

INTERRATER-RELIABILITÄT DES INDUKTIONSTESTS OS MANDIBULARE

Masterarbeit zur Erlangung des akademischen Grades

Master of Science

im Studium Osteopathie MSc

eingereicht von

Elisabeth Nauschnegg

Department für Gesundheitswissenschaften, Medizin und Forschung

an der Universität für Weiterbildung Krems

Betreuerin: Lisa-Teresa Woller-Mensshengen, BSc, MSc, MA

Betreuer: Raimund Engel, MSc D.O.

Frauental an der Laßnitz, 09. November 2025

EIDESSTÄTTLICHE ERKLÄRUNG

Ich, Elisabeth Nauschnegg, erkläre hiermit an Eides statt:

Ich habe meine Masterarbeit selbständig verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubten Hilfen bedient.

Folgende KI-gestützte Tools wurden unterstützend zur wissenschaftlichen Eigenleistung verwendet

- Chat-GPT für Textüberarbeitung

Ich bin für sämtliche Inhalte meiner Arbeit, einschließlich der durch KI-gestützte Tools generierten oder unterstützten Passagen, allein verantwortlich und die Einhaltung wissenschaftlicher Standards liegt in meiner alleinigen Verantwortung.

Ich habe meine Masterarbeit oder wesentliche Teile daraus bisher weder im In- noch im Ausland in irgendeiner Form als Prüfungsarbeit vorgelegt.

Ich habe, falls die Masterarbeit mein Unternehmen und/oder eine_n externe_n Kooperationspartner_in betrifft, diese über Titel, Form und Inhalt der Masterarbeit unterrichtet und ihr_sein Einverständnis eingeholt.

Datum: 09.11.2025

Unterschrift: 

DANKSAGUNGEN

An dieser Stelle möchte ich mich von Herzen bei all jenen bedanken, die mich auf meinem Weg begleitet und mich unterstützt haben.

Mein besonderer Dank gilt meiner Familie und meinen Freunden, die mir in den letzten Jahren stets zur Seite gestanden sind.

Ein riesiges Dankeschön geht vor allem an meinen Lebenspartner und besten Freund Markus. Du warst nicht nur während meiner gesamten Ausbildung meine größte Stütze, sondern hast mir auch in der intensiven Zeit des Schreibens dieser Masterarbeit den Rücken freigehalten. Danke, dass du dich so liebevoll um unsere beiden kleinen Töchter gekümmert hast und immer für mich da warst. Auch meinen Kindern Miriam und Olivia danke ich von Herzen für ihre Geduld und ihr Verständnis während dieser intensiven Zeit.

Ein großes Dankeschön geht auch an meine beiden Testerinnen Veronika und Sabrina. Ohne euch wäre diese Arbeit nicht möglich gewesen. Danke für eure Zeit, eure Unterstützung und euer Vertrauen. Veronika, dir möchte ich auch ganz besonders für die vielen Gespräche, deine ehrlichen Worte und deine Hilfe in so vielen Momenten danken.

Ein herzlicher Dank gebührt auch allen Proband_innen der Studie dafür, dass sie sich die Zeit genommen haben, mich zu unterstützen.

Zu guter Letzt möchte ich mich noch bei allen Menschen bedanken, die mich durch meine Ausbildungszeit hindurch begleitet haben, sei es im Unterricht, im Austausch oder einfach durch ihre Hilfe im richtigen Moment. Jede einzelne Unterstützung von euch hat dazu beigetragen, dass ich heute an diesem Punkt stehen darf.

ABSTRACT DEUTSCH

Wissenschaftlicher Hintergrund:

Das Ziel der vorliegenden Arbeit war es, die Interrater-Reliabilität des Induktionstests am Os mandibulare zu untersuchen. Dieser Test dient der qualitativen Beurteilung der Beweglichkeit der Mandibula in Außen- und Innenrotation. Bislang liegen noch keine Studien zur Reliabilität und Validität dieses Testverfahrens vor.

Methodik:

Bei der vorliegenden Studie handelt es sich um eine Interrater-Reliabilitätsstudie. Hierbei wurde der Induktionstest Os mandibulare an 42 Proband_innen (26 weiblich, 16 männlich) durchgeführt. Die Bewertung erfolgte anhand eines standardisierten Bogens mit drei Kategorien („unauffällig“, „auffällig rechts“, „auffällig links“). Um die Übereinstimmung zwischen den beiden Testerinnen zu beurteilen, wurde der Cohens-Kappa-Koeffizient berechnet.

Ergebnisse:

Je nach Auswertungskategorie lagen die Cohens-Kappa-Werte zwischen $\kappa = 0,25$ und $\kappa = 0,48$. Für eine substantielle Übereinstimmung wurde jedoch ein Wert von $\kappa \geq 0,61$ nach Landis und Koch (1977, S. 165) angestrebt. Die höchste Übereinstimmung wurde in der symptomatischen Subgruppe festgestellt ($\kappa = 0,48$).

Conclusio:

Die Übereinstimmung zwischen den Testerinnen ist als mäßig bis mittelmäßig zu bewerten. Demnach zeigen die Ergebnisse, dass dieser Test derzeit nur eingeschränkt in der klinischen Praxis eingesetzt werden kann. Aus diesem Grund sind weitere Studien zur Interrater-Reliabilität mit größeren Stichproben sowie zur Intrarater-Reliabilität und zur Validität empfehlenswert.

Stichworte:

Interrater-Reliabilität, Induktionstest, passiver Bewegungstest, Os mandibulare, Osteopathie

ABSTRACT ENGLISCH

Background:

The aim of this study was to investigate the interrater reliability of the induction test of the mandible. The test was used for the qualitative assessment of mandibular mobility in external and internal rotation. So far, no studies have been published on the reliability and validity of this test.

Research design and method:

This study was conducted as an interrater reliability study. The induction test was performed on 42 participants (26 female and 16 male). The assessment was completed using a standardized form with three categories ("normal", "right abnormal" and "left abnormal"). Cohen's kappa coefficient was calculated to measure the agreement between the raters.

Results:

The kappa values ranged between $\kappa = 0.25$ and $\kappa = 0.48$ depending on the evaluation category. According to Landis and Koch (1977, S. 165), a value of $\kappa \geq 0.61$ was aimed for substantial agreement. The highest agreement was found in the symptomatic subgroup ($\kappa = 0.48$).

Conclusion

The agreement between the testers was rated as fair to moderate. The results indicate that this test can currently only be used to a limited extent in clinical practice. Therefore, further studies with larger samples, as well as investigations of intrarater reliability and validity, are recommended.

Keywords:

Interrater reliability, induction test, passive motion test, mandible, osteopathy

INHALTSVERZEICHNIS

DANKSAGUNGEN	I
ABSTRACT DEUTSCH	II
ABSTRACT ENGLISCH	III
INHALTSVERZEICHNIS	1
1 Einleitung.....	6
2 Stand der Wissenschaft	9
2.1 Biomechanik und funktionelle Bedeutung des Os mandibulare	9
2.1.1 Biomechanik der Mandibula	9
2.1.2 Funktionelle Relevanz.....	9
2.2 Funktionelle Zusammenhänge.....	10
2.2.1 Mandibula und HWS.....	10
2.2.2 Mandibula und Os hyoideum	11
2.2.3 Mandibula und Temporomandibulargelenk (TMJ).....	12
2.2.4 Mandibula und Körperstatik	12
2.2.5 Mandibula und Nervensystem	12
2.2.6 Mandibula und fasziale Verbindungen	13
2.3 Dysfunktion des Os mandibulare	14
2.3.1 Symptome.....	14
2.3.2 Ursachen.....	14
2.4 Differenzialdiagnostik Craniomandibuläre Dysfunktion (CMD).....	15
2.5 Epidemiologie.....	16
2.6 Osteopathische Therapie und Behandlung	17
2.6.1 Behandlung der Dysfunktion in Außenrotation.....	17
2.6.2 Behandlung der Dysfunktion in Innenrotation.....	17
2.6.3 Behandlung der Kompressionsdysfunktion des kranio- mandibulären Komplexes	18
2.6.4 Behandlung der artikulären Dysfunktionen.....	18
2.6.4.1 Behandlung der anterioren Dysfunktion	18
2.6.4.2 Behandlung der posterioren Dysfunktion	19
2.6.5 Entspannung der Kaumuskelatur	19

2.7	Diagnostische Verfahren zur Untersuchung der Mandibula	19
2.7.1	Bildgebende Verfahren	19
2.7.1.1	Konventionelle Verfahren (OPT)	20
2.7.1.2	Magnetresonanztomographie (MRT)	20
2.7.1.3	Neuere Verfahren und technische Entwicklungen	20
2.7.2	Manuelle Diagnostik der Mandibula.....	22
2.7.2.1	Aktive, passive und spezifische Tests	23
2.7.2.2	Induktionstest Os mandibulare nach Corts	24
2.8	Gütekriterien und methodologische Grundlagen von Interrater-Reliabilitätsstudien	25
2.8.1	Gütekriterien und Messinstrumente.....	25
2.8.2	Design und Durchführung von Interrater-Reliabilitätsstudien.....	26
2.9	Forschungsstand zur Reliabilität osteopathischer Tests	27
3	Forschungsdesign und -methode	29
3.1	Stichprobenbeschreibung	29
3.2	Proband_innen.....	29
3.3	Ein- und Ausschlusskriterien.....	29
3.3.1	Einschlusskriterien	29
3.3.2	Ausschlusskriterien	30
3.4	Stichprobengröße.....	30
3.4.1	Demografische Daten der Stichprobe	30
3.4.2	Daten zur vorliegenden Symptomatik.....	31
3.5	Tester_innen	32
3.6	Messinstrumente und Testbeschreibung.....	33
3.6.1	Beschreibung des Testverfahrens.....	33
3.6.2	Materialien und Gerätschaften.....	34
3.7	Studienablauf	35
3.7.1	Trainingsphase	35
3.7.2	Studienphase	36
3.7.3	Verblindung.....	37
3.7.4	Randomisierung.....	37

3.8	Datenverarbeitung und -auswertung.....	37
3.8.1	Darstellung der Daten.....	37
3.8.2	Berechnung und Interpretation des Cohens-Kappa-Koeffizienten.....	38
3.8.3	Übersicht über die unterschiedlichen berechneten Cohens-Kappa-Koeffizienten.....	39
3.8.4	Testung der Hypothesen.....	40
4	Ergebnisse.....	41
4.1	Cohens-Kappa-Koeffizient mit drei Ergebniskategorien der gesamten Stichprobe.....	41
4.2	Cohens-Kappa-Koeffizient mit drei Ergebniskategorien der weiblichen Stichprobe.....	42
4.3	Cohens-Kappa-Koeffizient mit drei Ergebniskategorien der männlichen Stichprobe.....	42
4.4	Cohens-Kappa-Koeffizient mit zwei Ergebniskategorien der gesamten Stichprobe.....	43
4.5	Cohens-Kappa-Koeffizient mit zwei Ergebniskategorien der weiblichen Stichprobe.....	44
4.6	Cohens-Kappa-Koeffizient mit zwei Ergebniskategorien der männlichen Stichprobe.....	45
4.7	Cohens-Kappa-Koeffizient mit zwei Ergebniskategorien der asymptomatischen Stichprobe.....	45
4.8	Cohens-Kappa-Koeffizient mit zwei Ergebniskategorien der symptomatischen Stichprobe.....	46
4.9	Übersicht aller errechneten Cohens-Kappa-Koeffizienten.....	46
4.10	Häufigkeitsverteilung der Stichprobe.....	47
4.11	Überprüfung der Hypothese.....	49
5	Diskussion.....	50
5.1	Interpretation der Ergebnisse.....	50
5.1.1	Interpretation des Cohens-Kappa-Koeffizienten mit drei Ergebniskategorien der gesamten Stichprobe.....	50
5.1.2	Interpretation des Cohens-Kappa-Koeffizienten mit drei Ergebniskategorien der weiblichen Stichprobe.....	51

5.1.3	Interpretation des Cohens-Kappa-Koeffizienten mit drei Ergebniskategorien der männlichen Stichprobe	51
5.1.4	Interpretation des Cohens-Kappa-Koeffizienten mit zwei Ergebniskategorien der gesamten Stichprobe	52
5.1.5	Interpretation des Cohens-Kappa-Koeffizienten mit zwei Ergebniskategorien der weiblichen Stichprobe	53
5.1.6	Interpretation des Cohens-Kappa-Koeffizienten mit zwei Ergebniskategorien der männlichen Stichprobe	53
5.1.7	Interpretation des Cohens-Kappa-Koeffizienten mit zwei Ergebniskategorien der asymptomatischen Stichprobe.....	54
5.1.8	Interpretation des Cohens-Kappa-Koeffizienten mit zwei Ergebniskategorien der symptomatischen Stichprobe.....	55
5.2	Beantwortung der Forschungsfrage.....	55
5.3	Stärken der Studie	56
5.3.1	Standardisierung des Tests	57
5.3.2	Vergleichbarer Ausbildungsstand der Testerinnen	57
5.3.3	Einheitliche Rahmenbedingungen.....	57
5.3.4	Breite Zusammensetzung der Stichprobe	57
5.3.5	Vielfältige Auswertungsansätze.....	57
5.3.6	Detaillierte statistische Auswertung	57
5.4	Einschränkungen im Bereich der angewandten Methodik	58
5.4.1	Limitierende Faktoren bezüglich der Literaturrecherche	58
5.4.2	Kritische Reflexion bezüglich des Studiendesigns	58
5.4.3	Kritische Reflexion bezüglich der Stichprobe	59
5.4.4	Limitierende Faktoren der Testerinnen.....	60
5.4.5	Durchführung und Interpretation des Tests.....	61
5.4.6	Einfluss äußerer Faktoren auf die Untersuchung	61
5.5	Vergleich mit anderen Studien.....	62
5.6	Übertragung der Ergebnisse auf die Osteopathie	63
6	Zusammenfassung und Schlussbetrachtung.....	66
6.1	Empfehlungen für weitere Studien.....	67

6.1.1	Fallzahlberechnung.....	67
6.1.2	Stichprobengröße	67
6.1.3	Intrarater-Reliabilität.....	68
6.1.4	Weiterentwicklung der Testmethode	68
6.2	Fazit.....	68
	LITERATURVERZEICHNIS	70
	ABBILDUNGSVERZEICHNIS	78
	TABELLENVERZEICHNIS.....	79
	ABKÜRZUNGSVERZEICHNIS	80
	ANHANG ÜBERSICHT	81
	ANHANG A	82
	ANHANG B	83
	ANHANG C	84
	ANHANG D.....	85
	ANHANG E	86
	ANHANG F	88
	ANHANG G.....	89

1 Einleitung

Kiefergelenksproblemen liegt meist eine lange pathogenetische Geschichte zugrunde, in der viele verschiedene Faktoren eine Rolle spielen (Liem, 2019, S. 172). Als Auslöser einer Dysfunktion der Mandibula werden in der Literatur unter anderem perinatale Traumata, die den gesamten Schädel betreffen, sowie andere Traumata beschrieben. Zusätzlich werden zahnärztliche und kieferorthopädische Eingriffe als mögliche Ursachen genannt. Weitere mögliche Ursachen sind psychoemotionale Faktoren, Dysfunktionen der Synchondrosis sphenobasilaris (SSB) und der oberen Halswirbelsäule sowie statische Störungen, myofasziale Fehlspannungen und viszerale Störungen (Liem & Dobler, 2018, S. 316; Peters, 2024, S. 385). Eine Dysfunktion des Os mandibulare kann sich nicht nur auf das Os temporale auswirken. Sie kann auch das Tentorium cerebelli, die Synchondrosis sphenobasilaris (SSB), die intrakraniellen Membranen, das Os hyoideum und das Os sacrum betreffen (Hinkelthein & Weitendorff, 2022, S. 241; Liem, 2019, S. 220).

