis nahezu perfekt eingeordnet. Eine Zusammenfassung der Ergebnisse wird in **Tabelle 31** dargestellt.

Tabelle 31: Zusammenfassung der Ergebnisse

AutorIn	Kriterien	Körperregion	Studienart	Ergebnisse	Reliabilität
Benggaard et al. (2012)	A	Becken/SIAS	Interrater	0.028	Schwach
Comeaux (2005)	TTC, R	Wirbelsäule C2 - Th8	Interrater	0.12 - 0.56	Schwach bis ausreichend
Consorti et al. (2018)	TART	Sakrum	Interrater	Gesamt 0.17	Schwach
Degenhardt et al. (2005)	TART	LWS L1 – L4	Interrater (Pre- /posttraining)	Gewebeempfindlichkeit von 0.32 auf 0.68 Gewebebeschaffenheit von 0.12 auf 0.45 Bewegungsausmaß von 0.10 auf 0.20 Asymmetrie von 0.17 auf 0.34	Stark Moderat Schwach Ausreichend
Degenhardt et al. (2010)	TART	LWS L1 – L4	Interrater (Pre- /posttraining)	Gewebeempfindlichkeit von 0.61 auf 0.88 Gewebebeschaffenheit von 0.23 auf 0.55 Bewegungsausmaß von 0.30 auf 0.50 Asymmetrie von 0.17 auf 0.56 auf 0.72	Nahezu Perfekt Moderat Moderat Stark
Fryer (2006)	A	Malleolus medialis	Interrater	Alle Probandinnen 0.22 Nur Probandinnen mit Asymmetrie 0.94	Ausreichend Nahezu Perfekt
			Intrarater	Alle Probandinnen 0.31 Nur Probandinnen mit Asymmetrie 0.94	Ausreichend Nahezu Perfekt

Fryer et al. (2005)	A, R	Becken	Interrater Pre- /posttraining Intrarater Pre- /posttraining	Knochenpunkte Trainiert 0.04 – 0.31 Untrainierte -0.01 – 0.28 Flexionstest im Sitzen Trainiert 0.14 Untrainiert 0.07 Knochenpunkte Trainiert 0.07 – 0.68 Untrainiert 0.03 - 0.59 Flexionstest im Sitzen Trainiert 0.41 Untrainiert 0.02	Schwach bis ausreichend Schwach bis ausreichend Schwach Schwach Schwach bis stark Schwach bis moderat Moderat Schwach
Gibbons et al. (2002)	A	Becken, Beine	Interrater Intrarater	Keine Erhöhung 0.27 Mit 0,5cm Erhöhung 0.28 Mit 1cm Erhöhung 0.19 Keine Erhöhung 0.35 Mit 0,5cm Erhöhung 0.34 Mit 1cm Erhöhung 0.14	Ausreichend Ausreichend Schwach Ausreichend Ausreichend Schwach
Hafen (2019)	TTC	Extremitäten	Interrater Intrarater	0.61 0.72	Stark Stark
Hutchinson et al. (2017)	R	Obere Thoraxapertur	Interrater	Rotation 0.65, Translation 0.59 Sagittal 0.70	Stark Moderat Stark
Kmita und Lucas (2008)	A	Becken	Interrater Intrarater	-0.04 – 0.17 -0.028 – 0.63	Schwach Schwach bis nahezu Perfekt
Lee et. al. (2015)	A	Becken	Sensitivität Spezifität	5mm 82,5% 10mm 91,7% 31.0%	
McPartland & Goodridge (1997)	TTC und alle TART Kriterien	HWS	Interrater	SCS symptomatisch 0.45 SCS asymptomatisch 0.19 TART Kriterien Symptomatisch 0.35 Asymptomatisch 0.34	Moderat Schwach Ausreichend Ausreichend
O'Haire & Gibbons (2000)	A	Becken	Interrater Intrarater	0.04 – 0.08 0.02 – 0.33	Schwach Schwach bis ausreichend
Paulet & Fryer (2009)	TTC	BWS	Interrater	0.26	Schwach
Rajendran & Gallagher (2011)	A	Becken	Interrater	< 0.4 für alle Ergebnisse	Schwach
Spring et al. (2001)	A, R	LWS	Interrater/Intrarater	0.04	Schwach
Sutton et al. (2013)	A	Becken	Interrater	0.025 – 0.065	Schwach
Tong et al. (2006)	A, R	Becken	Interrater	Iliumposition 0.47 Sakrumbasisstellung 0.08	Moderat Schwach

				Seite der Dysfunktion 0.32	Ausreichend
Vincent-Smith & Gibbons (1991)	R	Becken	Interrater	0.052	Schwach
			Intrarater	0.46	Moderat

Um einen Überblick über die QAREL-Checklisten Ergebnisse zu bewahren, wurde in **Tabelle 32** eine Zusammenfassung aller Ergebnisse dargestellt.

Tabelle 32: Zusammenfassung der Ergebnisse aller QAREL-Checklisten

Autor	Ja	Nein	Unklar	N/A
Bengaard et al. (2012)	4	3	2	2
Comeaux et al. (2005)	6	2	1	2
Consorti et al. (2018)	7	0	2	2
Degenhardt et al. (2005)	6	2	1	2
Degenhardt et al. (2010)	8	0	1	2
Fryer (2006)	9	0	1	1
Fryer et al. (2005)	7	0	3	1
Gibbons et al. (2002)	6	1	3	1
Hafen (2019)	6	2	2	1
Hutchinson et al. (2017)	7	1	1	2
Kmita und Lucas (2008)	10	0	0	1
Lee et. al. (2015)	4	3	2	2
McPartland & Goodridge (1997)	7	1	1	2
O'Haire & Gibbons (2000)	8	1	1	1
Paulet & Fryer (2009)	5	1	2	3
Rajendran & Gallagher (2011)	6	1	2	2
Spring et al. (2001)	9	0	1	1
Sutton et al. (2013)	6	1	2	2
Tong et al. (2006)	4	1	4	2
Vincent-Smith & Gibbons (1991)	8	0	2	1
Gesamt	133	20	34	33

Insgesamt wurde die Asymmetrie mit 42% am häufigsten untersucht. Die Bewegungseinschränkung steht an zweiter Stelle mit 28%, gefolgt von der Gewebeveränderung

mit 19% und an letzter Stelle die Gewebeempfindlichkeit mit 11%. Dies wurde anhand eines Kreisdiagramms in **Abbildung 6** dargestellt. In **Abbildung 7** werden die Ergebnisse von schwach bis nahezu perfekt anhand einer Häufigkeitsverteilung zur besseren Übersicht abgebildet. Durch die **Abbildung 8** wird nochmals klar, wie sich die Verteilung der getesteten Körperregionen darstellt. Das Becken wird dabei mit 52% am häufigsten untersucht, die Wirbelsäule mit 29% am zweithäufigsten, die untere Extremität befindet sich mit 14% an dritter Stelle und andere Körperregionen werden mit 5% seltensten begutachtet. Untersuchungen an der oberen Extremität einschließlich der Schulterregion gab es keine.