Während in Bezug auf die kranio-mandibuläre Dysfunktion (CMD) Diagnosekriterien beschrieben werden, fehlen in der Literatur bislang klare diagnostische Kriterien für Dysfunktionen des Os mandibulare selbst (Liem, 2019, S. 239). Aufgrund der komplexen anatomischen und funktionellen Verbindungen von Strukturen des Unterkiefers gestaltet sich eine genaue Diagnostik oft schwierig (Corts, 2020, S. 157; Schünke et al., 2022, S. 73–92, 226–245). Daher sind zuverlässige Testverfahren zur Untersuchung des Os mandibulare von großer Bedeutung.

In der osteopathischen Diagnostik existieren bereits verschiedene Tests zur Untersuchung des Kiefergelenks, jedoch nur wenige spezifische Tests, mit denen sich die Beweglichkeit des Os mandibulare selbst ermitteln lässt. Die meisten diagnostischen Verfahren beziehen sich hingegen auf die Funktion des gesamten Kiefergelenks (*Articulatio temporomandibularis*) oder auch auf Fehlstellungen der Mandibula.

In der osteopathischen Literatur konnte lediglich der „Induktionstest Os mandibulare“ als spezifisches, jedoch bisher nicht standardisiertes Testverfahren zur passiven Bewegungsprüfung gefunden werden (Corts, 2020, S. 163). Dieser Test ermöglicht eine Beurteilung der Beweglichkeit des Os mandibulare in Außenrotation (Flexion) sowie

Innenrotation (Extension). Dabei wird untersucht, ob Asymmetrien, Bewegungseinschränkungen oder Kompressionen vorliegen.

Die Reliabilität osteopathischer Tests wird in der Forschung immer wieder kritisch diskutiert. So befassen sich unter anderem zwei frühere systematische Übersichtsarbeiten von Guillaud et al. (2016, 2018) damit, inwieweit eine zuverlässige Diagnosestellung sowie eine klinische Wirksamkeit in der kranialen und viszerale Osteopathie nachweisbar sind. Die Ergebnisse beider Studien zeigten, dass die Reliabilität diagnostischer Verfahren in diesen Bereichen generell unzureichend ist. Des Weiteren gab es keine Evidenz für die Wirksamkeit dieser Techniken.

Die Problematik der geringen Interrater-Reliabilität osteopathischer Tests wird auch durch Consorti et al. (2018) verdeutlicht. In ihrer Studie überprüften sie die Übereinstimmung von drei geschulten Osteopathiestudent_innen bei palpatorischen Tests am Sakrum anhand der sogenannten TART-Parameter (tissue texture abnormality, asymmetry, restriction of motion, tenderness). Die Ergebnisse zeigten nur eine geringe bis mäßige Übereinstimmung zwischen den Tester_innen ($\kappa = 0,06-0,34$), wobei der Tenderness-Parameter, also die durch Palpation ausgelöste Druckempfindlichkeit, die höchste Reliabilität aufwies. Damit verdeutlicht die Studie die eingeschränkte Reproduzierbarkeit palpatorischer Testergebnisse selbst bei standardisierter Vorgehensweise.

Die aktuelle Forschungslage wird unter anderem durch die Arbeiten von Consorti et al. (2018) und Guillaud et al. (2016, 2018) beschrieben. Aus diesen geht hervor, dass eine hohe Reliabilität in der osteopathischen Diagnostik wichtig ist, um sowohl die Qualität der Untersuchungen als auch die Patientensicherheit zu gewährleisten. Zudem können durch den Einsatz zuverlässiger diagnostischer Tests die Untersuchungsprozesse in der Praxis verbessert werden.

Die Mandibula ist mit vielen Strukturen funktionell verbunden. Daher ist eine zuverlässige Diagnostik in der Osteopathie besonders wichtig. In der aktuellen Literatur konnten bislang jedoch keine Untersuchungen zur Reliabilität des Induktionstests am

Os mandibulare gefunden werden. Dies verdeutlicht die Relevanz für eine weiterführende Forschung in diesem Bereich.

Der Induktionstest Os mandibulare wurde für diese Studie deswegen ausgewählt, weil er sich durch eine einfache und praktische Anwendung auszeichnet.

Die genannte Forschungslücke bildet die Grundlage der vorliegenden Masterarbeit. Das Ziel dieser Arbeit ist es, die Interrater-Reliabilität des Induktionstests Os mandibulare zu untersuchen.

Daraus ergibt sich folgende zentrale Forschungsfrage:

„Wie hoch ist die Übereinstimmung der Ergebnisse zweier Osteopathinnen mit gleichem Ausbildungsstand bei der passiven Untersuchung der Beweglichkeit der Mandibula in Außenrotation und Innenrotation mittels Induktionstest Os mandibulare?“

Zur Beantwortung dieser Frage wird folgende Hypothese geprüft:

- **Nullhypothese (H_0):** Die Übereinstimmung der Ergebnisse der beiden Testerinnen ist kleiner als der Cohens-Kappa-Koeffizient von 0,61 ($\kappa < 0,61$) und liegt somit unter dem beachtlichen Bereich (Landis & Koch 1977, S. 165).
- **Alternativhypothese (H_1):** Die Übereinstimmung der Ergebnisse der beiden Testerinnen ist gleich oder größer als der Cohens-Kappa-Koeffizient von 0,61 ($\kappa \geq 0,61$) und liegt somit im oder über dem beachtlichen Bereich (Grouven et al., 2007; Landis & Koch 1977, S. 165).

2 Stand der Wissenschaft

Dieses Kapitel gibt einen Überblick über den aktuellen wissenschaftlichen Kenntnisstand zur Funktion, Diagnostik und Therapie des Os mandibulare. Zudem werden die methodischen Grundlagen von Reliabilitätsstudien in der osteopathischen Diagnostik dargestellt.

2.1 Biomechanik und funktionelle Bedeutung des Os mandibulare

In diesem Abschnitt werden die biomechanischen Bewegungsmuster des Os mandibulare sowie deren funktionelle Bedeutung im Hinblick auf Belastungen, Stabilität und die Wechselwirkungen mit angrenzenden Strukturen beschrieben.

2.1.1 Biomechanik der Mandibula

Das Os mandibulare bewegt sich bei der kranialen Inspiration nach außen (Außenrotation) und bei der kranialen Exspiration nach innen (Innenrotation). Während der kranialen Inspiration bewegt sich die Kinnschuppe nach hinten (posterior), während die Anguli mandibulae nach außen (lateral) weichen und die Processus condylares nach innen (medial) wandern. Umgekehrt verlagert sich das Kinn bei der kranialen Exspiration nach vorne (anterior), wobei sich die Anguli mandibulae nach innen (medial) bewegen und die Processus condylares nach außen (lateral). Diese Rotation erfolgt um zwei vertikale Achsen, die sich auf der Höhe des ersten Backenzahns befinden (Corts, 2023, S. 38; Peters, 2024, S. 383).

2.1.2 Funktionelle Relevanz

Neben diesen osteopathisch beschriebenen Bewegungsmustern ist die Mandibula während der Kaufunktion verschiedenen mechanischen Belastungen ausgesetzt. Diese entstehen durch die Muskelzüge der Kaumuskulatur, die Gelenkbelastung im Kiefergelenk sowie den Kaukräften zwischen den Zahnreihen (Wong et al., 2010).

Wong et al. (2010) beschreiben außerdem, dass die Mandibula während jeder Kautätigkeit einer Kombination aus Biege- und Zugkräften am unteren Rand sowie Druckkräften am oberen Rand ausgesetzt ist. Die genaue Verteilung dieser Kräfte variiert jedoch je nach Position der Mandibula und Art der Belastung. In bestimmten

Situationen, wie etwa bei einer asymmetrischen Belastung oder nach strukturellen Defekten, können zusätzlich Rotationskräfte (Torsion) auftreten (Wong et al., 2010).

Auch Goh et al. (2008) betonen die mechanische Beanspruchung der Mandibula. So beschreiben sie, dass die Mandibula während der Kaufunktion eine tragende Rolle für die Stabilität des gesamten unteren Gesichtsschädels übernimmt. Bei Defekten oder nach Rekonstruktionen infolge von Traumata ist die Wiederherstellung der ursprünglichen Spannungsverhältnisse im Knochen ein zentrales Ziel. Eine ausreichende Stabilität ist deshalb erforderlich, um die bei der Kaufunktion auftretenden Kräfte aufzunehmen und weiterzuleiten. Die konkrete mechanische Belastung ist von der Lokalisation und dem Ausmaß eines Defekts abhängig (Goh et al., 2008).

2.2 Funktionelle Zusammenhänge

Aufbauend auf den beschriebenen biomechanischen Bewegungsmustern werden im Folgenden die wichtigsten funktionellen Verbindungen der Mandibula zu anderen anatomischen Strukturen dargestellt.

2.2.1 Mandibula und HWS

Die Kaumuskulatur stabilisiert das Kranium und unterstützt den Schluckakt. Allerdings können Fehlstellungen der Mandibula posturale Veränderungen sowie negative Auswirkungen auf die Körperstatik verursachen. Im Falle einer Prognathie (Vorverlagerung des Unterkiefers) kann eine verstärkte Okziput-Extension den Aufpressdruck im Kiefergelenk verändern, was Schmerzen sowie Knackgeräusche zur Folge hat. Dies kann wiederum eine Verspannung der subokzipitalen Muskulatur auslösen, die sich auf das visuelle System auswirken kann. Dadurch können neurovegetative Symptome wie Schwindel oder Kopfschmerzen hervorgerufen werden (Corts, 2023, S. 39). Zudem besteht eine neurologische Verbindung zwischen dem temporomandibulären Gelenk und der HWS, da der kaudale Trigeminuskern bis zu den Rückenmarkssegmenten C1 bis C3 reicht. Dadurch kann das Kiefergelenk die Haltung eines Menschen beeinflussen (Detoni et al., 2022).

Zusätzlich bestehen über die suprahyoidale Muskulatur und myofasziale Strukturen wie das Ligamentum stylohyoideum und den M. sternocleidomastoideus funktionelle

Verbindungen zwischen der Mandibula und der Halswirbelsäule. Diese ermöglichen eine mechanische Kopplung, durch die sich Spannungen oder Fehlstellungen wechselseitig beeinflussen können (Bordoni et al., 2018). Auch Hülse und Losert-Bruggner (2002) konnten aufzeigen, dass es infolge der reflektorischen Wechselwirkung zwischen dem Kopf- und dem Kiefergelenk bei einer Dysfunktion in einem Bereich auch zu Störungen im jeweils anderen kommen kann.

Neben der Verbindung mit der HWS spielt die Mandibula auch eine Rolle in der funktionellen Einheit mit der sphenobasilaren Synchondrose (SSB). Der M. pterygoideus medialis arbeitet als Mundschließer mit dem M. masseter und M. temporalis zusammen. Durch seinen Ursprung am Processus pterygoideus wird das Os sphenoidale ebenfalls in die Kaukräfte einbezogen. Dadurch können muskuläre Dysbalancen nicht nur zu Dysfunktionen der Mandibula, sondern auch zu Störungen des Os sphenoidale und der SSB führen (Corts, 2023, S. 39).

2.2.2 Mandibula und Os hyoideum

Die Mandibula steht zudem in einer engen funktionellen Verbindung mit dem Os hyoideum. In einer Studie von Jo et al. (2022) wurde die Distanz zwischen diesen beiden Strukturen untersucht. Hierbei wurde festgestellt, dass das Os hyoideum mit der Mandibula zusammenhängt und die Atemwege beeinflussen kann. Ein größerer Abstand zwischen beiden Strukturen kann wiederum die Zungenmuskulatur und die Stabilität der oberen Atemwege verändern. Wenn das Os hyoideum tiefer liegt, kann der Rachen leichter zusammensinken. Eine Vorverlagerung der Mandibula kann hingegen das Zungenbein anheben und die Atemwege offenhalten. Zudem wurde festgestellt, dass eine tiefere Position des Os hyoideum mit einer erhöhten Wahrscheinlichkeit für Atemaussetzer im Schlaf verbunden ist. Dies steht außerdem in Zusammenhang mit einer verringerten Sauerstoffsättigung und häufigeren Schlafunterbrechungen (Jo et al., 2022).

Die Position des Os hyoideum steht in direkter Verbindung mit der Position der Mandibula und der kraniocervikalen Haltung. Veränderungen hinsichtlich der Haltung der Halswirbelsäule können die Position des Zungenbeins beeinflussen, was wiederum

Auswirkungen auf die Stabilität und Funktion der Mandibula haben kann (Weber et al., 2012).

Des Weiteren steht das Os hyoideum in einer funktionellen Verbindung mit verschiedenen myofaszialen Diaphragmen des Körpers. Über diese faszialen Kontinuitäten kann es zu Wechselwirkungen zwischen dem Zungenbein, der Mandibula und anderen Strukturen wie dem Zwerchfell und dem Beckenboden kommen (Bordoni, 2020).

2.2.3 Mandibula und Temporomandibulargelenk (TMJ)

Das Kiefergelenk ist funktionell mit der Mandibula verbunden und lässt sich ebenfalls durch externe Faktoren beeinflussen. Im Rahmen einer Studie von Ohmure et al. (2008) konnte gezeigt werden, dass eine vorverlagerte Kopfhaltung („Forward Head Posture, FHP“) zu einer Verschiebung der Kondylenposition nach posterior führen kann. Dabei wurde beobachtet, dass sich die Aktivität der Kaumuskulatur, insbesondere des M. masseter und des M. digastricus, verändert. Diese Anpassungen könnten bei der biomechanischen Funktion des Kiefergelenks eine Rolle spielen (Ohmure et al., 2008).

2.2.4 Mandibula und Körperstatik

Neben ihren spezifischen Verbindungen mit anderen Strukturen wurde auch ein Einfluss der Mandibulaposition auf die gesamte Körperhaltung beobachtet. So ergab eine Untersuchung von Sakaguchi et al. (2007) zur Wechselwirkung zwischen der Mandibulaposition und der Körperhaltung, dass Veränderungen der Kieferstellung zugleich die Körperstatik beeinflussen können. In der Studie führte eine veränderte Kieferposition zu einer Verschiebung des Körperschwerpunkts. Zudem wurde umgekehrt festgestellt, dass eine veränderte Körperhaltung eine Verschiebung der Okklusionskräfte zur Folge haben kann (Sakaguchi et al., 2007).

2.2.5 Mandibula und Nervensystem

Auch dem Nervensystem kommt in der funktionellen Einheit der Mandibula eine wichtige Rolle zu. Demnach wird ein chronischer Gesichtsschmerz häufig mit Dysfunktionen des Kiefergelenks oder der Okklusion in Verbindung gebracht. Bei fortschreitender

Chronifizierung können zudem Einschränkungen im craniocervikalen und -sakralen System auftreten, die diesen Schmerz beeinflussen (Plato, 2001).

Der aurikulotemporale Nerv verläuft in unmittelbarer Nähe zum Kondylus der Mandibula und innerviert das Kiefergelenk sensorisch. Seine anatomische Position kann variieren, wodurch mechanische Belastungen zu Irritationen führen können (Fernandes et al., 2003).

2.2.6 Mandibula und fasziale Verbindungen

Das Os mandibulare ist über verschiedene fasziale Strukturen mit den umliegenden anatomischen Regionen verbunden. Dazu zählen die Fascia masseterica, die Fascia buccopharyngea, die Fascia cervicalis superior sowie die Aponeurosis pterygotemporomandibularis. Über die Linea mylohyoidea besteht außerdem eine Verbindung zur Schlundmuskulatur (Corts, 2023, S. 38).

Minelli et al. (2024) erläutern, dass die Fascia masseterica sowohl den M. masseter als auch die Parotisdrüse umfasst. Sie ist zudem mit der tiefen Fascia cervicalis verbunden. Dadurch ergibt sich eine anatomische Verbindung zwischen der Kiefer-, Gesichts- und Halsregion.

Funktionell ist das Os mandibulare über die Kaumuskulatur und die suprahyoidale Muskulatur in die tiefe myofasziale Frontallinie eingebunden. Zusätzlich besteht über die Fascia colli superficialis eine Verbindung zur sogenannten Bogensehne. Bei Spannungen oder auch Adhäsionen im Bereich der Fascia masseterica kann es zu lokalen Stauungen kommen. Die wechselnde Spannung des M. masseter während der Kaubewegung fördert durch einen mechanischen Ausstreichungseffekt die Drainage der Parotis (Corts, 2023, S. 38–40).

Im Rahmen einer anatomischen Studie mit Ultraschall- sowie MRT-Bildgebung konnten Pirri et al. (2021) aufzeigen, dass eine kontinuierliche Verbindung zwischen der Kaumuskulatur (M. masseter und M. temporalis) und dem M. sternocleidomastoideus besteht. Diese Verbindung wird als Teil einer craniocervikalen Einheit beschrieben, der eine bedeutende Rolle in der osteopathischen Diagnostik und Therapie einnimmt.

2.3 Dysfunktion des Os mandibulare

Eine Dysfunktion der Mandibula beschreibt eine eingeschränkte oder asymmetrische Beweglichkeit des Unterkiefers (Corts, 2020, S. 163). Nach dem osteopathischen Verständnis handelt es sich dabei um eine Form der somatischen Dysfunktion, das heißt um eine Funktionsstörung des Muskel-Skelett-Systems mitsamt seinen myofaszialen, vaskulären, lymphatischen und nervalen Komponenten (Coffey, 2005, S. 13).

Tramontano et al. (2021) definieren die somatische Dysfunktion anhand des sogenannten TART-Modells. Dieses umfasst die Aspekte, „Tissue texture changes“ (Gewebeveränderungen), „Asymmetry“ (Asymmetrie), „Restriction of motion“ (Bewegungseinschränkungen) sowie „Tenderness“ (Druckempfindlichkeit) und bildet die Grundlage für die osteopathische Diagnostik.

2.3.1 Symptome

Zu den Symptomen einer mandibulären Dysfunktion zählen Zahnschmerzen, allgemeine Schmerzen im Bereich des Kiefergelenks sowie Knack- und Knirschgeräusche. Darüber hinaus können Blockierungen in der OAA-Region (Occiput-Atlas-Axis), Kopfschmerzen und Verspannungen der Kaumuskulatur, auftreten. Weitere mögliche Beschwerden sind Trigeminusneuralgien, Bruxismus sowie statische Probleme. (Corts, 2020, S. 162; Peters, 2024, S. 385).

2.3.2 Ursachen

Die Ursachen einer mandibulären Dysfunktion sind aufgrund der zahlreichen anatomischen und funktionellen Zusammenhänge der Strukturen vielfältig und komplex (Corts, 2020, S. 162; Liem & Dobler, 2018, S. 316–317; Peters, 2024, S. 385). Als Auslöser hierfür werden unter anderem Traumata wie Gähntraumata, zahnärztliche Eingriffe und Operationen genannt. Auch funktionelle Störungen, beispielsweise der Synchrondrositis sphenobasilaris (SSB) oder der oberen Halswirbelsäule (HWS), können eine Rolle spielen. Ebenso können Bissstörungen, statische Ungleichgewichte, psychoemotionale Belastungen, myofasziale Dysbalancen und viszerale Ursachen zur Entstehung beitragen (Corts, 2020, S. 162; Liem & Dobler, 2018, S. 316–317; Peters, 2024, S. 385).

Ergänzend dazu beschreiben Liem und Dobler (2018, S. 316–317) verschiedene Formen von Dysfunktionen am Os mandibulare, wobei sie zwischen primären und sekundären Ursachen unterscheiden. Zu den primären Ursachen zählen angeborene Veränderungen oder traumatische Einwirkungen. Hierzu zählen zum Beispiel perinatale Traumata, Stürze, Schläge auf das Gesicht oder auch zahnärztliche Eingriffe. Sekundäre Ursachen hingegen entstehen häufig durch andere Funktionsstörungen. Dazu gehören Veränderungen der Schädelknochen, Bissstörungen, kieferorthopädische Geräte, Fehlhaltungen oder auch psychischer Stress. All diese Faktoren können die Beweglichkeit des Os mandibulare beeinflussen.

Des Weiteren führen Liem und Dobler (2018, S. 316–317) Ursachen wie eine Fixierung des Kondylus an. Je nach Lage ob anterior, posterior oder lateral, können diesbezüglich unterschiedliche Symptome auftreten. Dazu zählen zum Beispiel venöse Stauungen, Hörstörungen, Schluckbeschwerden oder eine eingeschränkte Kieferbewegung. Auch ein Frühbiss, fehlende Zahnkontakte oder ein muskuläres Ungleichgewicht können eine Rolle spielen. Besonders häufig sind dabei der M. masseter, der M. temporalis und der M. pterygoideus beteiligt.

Darüber hinaus gibt es angeborene oder auch erworbene Entwicklungsstörungen der Mandibula, die die Form und Funktion des Unterkiefers beeinflussen können. Solche Fehlentwicklungen führen häufig zu Asymmetrien, Bissproblemen oder Bewegungseinschränkungen (Hofmann et al., 2023).

2.4 Differenzialdiagnostik Craniomandibuläre Dysfunktion (CMD)

Während eine Dysfunktion der Mandibula primär die Beweglichkeit des Unterkiefers betrifft, handelt es sich bei der craniomandibuläre Dysfunktion (CMD) um eine komplexe Funktionsstörung des gesamten craniomandibulären Systems. Die CMD entsteht durch eine Regulationsstörung zwischen Kiefer, Kaumuskulatur sowie Zentralnervensystem und kann Schmerzen und Kiefergelenkgeräusche ebenso wie Bewegungseinschränkungen hervorrufen (Prodingler-Glöckl, 2013; Reißmann, 2017). Da sich die vorliegende Arbeit jedoch ausschließlich mit der Mandibula befasst, wird die CMD hier lediglich zur Unterscheidung vom eigentlichen Untersuchungsgegenstand beschrieben.

2.5 Epidemiologie

Im Vergleich zu anderen Strukturen des Kiefergelenks ist die Datenlage zur Mandibula bislang begrenzt. Daher werden in diesem Kapitel ergänzend auch Ergebnisse aus Studien zur craniomandibulären Dysfunktion (CMD) berücksichtigt, da die Mandibula hierbei funktionell mitbetroffen ist.

Laut Thiesen et al. (2016) beträgt die Prävalenz der mandibulären Asymmetrie in einer erwachsenen Population 17,4 % (95 %-Konfidenzintervall: 15,2–20,0). Ihre Untersuchung zeigte jedoch keine signifikanten Zusammenhänge zwischen mandibulärer Asymmetrie und dem Geschlecht, Alter oder einem Zahnverlust. Allerdings bestand ein statistisch signifikanter Zusammenhang mit der sagittalen Kieferrelation ($p = 0.045$).

Neben strukturellen Fehlstellungen wie einer mandibulären Asymmetrie sind auch allgemeine Malokklusionen weit verbreitet. Beispielsweise untersuchten Asiri et al. (2019) die Prävalenz von Zahnfehlstellungen in der erwachsenen Bevölkerung. Hierbei stellten sie fest, dass zwei Drittel der Erwachsenen mindestens eine Form einer klinisch bedeutsamen Zahnfehlstellung aufwiesen. Mit zunehmendem Alter nahm die Häufigkeit dieser Fehlstellungen zu. Zudem bestanden signifikante Unterschiede zwischen den Geschlechtern und ethnischen Gruppen, wobei Männer vermehrt Unregelmäßigkeiten im Unterkiefer zeigten.

Neben strukturellen Fehlstellungen spielen auch funktionelle Einschränkungen der Kieferbeweglichkeit eine wichtige Rolle. Studien zur CMD geben weitere Hinweise auf deren Häufigkeit und Ausprägung. Laut der Untersuchung von Lövgren et al. (2016), verändern sich Schmerzen sowie Funktionsstörungen des Kiefergelenks im Laufe des Lebens. Zudem sind Frauen im Vergleich zu Männern häufiger betroffen. So wurden wiederkehrende Kieferschmerzen bei 5,2 % der Frauen und 1,8 % der Männer festgestellt, schmerzhaftes Kieferbewegungen hingegen bei 2,5 % der Frauen bzw. 0,9 % der Männer.

Des Weiteren gibt Reißmann (2017) an, dass etwa 10% der allgemeinen Bevölkerung an einer craniomandibulären Dysfunktion (CMD) leiden. Zudem liege die

Neuerkrankungsrate (Inzidenz) bei etwa 3 % pro Jahr. Andere Studien untersuchten dabei spezifische Formen, wie beispielsweise temporomandibuläre Gelenkerkrankungen (TMJD). Eine systematische Übersichtsarbeit und Metaanalyse von Valesan et al. (2021) zeigte eine Gesamtprävalenz von 31,1 % für temporomandibuläre Gelenkerkrankungen (TMJD) bei Erwachsenen. Hierbei war die häufigste Diagnose die Diskusverlagerung mit Reposition, die bei 25,9 % der Erwachsenen vorkam. Diese Ergebnisse unterstreichen die hohe Verbreitung von strukturellen und funktionellen Kiefergelenkstörungen.

2.6 Osteopathische Therapie und Behandlung

Die osteopathische Behandlung der Mandibula richtet sich nach dem im betreffenden Fall vorliegenden Dysfunktionsmuster. Im Folgenden werden die häufigsten osteopathischen Befundmuster zusammengefasst und deren jeweilige Behandlungstechniken beschrieben.

2.6.1 Behandlung der Dysfunktion in Außenrotation

Bei einer Dysfunktion der Mandibula in Außenrotation liegt der_ die Patient_in in Rückenlage. Der_ die Therapeut_in sitzt am Kopfende der Liege und führt den Kinngriff durch. Während der kranialen Inspirationsphase wird die Außenrotation des Os mandibulare durch einen sanften Schub in kranial-posteriorer Richtung unterstützt. In der kranialen Expirationsphase hingegen wird die Innenrotation gebremst. Anschließend wird der Point of Balance gesucht und die Gewebereaktion beobachtet. Sobald eine Entspannung des Gewebes (Softening) eintritt und der primär respiratorische Rhythmus (PRM) wieder einsetzt, wird die Bewegung über mehrere Zyklen hinweg begleitet. Zuletzt erfolgt eine erneute Beurteilung der Amplitude und Frequenz (Peters, 2024, S. 386–387).

2.6.2 Behandlung der Dysfunktion in Innenrotation

Die Vorgehensweise bei einer Dysfunktion der Mandibula in Innenrotation ist grundsätzlich ähnlich wie bei der Behandlung einer Außenrotationsdysfunktion. Der Unterschied liegt in der Bewegungsausführung, wobei in der kranialen Expirationsphase die Innenrotation unterstützt wird und in der Inspirationsphase die Außenrotation gebremst wird (Peters, 2024, S. 387).

2.6.3 Behandlung der Kompressionsdysfunktion des kranio-mandibulären Komplexes

Auch bei der Behandlung einer Kompressionsdysfunktion des kranio-mandibulären Komplexes liegt der_ die Patient_in in Rückenlage. Der_ die Therapeut_in legt seine_ ihre Hände in entsprechender Ausgangsstellung auf das Os mandibulare und übt durch eine Supination der Hände eine sanfte Kompression in die kraniale Richtung aus. Dabei werden die Reaktionen des Kiefergelenks, der angrenzenden Suturen (Suturæ parietosquamosa und parietomastoidea) sowie des Tentoriums und der duralen Verbindung bis hin zum Sakrum beobachtet. Nach erfolgter Entspannung wird zur Dekompression wiederum durch die Pronation der Hände ein kaudal-anterior gerichteter Zug eingeleitet. Diese Technik wird bis zum Still Point begleitet. Nach dem Einsetzen des PRM wird die Bewegung über mehrere Zyklen hinweg beobachtet (Peters, 2024, S. 387).

2.6.4 Behandlung der artikulären Dysfunktionen

Artikuläre Dysfunktionen beziehen sich auf die Stellung und Beweglichkeit der Kondylen im Kiefergelenk. Sie können sowohl anterior als auch posterior ausgeprägt sein und entweder uni- oder bilateral auftreten. Nachfolgend werden die spezifischen Techniken zur Behandlung anteriorer und posteriorer artikulärer Dysfunktionen erläutert.

2.6.4.1 Behandlung der anterioren Dysfunktion

Bei der anterioren Dysfunktion befindet sich die Kondyle vor dem Tuberculum articulare. Diese Form ist häufig schmerzhaft. Eine bilaterale Ausprägung tritt typischerweise nach einem Gähntrauma auf. Der_ die Patient_in liegt in Rückenlage, der_ die Therapeut_in steht vor dem_ der Patient_in und nimmt den sogenannten „Pistolengriff“ ein. Dabei liegen die Daumen von kranial auf den Zahnreihen, während die Zeigefinger die Mandibula von kaudal stabilisieren. Zunächst wird eine Traktion im posterior-kaudalen Bereich ausgeführt und so lange gehalten, bis eine Gewebeentspannung eintritt. Anschließend folgt ein Schub nach posterior und eine kraniale Bewegung der Kondylen in Form einer Schubladenbewegung.

Bei einer unilateralen Dysfunktion wird diese Technik ausschließlich auf der betroffenen Seite angewendet. Hierbei erfolgt der Schub in latero-posteriorer Richtung (Peters, 2024, S. 387).

2.6.4.2 Behandlung der posterioren Dysfunktion

Die posteriore Dysfunktion entsteht meist infolge einer chronischen Fehlbelastung oder degenerativer Veränderungen. Dabei befindet sich die Kondyle hinter dem Diskus, was zu Schmerzen beim Essen, Sprechen oder Öffnen des Mundes führen kann. Bei der Behandlung liegt der/die Patient_in wie zuvor beschrieben in Rückenlage. Bei einer bilateralen Variante beginnt die Behandlung mit einer Traktion zur Entspannung der umliegenden Strukturen. Gegebenenfalls werden die Kondylen dekomprimiert und die Kaumuskulatur, insbesondere die Mm. pterygoidei medialis und lateralis, wird gelöst. Anschließend wird das Os mandibulare nach anterior, kranial und posterior mobilisiert, bis der Diskus wieder über dem Kondylus liegt.

Bei einer unilateralen posterioren Dysfunktion erfolgt die Technik hingegen nur auf der betroffenen Seite. Hierbei geht der Schub nach anterior-medial, um die Reposition des Diskus zu ermöglichen (Peters, 2024, S. 388).

2.6.5 Entspannung der Kaumuskulatur

Auch die Entspannung der Kaumuskulatur stellt einen wichtigen Bestandteil der osteopathischen Behandlung dar. Hierbei können der M. temporalis und der M. masseter durch verschiedene Techniken wie Quermassagen oder durch postisometrische Relaxation (PIR) behandelt werden. Intraoral können hingegen die Mm. pterygoidei medialis und lateralis behandelt werden, um Spannungen in der Tiefe zu lösen (Peters, 2024, S. 388–389).

2.7 Diagnostische Verfahren zur Untersuchung der Mandibula

In diesem Unterkapitel werden verschiedene diagnostische Verfahren beschrieben, die zur Untersuchung der Mandibula eingesetzt werden können. Hierzu zählen sowohl bildgebende Methoden als auch manuelle Untersuchungsverfahren.