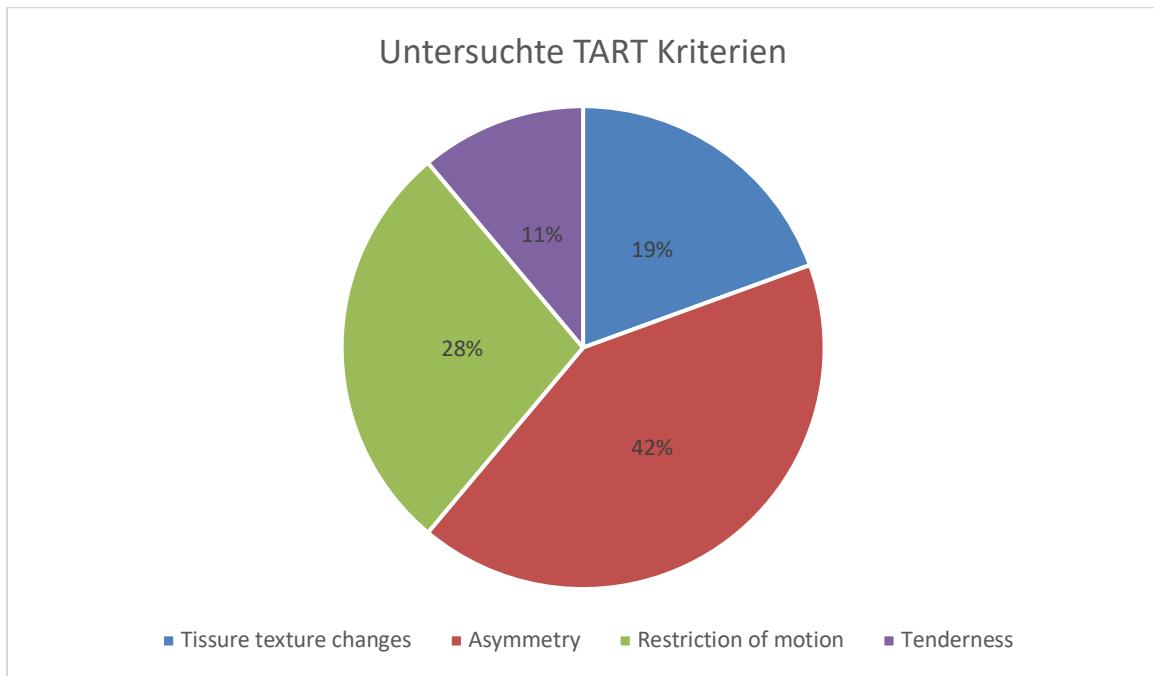


Abbildung 7: Kreisdiagramm aller untersuchten TART-Kriterien in Prozentangaben (eigene Darstellung)

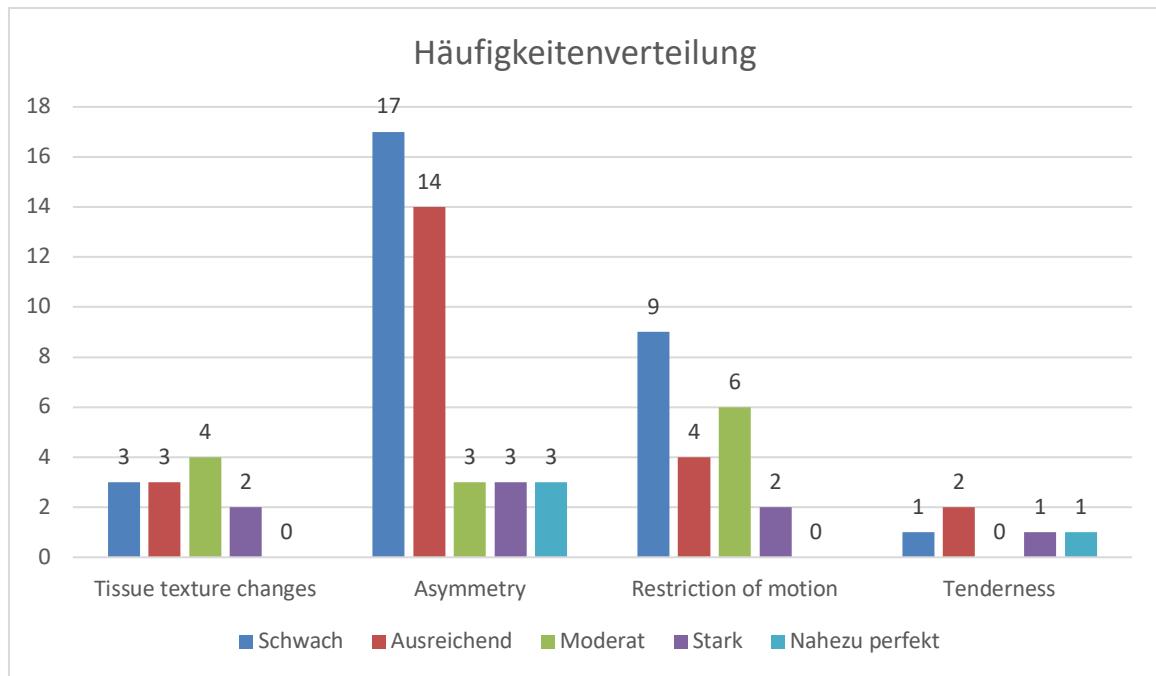


Abbildung 8: Diagramm zur Darstellung der Häufigkeitenverteilung (eigene Darstellung)