2.7.1 Bildgebende Verfahren

Für die diagnostische Beurteilung funktioneller bzw. struktureller Veränderungen des Os mandibulare und des temporomandibulären Gelenks (TMG) stehen verschiedene bildgebende Verfahren zur Verfügung. Da die Mandibula funktionell eng mit dem Kiefergelenk, der umgebenden Muskulatur und dem Discus articularis verbunden ist, beziehen sich viele bildgebende Verfahren auf das gesamte kranio-mandibuläre System.

Die Auswahl der Methode richtet sich nach der klinischen Fragestellung und dem zu untersuchenden Gewebe (das heißt, entweder Knochengewebe oder Weichteilgewebe).

2.7.1.1 Konventionelle Verfahren (OPT)

Zu den konventionellen Methoden zählen unter anderem das Orthopantomogramm (OPT), auch „Panoramaröntgen“ genannt, und die Digitale Volumentomographie (DVT). Diese Verfahren eignen sich insbesondere zur Beurteilung der knöchernen Strukturen wie der Kondylenform, Asymmetrien oder auch von knöchernen Umbauprozessen der Mandibula. Dabei ermöglicht die DVT eine dreidimensionale Darstellung bei geringer Strahlenbelastung. Allerdings ist eine Darstellung der Weichteile oder des Discus articularis mit diesen Verfahren nicht möglich (Brinkmann, 2022).

2.7.1.2 Magnetresonanztomographie (MRT)

Zur Beurteilung der Weichteilstrukturen im Bereich des temporomandibulären Gelenks, insbesondere für den Discus articularis, gilt die Magnetresonanztomographie (MRT) derzeit als diagnostischer Goldstandard. Diese Methode ist nicht invasiv, strahlungsfrei und erlaubt eine differenzierte Beurteilung des Discus articularis, des hinteren Bandapparats, der Gelenkkapsel sowie möglicher Entzündungen oder auch Gelenkergüsse. Insbesondere bei Verdacht auf Diskusverlagerungen oder diskogene Schmerzsyndrome wird die MRT als Methode der Wahl empfohlen (Emshoff & Rudisch, 2001).

Orhan et al. (2017) weisen in ihrer Übersichtsarbeit darauf hin, dass die MRT vor allem dafür geeignet ist, komplexe Zusammenhänge zwischen klinischer Symptomatik und strukturellen Veränderungen sichtbar zu machen. Da auffällige MRT-Ergebnisse auch bei beschwerdefreien Patient_innen vorkommen, ist jedoch eine zusätzliche Abstimmung mithilfe einer klinischen Untersuchung notwendig.

2.7.1.3 Neuere Verfahren und technische Entwicklungen

Neben der klassischen MRT werden zunehmend auch dynamische MRT-Verfahren eingesetzt. So wurde im Rahmen der Dissertation von Brinkmann (2022) der Einsatz sogenannter „Realtime-MRT-Techniken“ beschrieben, mit denen der Bewegungsablauf

des Discus während der Kieferöffnung beurteilt werden kann. Dies ist insbesondere bei funktionsabhängigen Beschwerden hilfreich.

Kopp et al. (2024) zeigen wiederum, dass auch MRT-Geräte mit niedriger Feldstärke (z. B. 0,55 T) zur Beurteilung des Kiefergelenks verwendet werden können. Zwar ist die Bildqualität im Vergleich zu Hochfeldsystemen (1,5 T oder 3 T) reduziert, jedoch reicht sie für klinische Zwecke aus. Limchaichana et al. (2006) weisen jedoch in ihrer systematischen Übersicht darauf hin, dass weitere hochwertige Studien notwendig sind, um die diagnostische Genauigkeit verschiedener MRT-Techniken weiter abzusichern.

Eine zusammenfassende Übersicht über die wichtigsten bildgebenden Verfahren zur Beurteilung des Kiefergelenks und der Mandibula ist in der nachstehenden Tabelle 1 dargestellt. Hierbei werden die Darstellbarkeit, die Strahlenbelastung sowie Besonderheiten der jeweiligen Methoden verglichen.

Tabelle 1: Übersicht bildgebender Verfahren

Verfahren	Darstellbare Strukturen	Strahlung	Besonderheiten	Quelle
Orthopantomogramm (OPT, Panoramaröntgen)	Knochen (Mandibula, Kondylen)	ja	geringe Aussagekraft für Weichteile	Brinkmann, 2022
Digitale Volumentomographie (DVT)	Hochauflösende Knochenbilder	ja	gute 3D-Darstellung, keine Weichteildiagnostik	Brinkmann, 2022
MRT	Weichteile, Diskus, Gelenkkapsel	nein	Goldstandard, gute Differenzierung, hohe Kosten	Emshoff & Rudisch, 2001
Realtime-MRT	Bewegungsanalyse des Discus	nein	dynamische Untersuchung möglich	Brinkmann, 2022
Low-Field-MRT	Gelenkstrukturen	nein	für klinische Diagnostik ausreichend, kostengünstiger	Kopp et al., 2024

2.7.2 Manuelle Diagnostik der Mandibula

Die manuelle Untersuchung der Mandibula wird in der osteopathischen Literatur nur selten beschrieben. Meist liegt der Fokus stattdessen auf dem Kiefergelenk (TMJ). Das Os mandibulare selbst wird hingegen meist indirekt über Palpation, oder auch durch das Bewegungsverhalten beurteilt. Ein standardisiertes Testverfahren existiert zurzeit noch nicht. In der osteopathischen Praxis wird jedoch der passive Induktionstest genutzt, um die Beweglichkeit der Mandibula zu beurteilen und Hinweise auf mögliche Einschränkungen zu erhalten.

2.7.2.1 Aktive, passive und spezifische Tests

In der kraniosakralen Osteopathie beschreibt die Induktion eine diagnostische Methode, bei der gezielte Bewegungen auf verschiedenen Ebenen auf eine bestimmte Struktur ausgeübt werden. Daraufhin beobachtet der_ die Therapeut_in die Reaktion dieser Struktur auf die induzierte Bewegung, um Hinweise auf potenzielle funktionelle Störungen oder Dysfunktionen zu erhalten. Diese Technik stellt eine Ergänzung für andere Verfahren wie das Listening und den Elastizitätstest dar (Peters, 2024, S. 301).

Zur Beurteilung der Mandibula kann der passive Induktionstest Os mandibulare eingesetzt werden. Da die Mandibula funktionell eng mit dem temporomandibulären Gelenk (TMJ), der Kaumuskulatur sowie faszialen Strukturen verbunden ist, zielen die meisten manuellen Testverfahren auf das gesamte kranio-mandibuläre System ab (Corts, 2020, S. 163; Peters, 2024, S. 386).

Die manuelle Diagnostik umfasst verschiedene Testarten, die sich grundsätzlich in aktive und passive Verfahren unterteilen lassen:

Aktive Tests: Der_ die Patient_in führt die Bewegung eigenständig aus, während der_ die Tester_in eventuelle Abweichungen, Schmerzen oder auch Geräusche beobachtet. Ein Beispiel hierfür ist der aktive Kieferöffnungstest, bei dem der_ die Patient_in seinen_ ihren Mund öffnet und der_ die Tester_in auf Deviationen bzw. Krepitationen achtet (Peters, 2024, S. 386).

Passive Tests: Bei passiven Verfahren wiederum führt der_ die Tester_in selbst die Bewegung aus, um das Gelenkspiel und die Weichteilspannung zu beurteilen. Hierzu zählt beispielsweise der passive 3D-Kiefergelenkstest, bei dem die Protrusion, Retrusion und Laterotrusion auf mögliche Bewegungseinschränkungen überprüft werden (Corts, 2020, S. 99).

Spezifische Gelenktests: Zur Differenzierung intraartikulärer Strukturen kommen hingegen passive Kompressionstests zum Einsatz, die gezielt auf die Gelenkflächen oder die hintere Gelenkszone (bilaminäre Zone) des TMJ abzielen (Liem, 2019, S. 252).

Da ein einzelner Test in der Regel jedoch nicht ausreicht, um eine Dysfunktion sicher zu diagnostizieren, empfiehlt es sich laut Patijn (2019), mehrere Tests miteinander zu kombinieren. Dies ermöglicht eine umfassendere funktionelle Beurteilung und erhöht auf diese Weise die diagnostische Sicherheit.

2.7.2.2 Induktionstest Os mandibulare nach Corts

Zuletzt soll an dieser Stelle insbesondere auf den passiven Induktionstest Os mandibulare nach Corts (2020, S. 163) eingegangen werden. Dieser Test dient der qualitativen Beurteilung der Beweglichkeit der Mandibula in Bezug auf die Amplitude, Frequenz und Symmetrie während der kranialen Atembewegung (CRI). Eine eingeschränkte oder auch asymmetrische Bewegung kann auf eine mögliche Dysfunktion hinweisen. Corts (2020, S. 163) nennt keine Kontraindikationen für die Durchführung dieses Tests. Eine ähnliche Testbeschreibung findet sich ebenfalls bei Peters (2024, S. 386). Die Durchführung und Zielsetzung gleichen im Grunde jenen bei Corts, jedoch unterscheiden sie sich leicht hinsichtlich der Terminologie und der Beschreibung der Bewegungsinduktion.

Da der Induktionstest Os mandibulare mit dem CRI zusammenhängt, stellt sich auch die Frage, wie weit dieser tatsächlich messbar ist. Dazu konnten Masoudi et al. (2024) in ihrer Studie zeigen, dass die Frequenz des Cranial Rhythmic Impulse (CRI) mithilfe einer dynamischen MRT zuverlässig messbar ist. Allerdings ließ sich die Amplitude nur an einer bestimmten Messstrecke, und zwar zwischen Sella turcica und Glabella, genau bestimmen. An anderen Punkten war der Messfehler hingegen zu hoch. Demnach deuten die Ergebnisse darauf hin, dass osteopathische Bewegungen wie der CRI zwar teilweise objektiv messbar sind, ihre Aussagekraft jedoch abhängig von der jeweiligen Struktur und Methode unterschiedlich ist.

Eine Übersicht über die wichtigsten Testverfahren ist in der folgenden Tabelle 2 dargestellt.

Tabelle 2: Übersicht osteopathischer Testverfahren

Testart	Testmethode	Durchführung	Zielstruktur	Quelle
Aktiver Test	Kieferöffnungstest	Patient_in öffnet den Mund selbstständig	Muskulatur, Kiefer, TMG	Peters, 2024
Passiver Test	3D- Kiefergelenkstest	Tester_in bewegt Kiefer passiv	TMG, Kiefer, Weichtteile	Corts, 2020
Spezifischer Gelenkstest	Kompression der bilaminären Zone	gezielte Kompression durch Tester_in	hintere Gelenkzone (TMG)	Liem, 2019
Induktionstest Os mandibulare	Palpation der mandibulären Eigenbewegung im Rahmen des kranialen Rhythmus	palpatorische Wahrnehmung durch Tester_in	Mandibula	Corts, 2020; Peters, 2024

2.8 Gütekriterien und methodologische Grundlagen von Interrater-Reliabilitätsstudien

In diesem Kapitel werden die zentralen Gütekriterien diagnostischer Verfahren behandelt und die Bedeutung der Interrater-Reliabilität für die Beurteilung manualtherapeutischer Tests erläutert.

2.8.1 Gütekriterien und Messinstrumente

Tests sind häufig Gegenstand der wissenschaftlichen Forschung. Dabei wird auch untersucht, wie aussagekräftig die Ergebnisse eines Tests sind. Ein wesentlicher Aspekt sind die Gütekriterien, die ein Messinstrument erfüllen muss. Hierzu werden in der Methodologie insbesondere die Objektivität, die Validität und die Reliabilität angegeben (Hussy et al., 2013, S. 23–24).

Objektivität beschreibt, wie unabhängig die Ergebnisse eines diagnostischen Verfahrens von dem_der Tester_in sind und ob sie generell frei von subjektiven Einflüssen bleiben. Das Kriterium der Objektivität ist eine wesentliche Voraussetzung für die Erfüllung der weiteren Gütekriterien (Hussy et al., 2013, S. 23–24). Validität wiederum bezieht sich auf die Gültigkeit eines Messinstrumentes, also darauf, wie gut es tatsächlich das misst, was es zu messen vorgibt (Scherfer et al., 2011, S. 242).

Reliabilität bezeichnet die Zuverlässigkeit eines Messinstrumentes. Sie ist dann gegeben, wenn bei wiederholten oder gleichzeitigen Messungen unter konstanten Bedingungen gleiche Ergebnisse erzielt werden (Scherfer et al., 2011, S. 242f). Hierbei kann zwischen zwei unterschiedlichen Formen unterschieden werden, nämlich der Intrarater- und der Interrater-Reliabilität. Die Intrarater-Reliabilität beschreibt, ob ein_e Tester_in bei mehrfacher Beurteilung desselben Merkmals unter gleichbleibenden Bedingungen konstante Ergebnisse erzielt. Während die Interrater-Reliabilität untersucht, ob zwei oder mehr Tester_innen bei der Beurteilung desselben Tests zu den gleichen Ergebnissen kommen (Scherfer et al., 2011, S. 250).

Zusätzlich wird bei Scherfer et al. (2011, S. 243f) auf die Praktikabilität hingewiesen, die zwar kein wissenschaftliches Gütekriterium darstellt, aber dennoch einen wichtigen Aspekt in Bezug auf die Durchführbarkeit und die Praxistauglichkeit einer Messmethode betrifft. Im Praxisalltag ist es wichtig, dass Tests und Untersuchungen mit möglichst geringem Aufwand durchgeführt werden können (Scherfer et al., 2011, S. 250).

2.8.2 Design und Durchführung von Interrater-Reliabilitätsstudien

Interrater-Reliabilitätsstudien sind ein wesentlicher Bestandteil der diagnostischen Forschung. Das Ziel solcher Studien ist es, Messinstrumente hinsichtlich ihrer Gütekriterien zu überprüfen, insbesondere auf Reliabilität und Objektivität (Scherfer et al., 2011, S. 242). Ein häufig verwendetes Studiendesign besteht darin, dass zwei oder mehr Tester_innen unabhängig voneinander dieselbe Struktur beurteilen. Daraufhin werden die Ergebnisse statistisch auf ihre Übereinstimmung analysiert. Dabei kommen Kennwerte wie der Cohens-Kappa-Koeffizient oder auch die Intraklassenkorrelation (ICC) zum Einsatz. Die International Academy of Manual/Musculoskeletal Medicine (IAMMM) hat einen Leitfaden für die Durchführung von Reliabilitäts- bzw.

Validitätsstudien veröffentlicht (Patijn, 2019), der als die methodologische Grundlage dieser Arbeit dient.

2.9 Forschungsstand zur Reliabilität osteopathischer Tests

In mehreren Studien wurde gezeigt, dass die Interrater-Reliabilität osteopathischer Tests stark von der untersuchten Körperregion, dem Erfahrungsniveau der Tester_innen und der Standardisierung des Testverfahrens abhängt (Consorti et al., 2018; Degenhardt et al., 2005; Guillaud et al., 2016, 2018; Nolet et al., 2021). Besonders im Bereich der kranialen und viszeralen Osteopathie konnten Guillaud et al. (2016, 2018) nur geringe bis mäßige Übereinstimmungswerte feststellen. Ähnliche Ergebnisse berichteten Consorti et al. (2018) bei palpatorischen Tests am Sakrum, wobei die Cohens-Kappa-Werte zwischen 0,06 und 0,34 lagen und somit lediglich eine geringe bis mittlere Reliabilität zeigten.

In einer weiteren Studie von Degenhardt et al. (2005) wurde eine kontrollierte Untersuchung zur Interobserver-Reliabilität von acht osteopathischen Tests an der Lendenwirbelsäule durchgeführt. Nach einer initialen Testphase mit geringen Übereinstimmungen (Cohens-Kappa-Werte zwischen $-0,02$ und $0,34$) zeigte sich nach einem gezielten Training eine signifikante Verbesserung der Reliabilität. Insbesondere bei den Tests zur Beurteilung der Gewebetextur ($\kappa = 0,45$) und des Druckschmerzes ($\kappa = 0,68$) konnten akzeptable Übereinstimmungswerte erreicht werden.

Eine systematische Übersichtsarbeit von Nolet et al. (2021) untersuchte wiederum die Reliabilität und Validität palpatorischer Tests bei Patient_innen mit lumbalen Beschwerden. Die Autor_innen analysierten insgesamt 49 Studien zu Weichteil-, Gelenk- und Bewegungspalpation. Die Ergebnisse zeigten, dass die Interrater-Reliabilität insbesondere für die Weichteil- und Gelenkpalpation uneinheitlich ist und nur vereinzelt hohe Cohens-Kappa-Werte ($> 0,6$) erreicht werden konnten. Zudem wurde auf methodische Schwächen im Bereich der Standardisierung und Schulung der Tester_innen hingewiesen.

Insgesamt ist der aktuelle Forschungsstand zur Reliabilität osteopathischer Tests als heterogen einzuschätzen. Demnach unterscheiden sich die Ergebnisse je nach Art des

Tests, der untersuchten Körperregion, der Qualifikation der Tester_innen und dem methodischen Aufbau der Studien. Trotz dieser Unterschiede sind sich allerdings viele Autor_innen darin einig, dass klar definierte Testabläufe, gezielte Schulungen und gut geplante Studien wichtig sind, um die Verlässlichkeit palpatorischer Verfahren in der Osteopathie zu verbessern (Degenhardt et al., 2005; Guillaud et al., 2016, 2018; Nolet et al., 2021).

3 Forschungsdesign und -methode

In diesem Kapitel werden das Forschungsdesign und die methodische Vorgehensweise der Interrater-Reliabilitätsstudie beschrieben. Ziel war die Überprüfung des Induktionstests Os mandibulare.

3.1 Stichprobenbeschreibung

Die Stichprobe wurde anhand vordefinierter Ein- und Ausschlusskriterien zusammengestellt. Dabei wurden sowohl symptomatische als auch asymptotische Proband_innen berücksichtigt, um eine möglichst praxisnahe Auswahl zu gewährleisten. Insgesamt nahmen 42 Proband_innen im Alter zwischen 30 und 65 Jahren an der Studie teil. Etwas mehr als die Hälfte der Proband_innen war asymptotisch, während rund 45 % Symptome aufwiesen. Dadurch wurde eine für den klinischen Alltag typische Verteilung erreicht.

3.2 Proband_innen

Die Proband_innen wurden über eine Ausschreibung in verschiedenen physiotherapeutischen Praxen rekrutiert. Zusätzlich wurden weitere Proband_innen aus dem Bekannten- und Verwandtenkreis der Autorin zur Studie herangezogen.

Interessierte meldeten sich per E-Mail bei der Autorin. Um das Einverständnis sowie die Ein- und Ausschlusskriterien zu überprüfen, wurden den Proband_innen ein Fragebogen und ein Aufklärungsbogen zugeschickt.

Wenn sämtliche Ausschlusskriterien mit „Nein“ beantwortet wurden, erhielt der_die jeweilige Proband_in einen Termin für die Testung. Alle Proband_innen nahmen freiwillig an der Studie teil.

3.3 Ein- und Ausschlusskriterien

Vorab wurden die Ein- und Ausschlusskriterien mittels des Fragebogens, per E-Mail oder auch per Telefonat abgefragt und überprüft.

3.3.1 Einschlusskriterien

Um eine möglichst praxisnahe Stichprobe zu gewährleisten, wurden in die Studie sowohl symptomatische als auch asymptotische Erwachsene einbezogen. Dadurch sollte sichergestellt werden, dass das Testverfahren sowohl bei Proband_innen mit

Beschwerden als auch bei beschwerdefreien Proband_innen angewendet werden kann. Dieses Vorgehen erhöht die externe Validität der Ergebnisse und ermöglicht eine bessere Übertragbarkeit auf den klinischen Alltag.

Das Mindestalter wurde auf 18 Jahre festgelegt, da bei Minderjährigen sowohl rechtliche als auch organisatorische Aspekte, wie etwa die Einholung der Einverständniserklärung der Eltern, sowie physiologische Unterschiede im Bereich des Schädel- und Kieferwachstums zu berücksichtigen sind. Diese Entwicklungsprozesse können die anatomischen und funktionellen Eigenschaften des kranio-mandibulären Systems beeinflussen und somit die Testergebnisse verändern (Schünke et al., 2022, S. 46–49).

3.3.2 Ausschlusskriterien

Zu den Ausschlusskriterien zählten akute Verletzungen, Traumata oder Operationen im Bereich des Os mandibulare oder des Schädelbereichs. Solche Zustände können die lokale Gewebespannung, Beweglichkeit und Schmerzempfindung verändern und dadurch die Palpationsergebnisse verfälschen.

Ein Ausschluss erfolgte dann, wenn der Eingriff weniger als sechs Wochen zurücklag oder keine ärztliche Abklärung vorlag. Dadurch sollten mögliche Komplikationen wie zerebrale Blutungen vermieden werden. Ebenso wurden Proband_innen ausgeschlossen, die aufgrund allgemeiner Beschwerden oder körperlicher Einschränkungen die Ausgangsposition für die Testung nicht einnehmen konnten, da in solchen Fällen eine standardisierte und sichere Durchführung der Untersuchung nicht gewährleistet gewesen wäre.

3.4 Stichprobengröße

Für diese Studie wurde, entsprechend der Empfehlung von Patijn (2019), eine Anzahl von 40 Proband_innen angestrebt, die letztlich auch erreicht werden konnte. Während der gesamten Studie gab es keinen Fall, bei dem die Testung vorzeitig abgebrochen werden musste.

3.4.1 Demografische Daten der Stichprobe

An der Studie nahmen insgesamt 42 Proband_innen teil. Davon waren 26 (61,90 %) Frauen und 16 (38,10 %) Männer. Die Proband_innen waren im Durchschnitt 44,6 Jahre

alt (SD = 11,8 Jahre), wobei das Alter von 30 bis 65 Jahren reichte. Der Median lag bei 39,5 Jahren. Damit war die Altersverteilung, mit einer größeren Anzahl jüngerer Proband_innen, leicht ungleich verteilt (siehe Tabelle 3). Da sich der Kieferbereich altersbedingt verändern kann, ist diese Information für die Einordnung der Ergebnisse wichtig. Ein möglicher Einfluss des Alters auf die Interrater-Reliabilität wurde in dieser Studie jedoch nicht untersucht.

Tabelle 3: Altersverteilung der teilnehmenden Proband_innen

Proband_innen	Minimum	Maximum	Median	Mittelwert	SD
(n = 42)	30	65	39,5	44,55	11,83

3.4.2 Daten zur vorliegenden Symptomatik

Anhand des am Studientag ausgefüllten Fragebogens konnten Daten zu den bestehenden Symptomatiken der Proband_innen erfasst werden. Die im Fragebogen berücksichtigten Symptome sind in *Kapitel 2.3.1* beschrieben. Basierend auf diesen Angaben wurde die Stichprobe in asymptomatische und symptomatische Proband_innen unterteilt. Dabei wurden Proband_innen, die ein Symptom angaben, zur symptomatischen Gruppe gezählt.

Wie in der nachfolgenden Abbildung dargestellt ist, nahmen insgesamt 42 Proband_innen an der Studie teil, davon waren 23 asymptomatisch (54,76 %) und 19 symptomatisch (45,24 %) (siehe Abbildung 1).

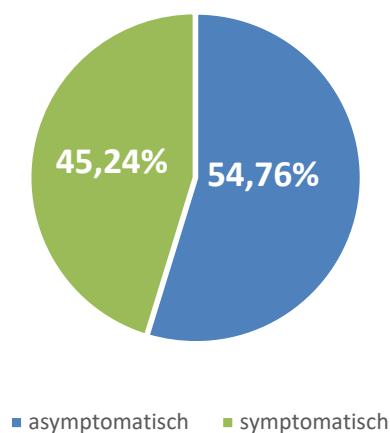


Abbildung 1: Anteil der asymptomatischen und symptomatischen Proband_innen

Die Auswertung ergab je nach Symptom unterschiedliche Verteilungen in der Stichprobe. Am häufigsten wurden hör- oder fühlbare Geräusche im Bereich des Kiefergelenks angegeben (11 Proband_innen). Kopfschmerzen, die einmal pro Woche oder häufiger auftraten, wurden wiederum von 6 Proband_innen angegeben. Verspannungen der Kaumuskulatur sowie Bewegungseinschränkungen oder auch Verspannungen im oberen Nackenbereich oder an der Schädelbasis wurden jeweils von 5 Proband_innen genannt. Zähneknirschen wurde von 3 und Schmerzen im Bereich des Kiefergelenks von 2 Proband_innen angegeben. Zwei weitere gaben Zahnschmerzen an. Dagegen gab keine_r der Proband_innen Schmerzen im Gesicht oder eine Trigeminusneuralgie an.

Die Häufigkeiten der genannten Symptome sind in Tabelle 4 zusammengefasst und geben einen Überblick über die Verteilung innerhalb der Stichprobe.

Tabelle 4: Übersicht der Symptome der Proband_innen

Symptom	Proband_innen
Schmerzen im Bereich des Kiefergelenks	2
Hör- bzw. fühlbare Geräusche im Kieferbereich	11
Zähneknirschen	3
Kopfschmerzen (einmal pro Woche oder häufiger)	6
Verspannung der Kaumuskulatur	5
Zahnschmerzen	2
Schmerzen im Gesicht bzw. Trigeminusneuralgie	0
Bewegungseinschränkung oder Verspannungen im oberen Nackenbereich oder an der Schädelbasis	5

3.5 Tester_innen

Für eine Interrater-Reliabilitätsstudie wird der Einsatz von zwei Tester_innen empfohlen (Patijn, 2019). Zu diesem Zweck wurden zwei Kolleginnen der Autorin für die Studie herangezogen. Beide üben seit Jahren die Tätigkeit als Physiotherapeutin aus und absolvieren zudem seit 2016 eine Osteopathie-Ausbildung. Zum Zeitpunkt der Studie

befanden sie sich zudem im Masterlehrgang für Osteopathie an der Universität für Weiterbildung Krems.

3.6 Messinstrumente und Testbeschreibung

In dieser Studie wurde der Induktionstest Os mandibulare als Messinstrument verwendet, um die passive Innen- und Außenrotation der Mandibula zu beurteilen. Die Testergebnisse wurden nach einem vorgegebenen Schema („unauffällig“, „auffällig rechts“, „auffällig links“) dokumentiert.

3.6.1 Beschreibung des Testverfahrens

In diesem Abschnitt erfolgt eine genaue Beschreibung des Induktionstests Os mandibulare, der in der vorliegenden Studie verwendet wurde.

Die Ausgangsstellung der Proband_innen war die Rückenlage. Die Testerin saß am Kopfende des_der Proband_in. Ihre Unterarme und Ellbogen waren auf der Liege abgelegt, und ihre Hände lagen rechts und links auf der Squama des Os temporale und auf den Wangen. Ihre Mittelfinger befanden sich rechts und links auf dem Corpus des Os mandibulare und die Zeigefinger waren am Kinn positioniert. Ihre Ring- und Kleinfinger lagen am kaudalen Rand der Mandibula auf. Dagegen hatten ihre Daumen keinen Kontakt mit der Mandibula.

Zur Testdurchführung leitete die Testerin in der kranialen Inspirationsphase mit sanftem Druck eine Außenrotation des Os mandibulare ein. Daraufhin wurde in der kranialen Expirationsphase eine Innenrotation des Os mandibulare eingeleitet. Auf diese Weise wurden die Außenrotation (Flexion) und die Innenrotation (Extension) des Os mandibulare getestet.

Der Test wurde pro Proband_in nur einmal durchgeführt, um einen möglichen mobilisierenden bzw. therapeutischen Effekt zu vermeiden. Demnach könnte ein mehrfaches passives Testen die Beweglichkeit beeinflussen und die Ergebnisse entsprechend verfälschen.

Wenn eine hinsichtlich Amplitude und Frequenz eingeschränkte, asymmetrische oder keine Bewegung festgestellt wurde, bestand der Verdacht auf eine Dysfunktion des Os mandibulare (Corts, 2020, S. 163).

Die Testergebnisse wurden von den Testerinnen in ein vorgefertigtes Bewertungsformular eingetragen:

- unauffällig
- auffällig rechts
- auffällig links

Die Handposition der Testerinnen während des Induktionstests Os mandibulare ist in Abbildung 2 dargestellt.



Abbildung 2: Handposition für den Induktionstest Os mandibulare

3.6.2 Materialien und Gerätschaften

Für die Durchführung des Induktionstests Os mandibulare wurden keine speziellen Messinstrumente oder auch technischen Geräte verwendet. Stattdessen erfolgte die Untersuchung ausschließlich manuell durch die beiden Testerinnen. In den zwei Untersuchungsräumen befanden sich identische Behandlungsliegen sowie Stühle, um eine einheitliche Position der Testerinnen während der Testdurchführung zu gewährleisten.

Der Induktionstest Os mandibulare findet in der osteopathischen Praxis häufig Anwendung, jedoch gibt es bislang keine Publikationen bezüglich seiner Gütekriterien (Reliabilität und Validität). Auch einen Goldstandard zur Beurteilung der mandibulären

Bewegung gibt es derzeit nicht. Infolgedessen basiert die Bewertung auf praktischer Erfahrung, einer standardisierten Schulung der Testerinnen sowie einem vorab festgelegten Testprotokoll. Die erhobenen Daten wurden zunächst in Microsoft Excel (Version 365, Microsoft Corporation, Redmond, USA) dokumentiert und für die weitere Analyse aufbereitet. Anschließend fand die statistische Auswertung mit IBM SPSS Statistics (Version 30.0.0.0 (172), IBM Corp., Armonk, USA) statt.

3.7 Studienablauf

In diesem Abschnitt wird der zeitliche und organisatorische Ablauf der Datenerhebung beschrieben, einschließlich der Testtage, der Zuteilung der Proband_innen sowie der Durchführung der diagnostischen Tests.

3.7.1 Trainingsphase

Nach Patijn (2019) wird eine Trainingsphase empfohlen, um die Zuverlässigkeit der Studie zu erhöhen. In dieser Phase soll eine Abstimmung der beiden Tester_innen in Bezug auf die folgenden Punkte erfolgen, und zwar auf die Ausgangsposition der Tester_innen, die Ausgangsposition der Proband_innen, anatomische Orientierungspunkte zur Erleichterung der Palpation, die Ausführung des Tests, die Anzahl der Wiederholungen sowie mögliche Interpretationsspielräume bei der Beurteilung des Testverfahrens.

Bereits vor dem geplanten Trainingstag fand ein erstes Treffen der beiden Testerinnen statt. Dort wurde der Induktionstest Os mandibulare aneinander geübt. Die gesamte Untersuchung, einschließlich Handhabung, Position, Druck, Ausführung und Testinterpretation, wurde dabei genau besprochen und auf einem Dokument festgehalten. Anschließend wurden die Testerinnen dazu angehalten, an so vielen Patient_innen wie möglich zu üben, um ihre Fähigkeiten diesbezüglich zu schulen.