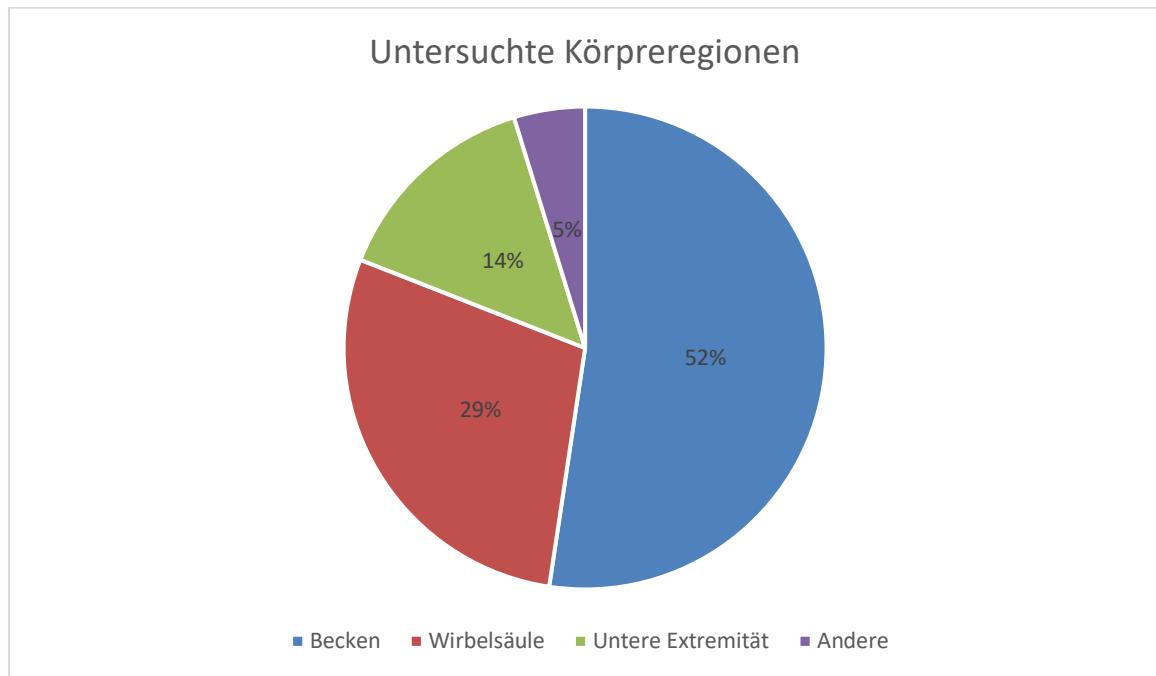


Abbildung 9: Kreisdiagramm zur Darstellung der untersuchten Körperregionen (eigene Darstellung)

5. Diskussion

Die 20 mit der QAREL-Checkliste analysierten Studien, konnten durchwegs eine gute bis sehr gute methodologische Qualität aufweisen. Die erste Frage der QAREL-Checkliste, ob die Studie auch an repräsentativen Personen durchgeführt wurde, konnte bis auf wenige Abweichungen mit einem „Ja“ beantwortet werden. Ausnahmen stellten dabei die Studien von Bengaard et al. (2012) und Lee et al. (2015) dar, bei denen jeweils ein Plastikmodell zum Einsatz kam, das nur eine bedingte Repräsentativität darstellen kann.

Die Einschlusskriterien dieses Reviews ließen nur Studien zu, in denen fertig ausgebildete oder sich noch im Studium befindende OsteopathInnen die Untersuchung durchführten. Dies war ein entscheidender Faktor für die Repräsentativität der UntersucherInnen und auch des Reviews selber, da die Diagnose der SD ein Alleinstellungsmerkmal der Osteopathie darstellt. Die Frage Nummer zwei der Checkliste, ob die UntersucherInnen auch repräsentativ wären, konnte somit durchgehend mit einem „Ja“ beantwortet werden. Dieses Vorgehen erhöhte zwar die Repräsentativität für die Osteopathie im Allgemeinen, jedoch würden deutlich mehr Studien zur Auswertung vorliegen, wenn Arbeiten aus dem Bereich verwandter Berufsgruppen, wie Chiropraktiker oder Manualtherapeuten, ebenfalls inkludiert werden würden. Vor allem zur segmentalen Untersuchung der Wirbelsäule gibt es im Bereich der Chiropraktik relevante Studien zu finden (French, Green, & Forbes, 2000).

Eine Verblindung wurde in den meisten Studien, sofern möglich durchgeführt. Nur selten kam es zu Unklarheiten oder fehlender Verblindung, was sich in der Beantwortung der Fragen drei bis sieben zeigte. Jedoch konnten nur vier der zwanzig Studien bei der Verblindung von zusätzlichen Hinweisen, die nicht Teil der Tests waren, wie Tattoos, Narben, Hautveränderungen, Piercings usw. gut abschließen. Dies war aber aufgrund mancher Studiendesigns nur sehr schwer umzusetzen. Die Reihenfolge der UntersucherInnen, die einen Test an den Probandinnen oder Probanden durchführten, variierten nur bei zwölf der zwanzig Studien. Bei den anderen Arbeiten war dies entweder unklar oder es wurde die Reihenfolge nicht beachtet, was beides die Qualität der Studien reduziert. Ein großes Problem stellte in den meisten Arbeiten der Punkt neun dar, da es durch einen Mangel an Referenzstandards zu keiner Stabilität oder theoretischen Stabilität der Testvariablen kommen kann. Einzig in der Arbeit von Kmita and Lucas (2008) wurde eine theoretische Stabilität gegeben, indem jeder Testdurchlauf zwischen einer und zwei Minuten dauerte und eine fünfminütige Pause zwischen den Testdurchläufen vorgegeben wurde.

Die Tests wurden, soweit dies möglich ist zu überprüfen, korrekt angewandt und interpretiert. Meist wurde ein Übereinstimmungstraining vor dem eigentlichen Untersuchungsablauf durchgeführt, was zur Verbesserung der Testausführung beitragen kann (Degenhardt et al., 2005). Eine angemessene statistische Auswertung konnte jede der untersuchten Studien vorweisen. Dies geschah in den meisten Arbeiten mit dem Kappa Koeffizienten. Als einzige Abweichung wurde bei der Studie von Lee et al. (2015) die Sensitivität und die Spezifität untersucht. Da die Ergebnisse nicht mit dem Kappa Koeffizienten direkt vergleichbar sind, wurden sie in der statistischen Aufbereitung nicht miteinbezogen und müssen gesondert betrachtet werden.