Zwei Monate vor der Testung fand ein offizieller Trainingstag statt. An diesem führten die beiden Testerinnen den Test an zehn verschiedenen Proband_innen durch, für die keine speziellen Ein- oder Ausschlusskriterien galten. Die Ergebnisse dieses Tests wurden auf einem eigens für diese Untersuchung entwickelten Bewertungsbogen (siehe Anhang G) dokumentiert. Im Anschluss an die Untersuchung wurden die Proband_innen

um ein Feedback zur Handhaltung, zum Druck sowie zur Positionierung der Testerinnen befragt. Zudem wurden etwaige Fragen, Unklarheiten oder auch Abweichungen bei der Durchführung der Tests durch die Testerinnen geklärt. Nach einer Einigung wurde der genaue Testablauf schriftlich festgehalten. Die Testerinnen wurden angewiesen, sich beim Üben in ihrer Praxis genau an diesen standardisierten Testablauf zu halten, um bei der Ausführung des Testverfahrens eine möglichst hohe Sicherheit zu erlangen. Bei diesem Trainingstag wurde die von Patijn (2019) empfohlene Übereinstimmung von 80 % erreicht.

3.7.2 Studienphase

Die Testung fand am 14. und 15. Juni 2023 in Frauental an der Laßnitz statt. Insgesamt nahmen 42 Proband_innen an der Studie teil. Alle 20 Minuten wurden zwei Proband_innen eingeladen. Die Studienleitung empfing diese im Wartebereich, wo sie die unterschriebene Einverständniserklärung sowie die Fragebögen zu den Ausschlusskriterien und zu ihrer Symptomatik abgaben. Die Testerinnen selbst hatten keinen Einblick in diese Dokumente.

Daraufhin wurden die Proband_innen per Losverfahren einem der zwei Untersuchungsräume zugewiesen und gebeten, sich dort mit dem Rücken auf die Liege zu legen. Erst nachdem die Proband_innen diese Ausgangsstellung eingenommen hatten, betraten die Testerinnen den Raum. Vor dem Betreten erhielten die Testerinnen von der Studienleitung zunächst den Ergebnisbogen mit einer Nummer für den_die jeweilige_n Proband_in. Anschließend wurde der Induktionstest Os mandibulare bei dem_der Proband_in durchgeführt. Da es sich bei diesem Test um eine rein passive palpatorische Untersuchung handelte, waren in dieser Studie keine potenziellen Schädigungen für die Proband_innen zu erwarten.

Im Anschluss wurden die Testergebnisse auf einem Ergebnisbogen dokumentiert, ohne dass die Proband_innen darauf Einsicht nehmen konnten. Zudem wurde der Ergebnisbogen verdeckt in einer Box abgelegt, um sicherzustellen, dass die zweite Testerin nicht davon beeinflusst wurde. Danach wechselten die Testerinnen den Raum und führten den Test erneut durch. Nach der Untersuchung verließen die

Proband_innen die Untersuchungsräume und erhielten als Dank für ihren Zeitaufwand ein kleines Präsent in Form von Schokolade.

Nach Abschluss aller Testungen fand eine Nachbesprechung mit der Studienleitung und den beiden Testerinnen statt, in der etwaige Limitationen besprochen wurden. Diese Eindrücke werden in *Kapitel 5.4* erneut aufgegriffen.

3.7.3 Verblindung

Die Verblindung wurde gewährleistet, indem kein Austausch zwischen den Testerinnen stattfand. Somit hatten die Testerinnen keine Möglichkeit, die Ergebnisse der jeweils anderen Testerin einzusehen. Zudem wurde die Verblindung zwischen den Testerinnen und den Proband_innen dadurch sichergestellt, dass kein verbaler Austausch über Symptome oder auch Testergebnisse stattfand. Ebenfalls waren die Proband_innen den Testerinnen völlig unbekannt. Wie von Patijn (2019) beschrieben, ist eine solche Verblindung essenziell, um einen „Beobachter-Bias“ zu vermeiden und die Objektivität der Beurteilung sicherzustellen.

3.7.4 Randomisierung

Es wurden zwei Testtage festgelegt. Die Zuordnung der Proband_innen zu einem dieser Testtage erfolgte entsprechend ihren Wünschen. Die genaue Uhrzeit wurde von der Studienleitung festgelegt und den Proband_innen im Voraus mitgeteilt. Des Weiteren erfolgte die Zuteilung zu Raum 1 oder Raum 2 und damit auch zur jeweiligen Testerin nach dem Losverfahren, um eine unbeeinflusste Verteilung sicherzustellen. Wie von Patijn (2019) empfohlen wird, trägt eine solche zufallsbasierte Zuteilung dazu bei, systematische Verzerrungen zu vermeiden und vergleichbare Untersuchungsbedingungen zu schaffen.

3.8 Datenverarbeitung und -auswertung

Die erhobenen Daten wurden wie im *Kapitel 3.6.2* beschrieben anhand von Microsoft Excel sowie dem Datenverarbeitungsprogramm IBM SPSS Statistics aufgearbeitet.

3.8.1 Darstellung der Daten

Die Darstellung der erhobenen Daten erfolgt über eine Kreuztabelle, genauer gesagt über eine Vierfeldtafel, wie von Patijn (2019) empfohlen. Die Vierfeldtafel wird für

zahlreiche statistische Berechnungen herangezogen (vgl. Tabelle 5). Dabei wird die Verteilung der Häufigkeiten zweier dichotomer Merkmale in Abhängigkeit voneinander dargestellt.

Tabelle 5: Vierfeldertafel (Kreuztabelle) - Quelle: adaptiert nach Schäfer & Schöttker-Königer (2015, S. 137)

		Tester_in 1		Summe
		Ergebnis	0	
Tester_in 2	0	a	b	a+b
	1	c	d	c+d
Summe		a+c	b+d	n

Für Merkmale mit mehr als zwei Ausprägungen werden Mehrfeldertafeln (vgl. Tabelle 6) empfohlen (Schäfer & Schöttker-Königer, 2015, S. 136).

Tabelle 6: Mehrfeldertafel – Quelle: adaptiert nach Schäfer und Schöttker-Königer (2015, S. 136)

		Tester_in 1			Summe
		Ergebnis	0	1	
Tester_in 2	0	a	b	c	a+b+c
	1	d	e	f	d+e+f
	2	g	h	i	g+h+i
Summe		a+d+g	b+e+h	c+f+i	n

3.8.2 Berechnung und Interpretation des Cohens-Kappa-Koeffizienten

Zur Bewertung der Interrater-Reliabilität wurde der Cohens-Kappa-Koeffizient herangezogen. Um diesen ermitteln zu können, müssen zuvor die folgenden Variablen berechnet werden (Patijn, 2019):

- p_0 (Gesamtübereinstimmung) = prozentuale Übereinstimmung
- p_c = erwartete Zufallsübereinstimmung

Zur Berechnung des Cohens-Kappa-Koeffizienten wird die folgende Formel verwendet:

$$k = \frac{p^0 - pc}{1 - pc}$$

(Patijn, 2019).

Die Berechnung des Cohens-Kappa-Koeffizienten erfolgte unter der Voraussetzung, dass beide Testerinnen dieselben nominalskalierten Kategorien („unauffällig“, „auffällig rechts“, „auffällig links“) verwendeten. Jede Bewertung wurde eindeutig einer Kategorie zugeordnet.

3.8.3 Übersicht über die unterschiedlichen berechneten Cohens-Kappa-Koeffizienten

Basierend auf den erhobenen Daten werden unterschiedliche Cohens-Kappa-Koeffizienten berechnet, um die Interrater-Reliabilität differenziert zu analysieren:

- **Gesamtanalyse (3 Kategorien):**

Der primäre Cohens-Kappa-Koeffizient wird unter Einbezug aller Proband_innen berechnet. Die Bewertung erfolgt anhand von drei Ergebniskategorien:

„unauffällig“, „auffällig rechts“ und „auffällig links“.

- **Auswertung nach Geschlecht (3 Kategorien):**

Die Daten der weiblichen und männlichen Proband_innen werden getrennt voneinander analysiert, um mögliche geschlechtsspezifische Unterschiede in der Beurteilung sichtbar zu machen.

- **Dichotome Analyse (2 Kategorien):**

Zusätzlich wird eine vereinfachte Auswertung mit zwei Ergebniskategorien durchgeführt („unauffällig“ vs. „auffällig“), um eine dichotome Interpretation zu ermöglichen.

- **Auswertung nach Geschlecht (2 Kategorien):**

Auch die dichotomen Daten werden geschlechtsspezifisch ausgewertet, um etwaige Unterschiede bei der Beurteilung zu identifizieren.

- **Auswertung nach Symptomatik/Asymptomatik (2 Kategorien):**

Es erfolgt eine getrennte Betrachtung symptomatischer sowie asymptomatischer Proband_innen, um potenzielle Einflüsse bestehender Beschwerden auf die Bewertung zu identifizieren.

Des Weiteren wird für jeden Cohens-Kappa-Koeffizienten das dazugehörige 95 %-Konfidenzintervall berechnet. Diese Koeffizienten können Werte zwischen -1 und $+1$ annehmen (Patijn, 2019). Die Interpretation erfolgt anhand der von Landis und Koch (1977, S. 165) beschriebenen Klassifikation (vgl. Tabelle 7):

Tabelle 7: Klassifikation der Cohens-Kappa-Koeffizienten – Quelle: adaptiert nach Landis und Koch (1977, S. 165)

Cohens-Kappa-Koeffizient κ	Grad der Übereinstimmung
$<0,00$	poor (mangelhaft)
$0,00 - 0,20$	slight (gering)
$0,21 - 0,40$	fair (mäßig)
$0,41 - 0,60$	moderate (mittelmäßig)
$0,61 - 0,80$	substantial (beachtlich)
$0,81 - 1,00$	almost perfect (fast vollkommen)

3.8.4 Testung der Hypothesen

Zur Überprüfung der in der Einleitung aufgestellten Hypothesen wird der Cohens-Kappa-Koeffizient (κ) für die gesamte Stichprobe ebenso wie für die verschiedenen Subgruppen berechnet. Dieser Kennwert misst die Übereinstimmung der beiden Testerinnen bei der Durchführung des Induktionstests Os mandibulare.

Die Testung findet anhand der folgenden Hypothesen statt:

- **Nullhypothese (H_0):** $\kappa < 0,61$ (keine substantielle Übereinstimmung)
- **Alternativhypothese (H_1):** $\kappa \geq 0,61$ (substantielle oder höhere Übereinstimmung)

Zusätzlich zum Cohens-Kappa-Koeffizienten wird für jede Auswertung das 95 %-Konfidenzintervall berechnet. Das Signifikanzniveau wird mit $\alpha = 0,05$ festgelegt.

4 Ergebnisse

In diesem Kapitel werden die Ergebnisse der Interrater-Reliabilitätsstudie dargestellt. Hierbei wird jeder der berechneten Cohens-Kappa-Koeffizienten in einem eigenen Unterkapitel einzeln erläutert.

Die Stichprobe der Studie wurde bereits im Kapitel 3.1 beschrieben. Aus Gründen der Übersichtlichkeit in Bezug auf die Studienergebnisse wird die Zusammensetzung der Stichprobe nachfolgend nochmals zusammengefasst: Insgesamt nahmen 42 Proband_innen ($n = 42$) an der Studie teil. Davon waren 26 weiblich (61,90 %) und 16 männlich (38,10 %). Die Einteilung nach Symptomatik ergab, dass 23 Proband_innen (54,76 %) der asymptomatischen Gruppe und 19 Proband_innen (45,24 %) der symptomatischen Gruppe zugeordnet wurden. Darüber hinaus lag das Durchschnittsalter der Stichprobe bei 44,6 Jahren bei einer Standardabweichung (SD) von 11,8 Jahren.

4.1 Cohens-Kappa-Koeffizient mit drei Ergebniskategorien der gesamten Stichprobe

In der untenstehenden Tabelle 8 wird ein Überblick über die beobachteten Häufigkeiten der Testergebnisse aller Proband_innen gegeben. Die Kategorie 0 beschreibt ein unauffälliges Ergebnis, die Kategorie 1 eine Auffälligkeit rechts und die Kategorie 2 eine Auffälligkeit links. Die grün markierten Felder zeigen die Übereinstimmungen der beiden Testerinnen an.

Absolut gesehen kam es zu 24 Übereinstimmungen. Diese teilen sich den Kategorien entsprechend wie folgt auf. In Kategorie 0 zeigten sich 14 Übereinstimmungen, in Kategorie 1 dagegen 2 Übereinstimmungen und in Kategorie 2 wiederum 8 Übereinstimmungen. Daraus ergibt sich eine prozentuelle Übereinstimmung von $p_0 = 57,14$ %. Die Zufallsübereinstimmung liegt bei $p_c = 36,22$ % und der p-Wert beträgt 0,002. Anhand dieser Daten konnte ein Cohens-Kappa-Wert von $\kappa = 0,33$ errechnet werden. Das dazugehörige *95 %-Konfidenzintervall* erstreckt sich von 0,123 bis 0,532.

Tabelle 8: Kreuztabelle – drei Ergebniskategorien der gesamten Stichprobe

		Testerin 1			Summe
		Kategorie	0	1	
Testerin 2	0	14	4	7	25
	1	2	2	0	4
	2	0	5	8	13
	Summe	16	11	15	42

4.2 Cohens-Kappa-Koeffizient mit drei Ergebniskategorien der weiblichen Stichprobe

Die unten abgebildete Tabelle 9 zeigt einen Überblick über die Ergebnisse der Testung der Probandinnen ($n = 26$). Absolut gesehen zeigten sich bei dieser Auswertung 16 Übereinstimmungen. Die Kategorie 0 weist 8 Übereinstimmungen, die Kategorie 1 dagegen 2 Übereinstimmungen und die Kategorie 2 wiederum 6 Übereinstimmungen auf. Dies entspricht bei dieser Stichprobe ($n = 26$) einer Gesamtübereinstimmung von $p_0 = 61,54\%$. Die Zufallsübereinstimmung liegt bei $p_c = 36,69\%$ und der p-Wert liegt bei 0,005. Der Cohens-Kappa-Koeffizient ergibt somit $\kappa = 0,39$ und das 95%-Konfidenzintervall reicht von 0,106 bis 0,665.

Tabelle 9: Kreuztabelle – drei Ergebniskategorien der weiblichen Stichprobe

		Testerin 1			Summe
		Kategorie	0	1	
Testerin 2	0	8	1	5	14
	1	2	2	0	4
	2	0	2	6	8
	Summe	10	5	11	26

4.3 Cohens-Kappa-Koeffizient mit drei Ergebniskategorien der männlichen Stichprobe

Nachdem die weibliche Stichprobe im vorangehenden Abschnitt 4.2 dargestellt wurde, zeigt die unten angeführte Tabelle 10 die Ergebnisse der männlichen Probanden.

Bei der männlichen Stichprobe ($n = 16$) stimmten die Testerinnen in 8 Fällen überein, die sich auf die drei Ergebniskategorien folgendermaßen aufteilen. In Kategorie 0 zeigten sich 6 Übereinstimmungen, in Kategorie 1 gab es 0 Übereinstimmungen und in Kategorie 2 kam es zu 2 Übereinstimmungen. Prozentual ergibt sich daraus eine Gesamtübereinstimmung von $p_0 = 50\%$, während die zufällige Übereinstimmung bei $p_c = 33,59\%$. Der p-Wert beträgt $p = 0,070$. Demzufolge beträgt der Cohens-Kappa-Koeffizient $\kappa = 0,25$ und das *95 %-Konfidenzintervall* erstreckt sich von 0,018 bis 0,521.

Tabelle 10: Kreuztabelle – drei Ergebniskategorien der männlichen Stichprobe

		Testerin 1			Summe
		Kategorie 0	Kategorie 1	Kategorie 2	
Testerin 2	Kategorie 0	6	3	2	11
	Kategorie 1	0	0	0	0
	Kategorie 2	0	3	2	5
	Summe	6	6	4	16

4.4 Cohens-Kappa-Koeffizient mit zwei Ergebniskategorien der gesamten Stichprobe

In der nachfolgenden Tabelle 11 sind die beobachteten Häufigkeiten der Testbeurteilung durch die beiden Testerinnen bei der Betrachtung mit zwei Ergebniskategorien (unauffällig oder auffällig) dargestellt. Hierbei wurden die Seiten (rechts/links) nicht berücksichtigt. Dies erfolgte, da in dieser Analyse nur das Vorliegen einer Auffälligkeit unabhängig von der Seite relevant war.

Die Kategorie 0 beschreibt ein unauffälliges Testergebnis, während Kategorie 1 ein auffälliges beschreibt. Die grün markierten Felder in Tabelle 12 zeigen die Fälle, in denen beide Testerinnen übereinstimmende Bewertungen abgegeben haben.

Insgesamt konnten 29 Übereinstimmungen festgestellt werden, wovon Kategorie 0 (unauffällig) 14 Übereinstimmungen und Kategorie 1 (auffällig) 15 Übereinstimmungen zeigt. Die berechnete Gesamtübereinstimmung (p_0) beträgt 69 %, die erwartete Übereinstimmung durch Zufall (p_c) lag bei 47,7 %. Der p-Wert liegt bei $p = 0,004$. Daraus

ergab sich ein Cohens-Kappa-Koeffizient von $\kappa = 0,41$ und das *95 %-Konfidenzintervall* liegt zwischen 0,153 bis 0,660.

Tabelle 11: Kreuztabelle – zwei Ergebniskategorien der gesamten Stichprobe

		Testerin 1		Summe
		Kategorie	0	
Testerin 2	0	14	11	25
	1	2	15	17
Summe		16	26	42

4.5 Cohens-Kappa-Koeffizient mit zwei Ergebniskategorien der weiblichen Stichprobe

Die untenstehende Tabelle 12 gibt einen Überblick über die Testergebnisse der weiblichen Stichprobe ($n = 26$). Wie bereits in Kapitel 4.4 erläutert wurde, steht Kategorie 0 für ein unauffälliges Testergebnis, während die Kategorie 1 ein auffälliges Testergebnis markiert. Die grün markierten Felder in der Tabelle 13 zeigen die Übereinstimmungen zwischen den beiden Testerinnen.

Insgesamt gab es 18 Übereinstimmungen. Die Gesamtübereinstimmung (p_0) lag bei 69,2 % und die erwartete Zufallsübereinstimmung (p_c) bei 49,1 %. Der errechnete p-Wert beträgt $p = 0,034$, woraus sich ein Cohens-Kappa-Koeffizient von $\kappa = 0,40$ ergibt. Das zugehörige *95 %-Konfidenzintervall* reicht von 0,053 bis 0,698.

Tabelle 12: Kreuztabelle – zwei Ergebniskategorien der weiblichen Stichprobe

		Testerin 1		Summe
		Kategorie	0	
Testerin 2	0	8	6	14
	1	2	10	12
Summe		10	16	26

4.6 Cohens-Kappa-Koeffizient mit zwei Ergebniskategorien der männlichen Stichprobe

In Tabelle 13 sind die beobachteten Häufigkeiten der Testbeurteilung für die männlichen Probanden aufgelistet. Bei dieser Darstellung wurden 11 Übereinstimmungen der Testerinnen festgestellt. Die Kategorie 0 weist hierbei 6 Übereinstimmungen auf, und die Kategorie 1 zählt 5 Übereinstimmungen. Prozentuell gesehen ergibt dies eine Übereinstimmung von $p_0 = 68,8\%$. Die zufällige Übereinstimmung p_c liegt bei $45,3\%$. Der p -Wert beträgt $0,037$. Somit ergibt der aus den vorliegenden Daten errechnete Cohens-Kappa-Koeffizient $\kappa = 0,43$ und das *95 %-Konfidenzintervall* erstreckt sich von $0,122$ bis $0,754$.

Tabelle 13: Kreuztabelle – zwei Ergebniskategorien der männlichen Stichprobe

		Testerin 1		Summe
		Kategorie 0	Kategorie 1	
Testerin 2	0	6	5	11
	1	0	5	5
Summe		6	10	16

4.7 Cohens-Kappa-Koeffizient mit zwei Ergebniskategorien der asymptomatischen Stichprobe

In der nachstehenden Tabelle 14 werden die Ergebnisse für die asymptomatischen Proband_innen dargestellt. Es gab 15 Übereinstimmungen der Testerinnen, davon 8 in Kategorie 0 und 7 in Kategorie 1. Prozentuell ergibt dies eine Gesamtübereinstimmung von $p_0 = 65,2\%$. Die erwartete Zufallsübereinstimmung (p_c) liegt bei $46,7\%$. Der p -Wert beträgt $p = 0,056$. Folglich ergibt der aus den vorliegenden Daten errechnete Cohens-Kappa-Koeffizient $\kappa = 0,35$ und das *95 %-Konfidenzintervall* erstreckt sich von $0,028$ bis $0,662$.

Tabelle 14: Kreuztabelle – zwei Ergebniskategorien der asymptomatischen Stichprobe

		Testerin 1		Summe
		Kategorie	0	
Testerin 2	0	8	7	15
	1	1	7	8
Summe		9	14	23

4.8 Cohens-Kappa-Koeffizient mit zwei Ergebniskategorien der symptomatischen Stichprobe

In Kapitel 3.4.2 Daten zur vorliegenden Symptomatik wurde aufgezeigt, dass 19 der Proband_innen Symptomatiken angegeben haben. In der folgenden Tabelle 15 sind die Ergebnisse der als symptomatisch bezeichneten Proband_innen dargestellt. Bei diesen stimmten die Testerinnen absolut gesehen 14 Mal überein. Die Kategorie 0 weist 6 Übereinstimmungen und die Kategorie 1 wiederum 8 Übereinstimmungen auf. Prozentuell gesehen ergibt sich daraus eine Gesamtübereinstimmung von $p_0 = 73,68\%$. Die zufällige Übereinstimmung beträgt $p_c = 49,31\%$. Der p-Wert liegt bei 0,027. Daraus ergibt sich ein Cohens-Kappa-Koeffizient von $\kappa = 0,48$ mit dem dazugehörigen *95 %-Konfidenzintervall* von 0,076 bis 0,791.

Tabelle 15: Kreuztabelle – zwei Ergebniskategorien der symptomatischen Stichprobe

		Testerin 1		Summe
		Kategorie	0	
Testerin 2	0	6	4	10
	1	1	8	9
Summe		7	12	19

4.9 Übersicht aller errechneten Cohens-Kappa-Koeffizienten

Zur besseren Übersicht werden in der Tabelle 16 alle sieben errechneten Cohens-Kappa-Koeffizienten mit dem dazugehörigen *95 %-Konfidenzintervall*

zusammengefasst dargestellt. Die berechneten Koeffizienten liegen zwischen $\kappa = 0,25$ und $\kappa = 0,48$.

Der niedrigste Cohens-Kappa-Wert wurde bei der männlichen Stichprobe mit drei Ergebniskategorien erreicht, während der höchste Wert bei der Auswertung der symptomatischen Stichprobe in zwei Ergebniskategorien erreicht wurde.

Tabelle 16: Übersicht über alle errechneten Cohens-Kappa-Koeffizienten mit zugehörigem 95%-Konfidenzintervall

Stichprobe (n) – Anzahl der Kategorien	κ	95%-KI	
		Untergrenze	Obergrenze
Gesamte Stichprobe (n = 42) – 3 Kategorien	0,33	0,123	0,532
Weibliche Stichprobe (n = 26) – 3 Kategorien	0,39	0,106	0,659
Männliche Stichprobe (n = 16) – 3 Kategorien	0,25	0,020	0,524
Gesamte Stichprobe (n = 42) – 2 Kategorien	0,41	0,153	0,660
Weibliche Stichprobe (n = 26) – 2 Kategorien	0,40	0,053	0,698
Männliche Stichprobe (n = 16) – 2 Kategorien	0,43	0,122	0,754
Stichprobe mit asymptomatischen Proband_innen (n = 23) – 2 Kategorien	0,35	0,020	0,662
Stichprobe mit symptomatischen Proband_innen (n = 19) – 2 Kategorien	0,48	0,076	0,791

4.10 Häufigkeitsverteilung der Stichprobe

Die Häufigkeitsverteilungen der Ergebnisse der beiden Testerinnen werden nach den Kategorien getrennt dargestellt. Die nachstehende Tabelle 17 bildet die entsprechenden Werte ab. Dabei stehen die Zahlen 0, 1 und 2 für die Ergebniskategorien des Induktionstests Os mandibulare (0 = unauffällig, 1 = auffällig rechts, 2 = auffällig links). Die folgende Übersicht zeigt, wie häufig jede dieser Kategorien vergeben wurde.

Tabelle 17: Verteilung der Bewertungen durch Testerin 1 und Testerin 2

Kategorie	Testerin 1		Testerin 2	
	<i>absolut</i>	<i>prozentuell</i>	<i>absolut</i>	<i>prozentuell</i>
0 (unauffällig)	16	38,1 %	25	59,5 %
1 (auffällig rechts)	11	26,2 %	4	9,5 %
2 (auffällig links)	15	35,7 %	13	31,0 %

Testerin 1 bewertete 16 von insgesamt 42 Fällen (38,1 %) als unauffällig, 11 Fälle (26,2 %) als auffällig rechts und 15 Fälle (35,7 %) als auffällig links. Bei Testerin 2 wurden dagegen 25 Fälle (59,5 %) als unauffällig eingestuft, 4 Fälle (9,5 %) als auffällig rechts und 13 Fälle (31,0 %) als auffällig links. Von beiden Testerinnen wurde am häufigsten die Kategorie 0 (unauffällig) vergeben. Am seltensten wurde wiederum Kategorie 1 (auffällig rechts) vergeben, insbesondere von Testerin 2.

Wird das Ergebnis in zwei Ergebniskategorien betrachtet, nämlich unauffällig (0) sowie auffällig (1 und 2 zusammengefasst), dann hat Testerin 1 insgesamt 16 Fälle (38,1 %) als unauffällig und 26 Fälle (61,9 %) als auffällig bewertet. Im Gegensatz dazu bewertete Testerin 2 insgesamt 25 Fälle (59,5 %) als unauffällig und 17 Fälle (40,5 %) als auffällig.

In der folgenden Tabelle 18 werden nun die auffälligen Testergebnisse der Kategorien 1 und 2 zusammengefasst dargestellt. Hierbei zeigt sich, dass sich die Verteilung der auffälligen Bewertungen zwischen den beiden Testerinnen sowie zwischen den Geschlechtern der Proband_innen unterscheidet. Während Testerin 1 bei Männern und Frauen nahezu gleich häufig Auffälligkeiten gefunden hat, bewertete Testerin 2 männliche Probanden (31,3 %) im Vergleich zu weiblichen Probanden (46,2 %) seltener als auffällig.

Tabelle 18: Geschlechtsverteilung der auffälligen Testergebnisse

Geschlecht	Testerin 1		Testerin 2	
	<i>absolut</i>	<i>prozentuell</i>	<i>absolut</i>	<i>prozentuell</i>
weiblich	16	61,5 %	12	46,2 %
männlich	10	62,5 %	5	31,3 %

4.11 Überprüfung der Hypothese

In der Einleitung wurden eine Nullhypothese (H_0) und eine Alternativhypothese (H_1) formuliert. Diese können aufgrund der erhobenen Daten überprüft werden.

Nullhypothese (H_0): Die Übereinstimmung der Ergebnisse der beiden Testerinnen ist kleiner als der Cohens-Kappa-Koeffizient von 0,61 ($\kappa < 0,61$) und liegt somit unter dem beachtlichen Bereich (Landis & Koch, 1977, S. 165).

Alternativhypothese (H_1): Die Übereinstimmung der Ergebnisse der beiden Testerinnen ist gleich oder größer als der Cohens-Kappa-Koeffizient von 0,61 ($\kappa \geq 0,61$) und liegt somit im oder über dem beachtlichen Bereich (Grouven et al., 2007; Landis & Koch, 1977, S. 165).

Da keiner der berechneten Cohens-Kappa-Werte größer oder gleich 0,61 ist, wird die Alternativhypothese (H_1) verworfen. Stattdessen wird die Nullhypothese (H_0) angenommen.

5 Diskussion

Das vorliegende Kapitel wird für eine klare Darstellung in verschiedenen Unterkapitel gegliedert. Dabei werden die Ergebnisse näher betrachtet, mit anderen Studien verglichen, ihre Bedeutung für die Osteopathie wird beschrieben und es werden mögliche Schwächen der Arbeit besprochen.

5.1 Interpretation der Ergebnisse

Im Folgenden sollen die in Kapitel 4 beschriebenen Ergebnisse interpretiert und hinterfragt werden. Zum Zweck dieser Analyse werden die Klassifikationen von Landis und Koch (1977, S. 165) herangezogen.

5.1.1 Interpretation des Cohens-Kappa-Koeffizienten mit drei Ergebniskategorien der gesamten Stichprobe

Der Cohens-Kappa-Koeffizient für alle Proband_innen beträgt $\kappa = 0,33$. Anhand dieses Wertes kann die Übereinstimmung der beiden Testerinnen als „mäßig“ eingestuft werden (Landis & Koch 1977, S. 165). Das dazugehörige *95 %-Konfidenzintervall (KI)* von 0,123 bis 0,532 zeigt, dass der wahre Cohens-Kappa-Wert mit *95 %iger* Wahrscheinlichkeit innerhalb dieses Bereichs liegt und somit zwischen einer „geringen“ und einer „mittelmäßigen“ Übereinstimmung variiert.

Obwohl die beobachtete Übereinstimmung deutlich über dem Zufall liegt, zeigt das Ergebnis nur eine begrenzte Einigkeit in der Bewertung der beiden Testerinnen auf. Patijn (2019) weist darauf hin, dass die Ergebnisse manueller Testverfahren von individuellen Faktoren wie Erfahrung, Wahrnehmung und subjektiver Interpretation beeinflusst werden können. Diese Einflüsse können zu unterschiedlichen Testergebnissen führen. Ein Grund für den „mäßigen“ Wert könnte auch die Einteilung in drei Ergebniskategorien sein. Sim & Wright (2005) weisen darauf hin, dass bei mehreren Kategorien die Wahrscheinlichkeit unterschiedlicher Bewertungen steigt, was den Cohens-Kappa-Wert verringern kann.

Auch die Verteilung der Bewertungen spielt eine Rolle. In dieser Studie wurden viele Testergebnisse als „unauffällig“ beurteilt, wodurch die erwartete Zufallsübereinstimmung ansteigt. Dadurch erscheint der Cohens-Kappa-Wert niedriger,

als es die tatsächliche Übereinstimmung vermuten lässt. Dieses Phänomen wird als Prävalenzeffekt oder „Kappa-Paradoxon“ bezeichnet (Donker et al., 1993; Sim & Wright, 2005).

5.1.2 Interpretation des Cohens-Kappa-Koeffizienten mit drei Ergebniskategorien der weiblichen Stichprobe

Der Cohens-Kappa-Wert der weiblichen Stichprobe ($n = 26$) mit drei Ergebniskategorien beträgt $\kappa = 0,39$ und weist damit auf eine „mäßige“ Übereinstimmung zwischen den beiden Testerinnen hin. Im Vergleich zur gesamten Stichprobe ($\kappa = 0,33$) fällt der Wert etwas höher aus, was auf eine größere Übereinstimmung innerhalb der weiblichen Gruppe hindeutet.