In einem systematischen Review von Basile, Scionti, and Petracca (2017) über die Reliabilität von osteopathischen Tests, in dem ebenfalls die QAREL-Checkliste zum Einsatz kam, wurde auch auf das Problem der theoretischen Stabilität aufmerksam gemacht. Das Review von Basile inkludierte jedoch auch Studien, die sich nicht ausschließlich mit der Untersuchung der TART-Kriterien beschäftigen, wie die Palpation des kranialen Rhythmus und Spannungsmuster im kranialen System. Von den insgesamt 20 Studien die dort bewertet wurden, sind 13 auch in diesem Review zur Bewertung herangezogen worden. Die Ergebnisse der QAREL-Checkliste überschneiden sich zum großen Teil und weichen nur in einzelnen Punkten ab, was erneut für die Reliabilität der QAREL-Checkliste spricht. Nach einem Artikel zur Überprüfung der Evidenz der QAREL-Checkliste vom Jahr 2013 stellt sie laut Autoren ein zuverlässiges Instrument zur Beurteilung von Studien für die diagnostische Reliabilität dar (Lucas et al., 2013). Die verwendete Checkliste stellt somit ein valides Instrument zur Beurteilung von Reliabilitätsstudien dar und hat sich auch in dieser Arbeit bewährt. Ein Bias bei der Bewertung von Studien mit einer Checkliste ist trotzdem möglich, da eine Subjektivität nicht ausgeschlossen werden kann. Um dies zu reduzieren, wäre es denkbar, die Studien unabhängig von weiteren Wissenschaftlern mit derselben Checkliste bewerten zu lassen und Abweichungen von Ergebnissen im Anschluss zu diskutieren.

In den 20 untersuchten Studien, wurden bei zwölf Arbeiten nur eines der TART-Kriterien getestet, bei vier wurden zwei der Kriterien untersucht und bei weiteren vier wurden alle TART-Kriterien einer Untersuchung unterzogen. Für die Umsetzbarkeit ist es fraglich, welche Aussage eine Studie hat, die sich nur auf ein Kriterium bezieht, da in der Praxis alle Kriterien untersucht werden und der Ausdruck einer SD des Patienten sehr individuell und patientenzentriert zu betrachten ist. Warum diese ungleiche Verteilung vorhanden ist, ist unklar. Wie in der Einleitung schon beschrieben, ist laut Greenman die

Bewegungseinschränkung und die dabei palpierte Qualität und Quantität die entscheidendste Komponente (DeStefano, 2011). Auch dass eine Asymmetrie normal sein kann und kein Zeichen von Dysfunktion darstellen muss, spricht gegen eine solche Verteilung (Robert, 2019). Laut dem Lehrbuch Manuelle Medizin (Bischoff & Moll, 2007) stellt die Schmerzprovokation und damit die Gewebeempfindlichkeit die evidenteste Methode der Diagnostik dar. Die Meinung des Autors und seine empirische Erfahrung zeigen auch, dass die Qualität und Quantität der Bewegung, die Gewebebeschaffenheit und die Sensibilitätsänderungen des Gewebes in ihrer Kombination die bedeutendsten Komponenten der SD widerspiegeln. Die Asymmetrie dient dabei als zusätzliche Information und sollte nicht als Alleinstellungsmerkmal zur Identifikation einer SD genutzt werden. Für dieses Argument spricht auch die Verteilung der Übereinstimmungen im schwachen und ausreichenden Bereich beim Kriterium der Asymmetrie. Doch sogar bei der Asymmetrie konnten drei nahezu perfekte Ergebnisse bei der Testung des Malleolus lateralis auf Höhenunterschied im Seitenvergleich erzielt werden. Die moderaten bis starken Übereinstimmungen wurden entweder bei der Testung der LWS oder durch vorheriges Übereinstimmungstraining erreicht. Damit kann man zusammenfassen, dass die Asymmetrie als Zeichen der SD nur bedingt zur Diagnostik benutzt werden sollte. Die Untersuchung auf Asymmetrie im Rahmen aller TART-Kriterien kann an der LWS zur Positionsbestimmung und des Malleolus medialis zur Testung der Beinlänge von schwach bis nahezu perfekt alle Ergebnisse aufweisen. Von einer isolierten statischen Untersuchung des Beckens mit dem Rückschluss auf eine SD kann durch die vorliegenden Ergebnisse vorerst abgeraten werden.

Die Untersuchung des Bewegungsausmaßes ergab zwei starke und eine moderate Übereinstimmung bei der myofaszialen Testung der oberen Thoraxapertur (Hutchinson et al., 2017). Da diese Studie durchwegs gute Ergebnisse hatte, sollten in Zukunft mehrere Studien zu diesem Thema in Erwägung gezogen werden. Tests an Becken und der Wirbelsäule weisen starke Schwankungen von schwach bis moderat auf und sollten auch mit diesem Hintergrundwissen angewandt werden. Erfahrung und ein vorheriges Training zur Standardisierung und Interpretation können die Ergebnisse der Testung auf Bewegungseinschränkungen positiv beeinflussen. Zu bemängeln ist hierbei allerdings, dass keine einzige Studie zur Testung der Beweglichkeit an den Extremitäten zu finden war. Es wäre denkbar, dass die Ergebnisse besser ausfallen könnten, da die sehr geringe Beweglichkeit von einem einzelnen Wirbelsäulensegment oder dem SIG möglicherweise schwerer zu beurteilen ist als die deutlich größere Beweglichkeit der peripheren Gelenke.

Bei der Untersuchung auf Gewebeveränderungen wurden mit der Methode des „Mechanical Link“ zwei starke Ergebnisse erzielt. Mit vier weiteren moderaten Ergebnissen bei der Untersuchung von Wirbelsäule und Becken kann die Testung der Gewebebeschaffenheit als bisher bestes Mittel zur Identifikation der SD angewandt werden. Studien die spezifisch die TTC untersuchten oder bei denen ein Übereinstimmungstraining durchgeführt wurde, konnten dabei allgemein bessere Ergebnisse erzielen als Studien, die alle Kriterien untersuchten.

Die Gewebeempfindlichkeit wurde nur im Rahmen von Studien untersucht, die alle TART-Kriterien getestet haben. Keine einzige Studie konzentrierte sich alleine auf dieses Kriterium. Dies wäre durchaus wünschenswert gewesen, da die Ergebnisse von schwach bis nahezu perfekt stark schwanken. Ein schwaches Ergebnis konnte bei der Testung des Sakrums in der Studie von Consorti et al. (2018) erzielt werden. Zwei ausreichende Ergebnisse wurden in der Arbeit von McPartland and Goodridge (1997) erreicht, dort wurde jedoch die TART-Untersuchung als Kontrolltestung zur SCS-Methode angewandt und nicht genauer beschrieben. Das starke bis nahezu perfekte Ergebnis der Reliabilität konnte an der LWS unter Einbeziehung eines Übereinstimmigkeitstrainings erzielt werden. Mit diesem Hintergrund kann die Testung auf Gewebeempfindlichkeit neben der Gewebeveränderung als zweitbestes Mittel zur Identifikation der SD angewandt werden.

Zusammenfassend sind die Ergebnisse sehr heterogen. Die schlechtesten Ergebnisse waren im Bereich der statischen Untersuchung des Beckens zu finden, zu diesem Thema werden auch die meisten Studien gefunden. Die besten Ergebnisse wurden bei der Untersuchung von Wirbelsäule und Extremitäten nach den Kriterien Gewebebeschaffenheit, Gewebeempfindlichkeit und Beweglichkeit erzielt. Eine weitere gute Reliabilität konnte mit der Untersuchung der Gewebebeschaffenheit erzielt werden, gefolgt von der Gewebeempfindlichkeit an zweiter Stelle. Myofasziale Spannungen zu testen, hatte bei der Untersuchung des Bewegungsausmaßes die optimalsten Ergebnisse und ist somit nach aktuellem Stand zu empfehlen. Die Untersuchungsergebnisse der Asymmetrie sind sehr durchwachsen und werden nach Analyse der vorliegenden Studien nur bedingt als Diagnosekriterium empfohlen. Eine Reliabilität ist durchaus gegeben, jedoch sollte sich der Leser der wechselhaften Ergebnisse bewusst sein. Die Arbeit konnte gut darstellen, welche Reliabilität bei der Diagnostik der SD vorhanden ist und wo noch Schwächen zu finden sind. Weitere Arbeiten im Bereich der Reliabilitätstestung sind daher wünschenswert, um in zukünftigen Reviews das Thema weiter zu untersuchen.

Dass ein vorheriges Übereinstimmkeitstraining zu besseren Endergebnissen der Studien führte, war eine weitere wichtige Erkenntnis dieser Arbeit. Zur Diskussion steht offen, ob ein vorheriges Training die Gesamtergebnisse nicht fälschlicherweise zum Positiven verändern könnte. Zu wissen, dass ein vorheriges Training zu besseren Ergebnissen führt, kann aber auch in Aus- und Fortbildungen genutzt werden. Somit wäre es wünschenswert, dass ein intensiveres Training von Tests und deren Interpretation schon an Ausbildungsstätten gelehrt werden sollte.

Grundlegend wären größere Fallzahlen in den Studien mit Einbeziehung aller TART-Parameter und vorherigem Übereinstimmigkeitstraining interessante Studien für die Zukunft. Die somatische Dysfunktion in ihrem ganzen Ausmaß, unter Berücksichtigung aller TART-Kriterien und unter standardisierten Vorgaben, als ganze Testreihe zu untersuchen, wäre ein wichtiger Schritt in Richtung weiterer Evidenz.

Literaturverzeichnis

- Basile, F., Scionti, R., & Petracca, M. (2017). Diagnostic reliability of osteopathic tests: A systematic review. *International Journal of Osteopathic Medicine*, 25, 21-29. doi:10.1016/j.ijosm.2017.03.004
- Bengaard, K., Bogue, R., & Crow, W. (2012). Reliability of diagnosis of somatic dysfunction among osteopathic physicians and medical students. *Osteopathic Family Physician*, 4. doi:10.1016/j.osfp.2011.08.003
- Bischoff, H. P., & Moll, H. (2007). *Kurz gefasstes Lehrbuch der manuellen Medizin*: Spitta.
- Bogduk, N. (2005). *Clinical Anatomy of the Lumbar Spine and Sacrum*: Elsevier/Churchill Livingstone.
- Bogduk, N., April, C., & Derby, R. (2013). Lumbar discogenic pain: state-of-the-art review. *Pain Med*, 14(6), 813-836. doi:10.1111/pme.12082
- Böhni, U. W., Lauper, M., & Locher, H. A. (2015). *Manuelle Medizin 1: Fehlfunktion und Schmerz am Bewegungsorgan verstehen und behandeln*: Thieme.
- Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., . . . group, f. t. S. (2011). Vollständiges und präzises Berichten von Studien zur diagnostischen Genauigkeit: Die STARD-Initiative. doi:10.1055/s-0031-1272981
- Breuer, J.-P., Seeling, M., Barz, M., Baldini, T., Scholtz, K., & Spies, C. (2012). A standardised German translation of the STAndards for Reporting of Diagnostic accuracy studies (STARD statement)*: Methodological aspects. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, 106(7), 500-508. doi:10.1016/j.zefq.2012.05.021
- Brinjikji, W., Luetmer, P. H., Comstock, B., Bresnahan, B. W., Chen, L. E., Deyo, R. A., . . . Jarvik, J. G. (2015). Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. *AJNR Am J Neuroradiol*, 36(4), 811-816. doi:10.3174/ajnr.A4173
- Carreiro, J. E. (2011). *Osteopathie bei Kindern und Jugendlichen: Grundlagen, Krankheitsbilder und Behandlungstechniken*: Elsevier, Urban & Fischer.

Chaitow, L. (2001). *Palpationstechniken und Diagnostik: Lehr- und Arbeitsbuch für Osteopathen*: Urban und Fischer.

Colloca, C. J., Keller, T. S., Moore, R. J., Gunzburg, R., & Harrison, D. E. (2007). Intervertebral disc degeneration reduces vertebral motion responses. *Spine*, 32(19), E544-550. doi:10.1097/BRS.0b013e318145ac39

Comeaux, Z. (2005). Somatic Dysfunction – a Reflection on the Scope of Osteopathic Practice. *The AAO Journal*.

Comeaux, Z., Eland, D., Chila, A., Pheley, A., & Tate, M. (2001). Measurement challenges in physical diagnosis: refining inter-rater palpation, perception and communication. *Journal of Bodywork and Movement Therapies*, 5(4), 245-253. doi:10.1054/jbmt.2001.0236

Consorti, G., Basile, F., Pugliese, L., & Petracca, M. (2018). Interrater Reliability of Osteopathic Sacral Palpatory Diagnostic Tests Among Osteopathy Students. *The Journal of the American Osteopathic Association*, 118(10), 637-644. doi:10.7556/jaoa.2018.132

Cramer, G. D., Henderson, C. N., Little, J. W., Daley, C., & Grieve, T. J. (2010). Zygapophyseal joint adhesions after induced hypomobility. *J Manipulative Physiol Ther*, 33(7), 508-518. doi:10.1016/j.jmpt.2010.08.002

Degenhardt, B. F., Johnson, J. C., Snider, K. T., & Snider, E. J. (2010). Maintenance and Improvement of Interobserver Reliability of Osteopathic Palpatory Tests Over a 4-Month Period. *The Journal of the American Osteopathic Association*, 110(10), 579-586. doi:10.7556/jaoa.2010.110.10.579

Degenhardt, B. F., Snider, K. T., Snider, E. J., & Johnson, J. C. (2005). Interobserver Reliability of Osteopathic Palpatory Diagnostic Tests of the Lumbar Spine: Improvements From Consensus Training. *The Journal of the American Osteopathic Association*, 105(10), 465-473. doi:10.7556/jaoa.2005.105.10.465

Denslow, J. S. (1953). The somatic component. *The Journal of the American Osteopathic Association*, 52(5), 258-261.

DeStefano, L. A. (2011). *Greenman's Principles of Manual Medicine*: Lippincott Williams & Wilkins/Wolters Kluwer.

Ellingson, A. M., Shaw, M. N., Giambini, H., & An, K. N. (2016). Comparative role of disc degeneration and ligament failure on functional mechanics of the lumbar spine. *Comput Methods Biomed Engin*, 19(9), 1009-1018. doi:10.1080/10255842.2015.1088524

Frank, C. B., Hart, D. A., & Shrive, N. G. (1999). Molecular biology and biomechanics of normal and healing ligaments--a review. *Osteoarthritis Cartilage*, 7(1), 130-140. doi:10.1053/joca.1998.0168

Freeman, M. A., & Wyke, B. (1967). The innervation of the knee joint. An anatomical and histological study in the cat. *Journal of anatomy*, 101(Pt 3), 505-532. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/6051731>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1270929/>

French, S. D., Green, S., & Forbes, A. (2000). Reliability of chiropractic methods commonly used to detect manipulable lesions in patients with chronic low-back pain. *Journal of manipulative and physiological therapeutics*, 23(4), 231-238. doi:10.1067/mmt.2000.106101

Fryer, & Adams, J. H. (2011). Magnetic resonance imaging of subjects with acute unilateral neck pain and restricted motion: a prospective case series. *The spine journal : official journal of the North American Spine Society*, 11(3), 171-176. doi:10.1016/j.spinee.2010.12.002

Fryer, G. (1999). Somatic dysfunction: Updating the concept.

Fryer, G. (2003). Intervertebral dysfunction: a discussion of the manipulable spinal lesion. *Journal of Osteopathic Medicine*, 6(2), 64-73. doi:10.1016/S1443-8461(03)80016-3

Fryer, G. (2006). Factors affecting the intra-examiner and inter-examiner reliability of palpation for supine medial malleoli asymmetry. *International Journal of Osteopathic Medicine*, 9(2), 58-65. doi:10.1016/j.ijosm.2005.11.004

Fryer, G. (2016). Somatic dysfunction: An osteopathic conundrum. *International Journal of Osteopathic Medicine*, 22, 52-63. doi:10.1016/j.ijosm.2016.02.002

Fryer, G., McPherson, H. C., & O'Keefe, P. (2005). The effect of training on the inter-examiner and intra-examiner reliability of the seated flexion test and assessment of pelvic anatomical landmarks with palpation. *International Journal of Osteopathic Medicine*, 8(4), 131-138. doi:10.1016/j.ijosm.2005.08.004

Gibbons, P., Dumper, C., & Gosling, C. (2002). Inter-examiner and intra-examiner agreement for assessing simulated leg length inequality using palpation and observation during a standing assessment. *Journal of Osteopathic Medicine*, 5(2), 53-58. doi:10.1016/S1443-8461(02)80002-8

Hafen, G. I. (2019). Reliabilitätsstudie des Befunds nach dem Mechanical Link: Methodologische Studie über die Befunderhebung der Knochen und Gelenke der Extremitäten nach der Methode des Mechanical Link. *Osteopathische Medizin*, 20(1), 21-26. doi:[https://doi.org/10.1016/S1615-9071\(19\)30013-9](https://doi.org/10.1016/S1615-9071(19)30013-9)

Herbert, M. K., & Holzer, P. (2002). [Neurogenic inflammation. I. Basic mechanisms, physiology and pharmacology]. *Anesthesiol Intensivmed Notfallmed Schmerzther*, 37(6), 314-325. doi:10.1055/s-2002-32233

Hutchinson, D., Hines, S., Vijayaraghavan, N., Sammond, A., Metzler-Wilson, K., & Kuchera, M. L. (2017). Interexaminer reliability study of a standardized myofascial diagnostic technique of the superior thoracic inlet. *Journal of Bodywork and Movement Therapies*, 21(3), 658-663. doi:10.1016/j.jbmt.2017.05.004

Indahl, A., Kaigle, A., Reikerås, O., & Holm, S. (1995). Electromyographic response of the porcine multifidus musculature after nerve stimulation. *Spine*, 20(24), 2652-2658. doi:10.1097/00007632-199512150-00006

Jung, H. J., Fisher, M. B., & Woo, S. L. (2009). Role of biomechanics in the understanding of normal, injured, and healing ligaments and tendons. *Sports Med Arthrosc Rehabil Ther Technol*, 1(1), 9. doi:10.1186/1758-2555-1-9

Kia, T. (2013). *Integrative Osteopathie bei Rückenschmerz: Systematische Diagnostik und ganzheitliche Therapie - mit Zugang zum Elsevier-Portal*: Urban & Fischer in Elsevier.

Kmita, A., & Lucas, N. P. (2008). Reliability of physical examination to assess asymmetry of anatomical landmarks indicative of pelvic somatic dysfunction in subjects with and

without low back pain. *International Journal of Osteopathic Medicine*, 11(1), 16-25. doi:10.1016/j.ijosm.2008.01.003

Korr, I. M. (1947). The neural basis of the osteopathic lesion. *The Journal of the American Osteopathic Association*, 47(4), 191-198.

Kushchayev, S. V., Glushko, T., Jarraya, M., Schuleri, K. H., Preul, M. C., Brooks, M. L., & Teytelboym, O. M. (2018). ABCs of the degenerative spine. *Insights Imaging*, 9(2), 253-274. doi:10.1007/s13244-017-0584-z

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

Latremoliere, A., & Woolf, C. J. (2009). Central sensitization: a generator of pain hypersensitivity by central neural plasticity. *J Pain*, 10(9), 895-926. doi:10.1016/j.jpain.2009.06.012

Lederman, E. (2000). Facilitated segments: a critical review. *British Osteopathic Journal*, 7-10.

Lee, A. S., Pyle, C. W., & Redding, D. (2015). Accuracy of Anterior Superior Iliac Spine Symmetry Assessment by Routine Structural Examination. *The Journal of the American Osteopathic Association*, 115(8), 482-489. doi:10.7556/jaoa.2015.102

Lewit, K. (2012). *Manuelle Medizin: bei Funktionsstörungen des Bewegungsapparates*: Elsevier Health Sciences Germany.

Liem, T. (2016). A.T. Still's Osteopathic Lesion Theory and Evidence-Based Models Supporting the Emerged Concept of Somatic Dysfunction. *The Journal of the American Osteopathic Association*, 116(10), 654-661. doi:10.7556/jaoa.2016.129

Liem, T. (2018). Von A.T. Stills Theorie der osteopathischen Läsion zur somatischen Dysfunktion: Geschichtliche Entwicklung und evidenzbasierte Modelle, die das aktuelle Konzept der somatischen Dysfunktion unterstützen. *Osteopathische Medizin*, 19(3), 21-26. doi:[https://doi.org/10.1016/S1615-9071\(18\)30083-2](https://doi.org/10.1016/S1615-9071(18)30083-2)

Liem, T., & Dobler, T. K. (2017). *Leitfaden Osteopathie: Parietale Techniken*: Elsevier Health Sciences Germany.

Lucas, Macaskill, P., Irwig, L., Moran, R., Rickards, L., & Turner, R. (2013). The reliability of a quality appraisal tool for studies of diagnostic reliability (QAREL). *BMC medical research methodology*, 13, 111. doi:10.1186/1471-2288-13-111

Lucas, N. P., Macaskill, P., Irwig, L., & Bogduk, N. (2010). The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). *Journal of clinical epidemiology*, 63(8), 854-861. doi:10.1016/j.jclinepi.2009.10.002

McCole, G. M. (1935). *An Analysis of the Osteopathic Lesion: A Study in Pathology, Physiology and Anatomy*: GM McCole.

McPartland, J. M., & Goodridge, J. P. (1997). Counterstrain and traditional osteopathic examination of the cervical spine compared. *Journal of Bodywork and Movement Therapies*, 1(3), 173-178. doi:10.1016/S1360-8592(97)80037-X

Meert, G. F. (2009). *Das Becken aus osteopathischer Sicht : Funktionelle Zusammenhänge nach dem Tensegrity-Modell*: Urban & Fischer.

Mooney, V. (1987). Presidential address. International Society for the Study of the Lumbar Spine. Dallas, 1986. Where is the pain coming from? *Spine*, 12(8), 754-759. doi:10.1097/00007632-198710000-00008

Nazarian, L. N., Zegel, H. G., Gilbert, K. R., Edell, S. L., Bakst, B. L., & Goldberg, B. B. (1998). Paraspinal ultrasonography: lack of accuracy in evaluating patients with cervical or lumbar back pain. *J Ultrasound Med*, 17(2), 117-122. doi:10.7863/jum.1998.17.2.117

O'Haire, C., & Gibbons, P. (2000). Inter-examiner and intra-examiner agreement for assessing sacroiliac anatomical landmarks using palpation and observation: pilot study. *Manual therapy*, 5(1), 13-20. doi:10.1054/math.1999.0203

Pape, H. C., Kurtz, A., & Silbernagl, S. (2018). *Physiologie*: Thieme.

Paulet, T., & Fryer, G. (2009). Inter-examiner reliability of palpation for tissue texture abnormality in the thoracic paraspinal region. *International Journal of Osteopathic Medicine*, 12(3), 92-96. doi:10.1016/j.ijosm.2008.07.001

Peng, Y. B., Kenshalo, D. R., & Gracely, R. H. (2003). Periaqueductal gray-evoked dorsal root reflex is frequency dependent. *Brain Res*, 976(2), 217-226. doi:10.1016/s0006-8993(03)02718-5

Principles, A. A. o. C. o. O. M. E. C. o. O., & Medicine, M. S. U. C. o. O. (1982). *Glossary of Osteopathic Terminology*: Michigan State University College of Osteopathic Medicine.

Rajendran, D., & Gallagher, D. (2011). The assessment of pelvic landmarks using palpation: A reliability study of undergraduate students. *International Journal of Osteopathic Medicine*, 14(2), 57-60. doi:10.1016/j.ijosm.2010.10.005

Robert, V. (2019). Ist Asymmetrie Dysfunktion? *Osteopathische Medizin*, 20(1), 4-8. doi:[https://doi.org/10.1016/S1615-9071\(19\)30009-7](https://doi.org/10.1016/S1615-9071(19)30009-7)

Scherfer, E., & Hempel, C. (2006). *Forschung verstehen: ein Grundkurs in evidenzbasierter Praxis*: Pflaum.

Schulte, E., Schumacher, U., & Schünke, M. (2018). *PROMETHEUS Kopf, Hals und Neuroanatomie: LernAtlas Anatomie*: Thieme.

Spring, F., Gibbons, P., & Tehan, P. (2001). Intra-examiner and inter-examiner reliability of a positional diagnostic screen for the lumbar spine. *Journal of Osteopathic Medicine*, 4(2), 47-55. doi:10.1016/S1443-8461(01)80002-2

Steilen, D., Hauser, R., Woldin, B., & Sawyer, S. (2014). Chronic neck pain: making the connection between capsular ligament laxity and cervical instability. *Open Orthop J*, 8, 326-345. doi:10.2174/1874325001408010326

Still, A. T. (1908). *Autobiography of Andrew T. Still: With a History of the Discovery and Development of the Science of Osteopathy, Together with an Account of the Founding of the American School of Osteopathy; and Lectures Delivered Before that Institution*: The author.

Sutton, C., Nono, L., Johnston, R. G., & Thomson, O. P. (2013). The effects of experience on the inter-reliability of osteopaths to detect changes in posterior superior iliac spine levels using a hidden heel wedge. *Journal of Bodywork and Movement Therapies*, 17(2), 143-150. doi:10.1016/j.jbmt.2012.07.005

Tong, H. C., Heyman, O. G., Lado, D. A., & Isser, M. M. (2006). Interexaminer reliability of three methods of combining test results to determine side of sacral restriction, sacral base position, and innominate bone position. *The Journal of the American Osteopathic Association*, 106(8), 464-468. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/16943516>

Tozzi, P. (2015). A unifying neuro-fasciogenic model of somatic dysfunction - underlying mechanisms and treatment - Part I. *J Bodyw Mov Ther*, 19(2), 310-326. doi:10.1016/j.jbmt.2015.01.001

Trepel, M., & Dalkowski, K. (2011). *Neuroanatomie: Struktur und Funktion*: Elsevier Health Sciences Germany.

Van Buskirk, R. L. (1990). Nociceptive reflexes and the somatic dysfunction: a model. *The Journal of the American Osteopathic Association*, 90(9), 792-794, 797-809. Retrieved from <http://europepmc.org/abstract/MED/2211195>

<http://www.jaoa.org/cgi/content/abstract/90/9/792>

Vincent-Smith, B., & Gibbons, P. (1999). Inter-examiner and intra-examiner reliability of the standing flexion test. *Manual therapy*, 4(2), 87-93. doi:10.1054/math.1999.0173

Webb, A. L., Collins, P., Rassoulian, H., & Mitchell, B. S. (2011). Synovial folds - a pain in the neck? *Manual therapy*, 16(2), 118-124. doi:10.1016/j.math.2010.11.004

Whiting, P., Rutjes, A. W., Reitsma, J. B., Bossuyt, P. M., & Kleijnen, J. (2003). The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC medical research methodology*, 3, 25. doi:10.1186/1471-2288-3-25

Willis, W. D., Jr. (1999). Dorsal root potentials and dorsal root reflexes: a double-edged sword. *Exp Brain Res*, 124(4), 395-421. doi:10.1007/s002210050637

Wolff, H. D. (2013). *Neurophysiologische Aspekte der manuellen Medizin*: Springer Berlin Heidelberg.

Tabellenverzeichnis

Tabelle 1: Nervenfasereinteilung nach Lloyd und Hunt (Pape et al., 2018).....	9
Tabelle 2: Nervenfasereinteilung nach Erlanger und Gasser (Pape et al., 2018).....	10
Tabelle 3: Ergebnisse der Literatursuche	23
Tabelle 4: Literatursuche.....	23
Tabelle 5: Inkludierte Studien	25
Tabelle 6: QAREL-Checkliste (Lucas et al., 2010)	27
Tabelle 7: Interpretation des Kappa Koeffizienten nach Landis and Koch (1977).....	34
Tabelle 8: Reliability of diagnosis of somatic dysfunction among osteopathic physicians and medical students (QAREL-Checkliste).....	36
Tabelle 9: Measurement challenges in physical diagnosis: refining inter-rater palpation, perception and communication (QAREL-Checkliste)	38
Tabelle 10: Interrater Reliability of Osteopathic Sacral Palpatory Diagnostic Tests Among Osteopathy Students (QAREL-Checkliste).....	40
Tabelle 11: Interobserver Reliability of Osteopathic Palpatory Diagnostic Tests of the Lumbar Spine: Improvements from Consensus Training (QAREL-Checkliste).....	41
Tabelle 12: Maintenance and Improvement of Interobserver Reliability of Osteopathic Palpatory Tests Over a 4-Month Period (QAREL-Checkliste).....	43
Tabelle 13: Factors affecting the intra-examiner and inter-examiner reliability of palpation for supine medial malleoli asymmetry (QAREL-Checkliste)	45
Tabelle 14: Intrarater Reliabilität	46
Tabelle 15: Interrater Reliabilität	47
Tabelle 16: The effect of training on the inter-examiner and intra-examiner reliability of the seated flexion test and assessment of pelvic anatomical landmarks with palpation (QAREL-Checkliste).....	48

Tabelle 17: Inter-examiner and intra-examiner agreement for assessing simulated leg length inequality using palpation and observation during a standing assessment (QAREL-Checkliste)	49
Tabelle 18: Reliabilitätsstudie des Befunds nach dem Mechanical Link: Methodologische Studie über die Befunderhebung der Knochen und Gelenke der Extremitäten nach der Methode des Mechanical Link (QAREL-Checkliste)	51
Tabelle 19: Tabelle 19: Interexaminer reliability study of a standardized myofascial diagnostic technique of the superior thoracic inlet (QAREL-Checkliste).....	53
Tabelle 20: Ergebnisse der intrarater- und interrater Reliabilitätstestung.....	54
Tabelle 21: Reliability of physical examination to assess asymmetry of anatomical landmarks indicative of pelvic somatic dysfunction in subjects with and without low back pain (QAREL-Checkliste).....	55
Tabelle 22: Accuracy of Anterior Superior Iliac Spine Symmetry Assessment by Routine Structural Examination (QAREL-Checkliste)	56
Tabelle 23: Counterstrain and traditional osteopathic examination of the cervical spine compared (QAREL-Checkliste)	58
Tabelle 24: Inter-examiner and intra-examiner agreement for assessing sacroiliac anatomical landmarks using palpation and observation: pilot study (QAREL-Checkliste)	60
Tabelle 25: Inter-examiner reliability of palpation for tissue texture abnormality in the thoracic paraspinal region (QAREL-Checkliste)	61
Tabelle 26: The assessment of pelvic landmarks using palpation: A reliability study of undergraduate students (QAREL-Checkliste).....	63
Tabelle 27: Intra-examiner and inter-examiner reliability of a positional diagnostic screen for the lumbar spine (QAREL-Checkliste)	64
Tabelle 28: The effects of experience on the inter-reliability of osteopaths to detect changes in posterior superior iliac spine levels using a hidden heel wedge (QAREL-Checkliste)	66

Tabelle 29: Interexaminer reliability of three methods of combining test results to determine side of sacral restriction, sacral base position, and innominate bone position (QAREL-Checkliste).....	67
Tabelle 30: Inter-examiner and intra-examiner reliability of the standing flexion test (QAREL-Checkliste).....	69
Tabelle 31: Zusammenfassung der Ergebnisse	70
Tabelle 32: Zusammenfassung der Ergebnisse aller QAREL-Checklisten	72

Abbildungsverzeichnis

Abbildung 1: Faktoren der somatischen Dysfunktion (eigene Darstellung).....	6
Abbildung 2: Graue Substanz im Rückenmark (eigene Darstellung)	13
Abbildung 3: WDR-Neuron (eigene Darstellung)	14
Abbildung 4: Somatoafferzen, Visceraoafferzen, Somatoefferenzen und Visceroefferenzen (eigene Darstellung).....	16
Abbildung 5: Mind-Map zum Forschungsdesign (eigene Darstellung).....	20
Abbildung 6: Flow-Chart (eigene Darstellung)	24
Abbildung 7: Kreisdiagramm aller untersuchten TART-Kriterien in Prozentangaben (eigene Darstellung).....	73
Abbildung 8: Diagramm zur Darstellung der Häufigkeitenverteilung (eigene Darstellung)...	74
Abbildung 9: Kreisdiagramm zur Darstellung der untersuchten Körperregionen (eigene Darstellung).....	74

Abkürzungsverzeichnis

AIL.....Angulus inferior lateralis

BWS.....Brustwirbelsäule

CGRP.....Calcitonin-gene-related-peptide

HWS.....Halswirbelsäule

KCOM.....Kirksville College of Osteopathic Medicine

LWS.....Lendenwirbelsäule

PNS.....Peripheres Nervensystem

QAREL....Quality Appraisal of Reliability Studies-Checkliste

ROM.....Range of motion

SCS.....Strain Counterstrain

SD.....Somatische Dysfunktion

SIAS.....Spina iliaca anterior superior

SIPS.....Spina iliaca posterior superior

STARD.....Standards for Reporting of Diagnostic Accuracy-Checkliste

TART.....Tissue texture changes, asymmetry, restriction of motion, tenderness

WDR.....Wide dynamic range neuron

ZNS.....Zentrales Nervensystem