Die statistische Signifikanz ($p = 0,005$) zeigt, dass die beobachtete Übereinstimmung nicht zufällig entstanden ist. Eine mögliche Erklärung für den leicht höheren Cohens-Kappa-Wert könnte darin liegen, dass die weibliche Gruppe homogener zusammengesetzt war oder die Ergebnisse eindeutiger ausfielen, was die Beurteilung erleichtert (Sim & Wright, 2005). Insgesamt bleibt die Übereinstimmung jedoch im „mäßigen“ Bereich, was in der Literatur für manuelle Testverfahren als typisch beschrieben wird (Basile et al., 2017; Patijn, 2019).

5.1.3 Interpretation des Cohens-Kappa-Koeffizienten mit drei Ergebniskategorien der männlichen Stichprobe

Der Cohens-Kappa-Koeffizient der männlichen Stichprobe ($n = 16$) mit drei Ergebniskategorien ergibt einen Wert von $\kappa = 0,25$ und kann damit der Kategorie „mäßig“ zugeordnet werden (Landis & Koch 1977, S. 165). Der p-Wert von 0,070 liegt knapp über dem üblichen Signifikanzniveau von 0,05. Daher gilt die beobachtete Übereinstimmung statistisch nicht als signifikant und könnte auch zufällig zustande gekommen sein.

Im Vergleich zu den anderen Gruppen stellt dieser Wert die niedrigste Übereinstimmung in der gesamten Studie dar. Dies deutet darauf hin, dass die Ergebnisse innerhalb der männlichen Gruppe weniger übereinstimmend ausfielen. Ein möglicher Grund dafür ist die kleine Stichprobengröße, da sich bei weniger Proband_innen einzelne

Abweichungen stärker auf den Cohens-Kappa-Wert auswirken und die Reliabilität verringern können (Sim & Wright, 2005).

Hinzu kommt, dass in dieser Gruppe die Kategorie „auffällig rechts“ von Testerin 2 nicht vergeben wurde. Solche unausgeglichene Kategorienverteilungen können den Cohens-Kappa-Wert zusätzlich mindern, da sie die erwartete Zufallsübereinstimmung erhöhen (Donker et al., 1993; Mayer et al., 2004). Das breite 95 %-Konfidenzintervall (0,018–0,521) zeigt zudem auf, dass die tatsächliche Übereinstimmung sehr unterschiedlich ausfallen kann. Für zukünftige Untersuchungen wäre eine größere männliche Stichprobe sinnvoll, um zuverlässiger und aussagekräftigere Ergebnisse zu erzielen.

5.1.4 Interpretation des Cohens-Kappa-Koeffizienten mit zwei Ergebniskategorien der gesamten Stichprobe

Im Vergleich zur Auswertung mit drei Ergebniskategorien ($\kappa = 0,33$) fällt die Übereinstimmung bei zwei Ergebniskategorien ($\kappa = 0,41$) etwas höher aus und kann somit der Kategorie „mittelmäßig“ zugeordnet werden (Landis & Koch 1977, S. 165).

Dies lässt sich dadurch erklären, dass bei einer dichotomen Bewertung mit den Kategorien „unauffällig“ und „auffällig“ der Interpretationsspielraum geringer ist und dadurch weniger Unterschiede in der Beurteilung auftreten. Eine geringere Anzahl an Kategorien führt häufig zu einer höheren Reliabilität, da die Entscheidung einfacher und eindeutiger ausfällt (Sim & Wright, 2005). Durch die Vereinfachung verliert die Auswertung an Detailgenauigkeit, weil keine Seitenlokalisierung berücksichtigt wird. Dies wird auch in der Literatur beschrieben. Eine reduzierte Anzahl an Bewertungskategorien kann die Übereinstimmung verbessern, verringert jedoch gleichzeitig die Aussagekraft der Testergebnisse (Mayer et al., 2004). Insgesamt deutet der höhere Cohens-Kappa-Wert darauf hin, dass die Beurteilung einheitlicher ausfällt, wenn nur zwischen zwei Bewertungskategorien unterschieden wird.

5.1.5 Interpretation des Cohens-Kappa-Koeffizienten mit zwei Ergebniskategorien der weiblichen Stichprobe

Der Cohens-Kappa-Wert der weiblichen Stichprobe ($n = 26$) mit zwei Ergebniskategorien beträgt $\kappa = 0,40$. Dieses Ergebnis zeigt, ähnlich wie in der Gesamtauswertung, eine „mäßige“ Übereinstimmung zwischen den Testerinnen. Im Vergleich zur Analyse mit drei Ergebniskategorien ($\kappa = 0,39$) ändert sich der Wert kaum. Das spricht dafür, dass die Vereinfachung auf zwei Kategorien die Beurteilung zwar etwas stabiler, aber nicht genauer gemacht hat. Ein möglicher Grund dafür ist, dass die Testerinnen in dieser Gruppe schon bei der dreistufigen Einteilung ähnlich entschieden haben. Die Reduktion auf zwei Kategorien verringert mögliche Unterschiede, ändert aber nichts an der grundsätzlichen Einschätzung (Sim & Wright, 2005).

Wie auch in anderen Studien beschrieben wird, liegen Reliabilitätswerte manueller Tests meist im mäßigen Bereich. Das hängt vor allem damit zusammen, dass palpatorische Verfahren stark von der Erfahrung und Wahrnehmung der Tester_innen abhängen (Basile et al., 2017; Patijn, 2019).

Insgesamt zeigt sich, dass die Testerinnen bei den weiblichen Probandinnen ähnlich bewertet haben, egal ob mit drei oder zwei Kategorien. Das deutet darauf hin, dass der Induktionstest Os mandibulare in dieser Gruppe relativ zuverlässig angewendet werden kann, wenn die Testbedingungen klar festgelegt sind.

5.1.6 Interpretation des Cohens-Kappa-Koeffizienten mit zwei Ergebniskategorien der männlichen Stichprobe

Der Cohens-Kappa-Wert der männlichen Stichprobe ($n = 16$) mit zwei Ergebniskategorien beträgt $\kappa = 0,43$ und zeigt damit eine „mittelmäßige“ Übereinstimmung zwischen den Testerinnen (Landis & Koch 1977, S. 165). Im Vergleich zur dreistufigen Auswertung ($\kappa = 0,25$) fällt die Übereinstimmung deutlich höher aus. Dieser Unterschied lässt sich dadurch erklären, dass bei einer dichotomen Bewertung („unauffällig“ oder „auffällig“) der Interpretationsspielraum geringer ist und sich die Testerinnen dadurch häufiger einig waren (Sim & Wright, 2005). Die höhere Übereinstimmung spricht dafür, dass eine dichotome Bewertung zu tendenziell

besseren Ergebnissen führen kann. Allerdings sollte berücksichtigt werden, dass die männliche Stichprobe sehr klein war. Bei wenigen Proband_innen wirken sich einzelne abweichende Einschätzungen stärker auf den Cohens-Kappa-Wert aus, was die Reliabilität empfindlicher macht (Mayer et al., 2004). Eine größere und ausgewogenere Stichprobe könnte daher zu verlässlicheren Ergebnissen führen.

5.1.7 Interpretation des Cohens-Kappa-Koeffizienten mit zwei Ergebniskategorien der asymptomatischen Stichprobe

Der Cohens-Kappa-Wert der asymptomatischen Stichprobe ($n = 23$) mit zwei Ergebniskategorien liegt mit $\kappa = 0,35$ im Bereich einer „mäßigen“ Übereinstimmung laut der Klassifikation nach Landis und Koch (1977, S. 165). Im Vergleich zu den anderen Gruppen fällt die Reliabilität etwas geringer aus. Ein möglicher Grund für die geringere Übereinstimmung könnte sein, dass bei asymptomatischen Proband_innen oft nur geringe oder keine tastbaren Auffälligkeiten vorhanden sind. Dadurch wird die Befundung schwieriger, und es kommt häufiger zu unterschiedlichen Einschätzungen der Tester_innen. Dies zeigt auch die systematische Übersichtsarbeit von Basile et al. (2017), in der aufgezeigt wird, dass osteopathische Tests bei asymptomatischen Proband_innen meist nur eine geringe bis mäßige Übereinstimmung mit Cohens-Kappa-Werten zwischen 0,20 und 0,50 aufweisen.

In dieser Gruppe wurde die Kategorie „unauffällig“ besonders häufig vergeben. Dadurch entstand eine ungleiche Verteilung der Bewertungen, die den Cohens-Kappa-Wert beeinflussen kann. Wenn eine Kategorie sehr oft gewählt wird, steigt die Zufallsübereinstimmung, wodurch der berechnete Cohens-Kappa-Wert niedriger ausfällt. Dieses Phänomen wird als Prävalenzeffekt bezeichnet (Donker et al., 1993; Sim & Wright, 2005).

Das Ergebnis mit einem p-Wert von 0,056 ist statistisch nicht signifikant, liegt jedoch knapp über dem Signifikanzniveau. Vermutlich hängt dies mit der homogenen Zusammensetzung der asymptomatischen Gruppe und damit zusammen, dass die meisten Testergebnisse „unauffällig“ waren.

5.1.8 Interpretation des Cohens-Kappa-Koeffizienten mit zwei Ergebniskategorien der symptomatischen Stichprobe

Der Cohens-Kappa-Wert der symptomatischen Stichprobe ($n = 19$) liegt mit $\kappa = 0,48$ im Bereich einer „mittelmäßigen“ Übereinstimmung laut der Klassifikation nach Landis und Koch (1977, S. 165). Damit zeigt diese Gruppe den höchsten Wert im Vergleich zu den anderen Stichproben. Ein möglicher Grund dafür ist, dass bei symptomatischen Proband_innen palpatorische Auffälligkeiten deutlicher ausgeprägt sind, was die Beurteilung erleichtert und die Übereinstimmung erhöht (Basile et al., 2017; Patijn, 2019). Auch die dichotome Bewertung („auffällig“/„unauffällig“) kann zu einer besseren Übereinstimmung beigetragen haben, da sie den Interpretationsspielraum verringert (Sim & Wright, 2005).

Insgesamt zeigen die Ergebnisse, dass der Induktionstest Os mandibulare bei symptomatischen Proband_innen eine höhere Reproduzierbarkeit aufweist. Dies weist darauf hin, dass der Test bei bestehenden Beschwerden zuverlässiger angewendet werden kann, während bei asymptomatischen Proband_innen größere Unterschiede bei der Bewertung auftreten können.

5.2 Beantwortung der Forschungsfrage

Die Forschungsfrage dieser Arbeit lautete:

„Wie hoch ist die Übereinstimmung der Ergebnisse zweier Osteopathinnen mit gleichem Ausbildungsstand bei der passiven Untersuchung der Beweglichkeit der Mandibula in Außenrotation und Innenrotation mittels Induktionstest Os mandibulare?“

Zur Beantwortung dieser Frage wurde in erster Linie der Cohens-Kappa-Wert der gesamten Stichprobe mit drei Ergebniskategorien herangezogen. Diese Auswertung bildet deswegen die zentrale Grundlage, weil sie sich auf alle Proband_innen bezieht und die ursprüngliche Einteilung des Tests („unauffällig“, „auffällig rechts“, „auffällig links“) berücksichtigt. Dabei ergab sich ein Cohens-Kappa-Wert von $\kappa = 0,33$, was laut der Klassifikation nach Landis und Koch (1977, S. 165) als eine „mäßige“ Übereinstimmung gilt. Das dazugehörige 95 %-Konfidenzintervall reicht von 0,123 bis 0,532. Dies zeigt auf, dass die Übereinstimmung bei einer Wiederholung der Studie

entweder niedriger oder auch etwas höher ausfallen könnte, aber wahrscheinlich trotzdem relativ gering wäre.

Zusätzlich zur Cohens-Kappa-Berechnung wurden auch die Gesamtübereinstimmung ($p_0 = 57,14\%$) und die Zufallsübereinstimmung ($p_c = 36,22\%$) betrachtet. Der p -Wert von 0,002 verdeutlicht, dass die Übereinstimmung nicht zufällig war. Folglich besteht tatsächlich eine gewisse Verlässlichkeit der Beurteilungen der beiden Osteopathinnen, auch wenn die Ergebnisse nicht stark übereinstimmen. Dies bedeutet, dass die beiden Testerinnen öfter zu dem gleichen Ergebnis kamen, als es allein durch Zufall zu erwarten gewesen wäre.

Um die Ergebnisse besser interpretieren zu können, wurde die Übereinstimmung zusätzlich getrennt nach Geschlecht und Symptomatik berechnet sowie in zwei Ergebniskategorien zusammengefasst. Dabei zeigten sich etwas höhere Cohens-Kappa-Werte in der dichotomen Auswertung, mit dem höchsten Wert bei den symptomatischen Proband_innen ($\kappa = 0,48$) und dem niedrigsten bei den männlichen Probanden ($\kappa = 0,25$). Alle Cohens-Kappa-Werte lagen unter dem Schwellenwert von 0,61, der laut Landis und Koch (1977, S. 165) für eine „beachtliche“ Übereinstimmung stehen würde.

Im Hinblick auf die Forschungsfrage zeigte sich, dass die Übereinstimmung der beiden Osteopathinnen nur im „mäßigen“ bis „mittelmäßigen“ Bereich lag und somit keine hohe Reproduzierbarkeit erkennen ließ. Zwar gab es eine gewisse Übereinstimmung zwischen den Ergebnissen der beiden Testerinnen, diese fiel insgesamt gering aus. Die beiden Testerinnen kamen häufiger zu dem gleichen Ergebnis, als es zufällig zu erwarten gewesen wäre. Trotzdem war diese Übereinstimmung nicht hoch genug, um von einer zuverlässigen Reproduzierbarkeit zu sprechen. Des Weiteren wird die Aussagekraft der Ergebnisse durch die kleine Stichprobe, die breiten Konfidenzintervalle und das Studiendesign eingeschränkt.

5.3 Stärken der Studie

Im Folgenden werden die zentralen Stärken der vorliegenden Studie dargestellt, die zur Qualität, Aussagekraft und Übertragbarkeit der Ergebnisse beigetragen haben.

5.3.1 Standardisierung des Tests

Der Induktionstest Os mandibulare wurde vorab durch eine gemeinsame Schulung und eine klar definierte Testanleitung standardisiert. Das entspricht einer üblichen Methode zur Erhöhung der Reliabilität diagnostischer Verfahren (Demers et al., 2021; Patijn, 2019). Dadurch konnten Unterschiede in der Durchführung minimiert werden.

5.3.2 Vergleichbarer Ausbildungsstand der Testerinnen

Beide Testerinnen verfügten über einen ähnlichen Ausbildungsstand. Dies verringert die Unterschiede, die durch verschiedene Berufserfahrungen entstehen können (Colonna & Mazzanti, 2024).

5.3.3 Einheitliche Rahmenbedingungen

Die Untersuchung wurde in einer kontrollierten Studienumgebung mit identischen Bedingungen für alle Proband_innen durchgeführt. Die gleichen Rahmenbedingungen helfen dabei, die Einflüsse durch äußere Faktoren zu verringern und die Ergebnisse vergleichbarer zu machen.

5.3.4 Breite Zusammensetzung der Stichprobe

Es wurden sowohl symptomatische als auch asymptomatische Proband_innen in die Untersuchung einbezogen. Dadurch entstand eine breitere und praxisnähere Stichprobe, die eine bessere Übertragbarkeit der Ergebnisse auf den klinischen Alltag ermöglicht.

5.3.5 Vielfältige Auswertungsansätze

Die Ergebnisse wurden in unterschiedlichen Auswertungsvarianten (drei Kategorien und dichotom) analysiert. Dadurch können die Ergebnisse genauer interpretiert und die Unterschiede zwischen den verschiedenen Auswertungskategorien erkannt werden.

5.3.6 Detaillierte statistische Auswertung

Zusätzlich zu den Cohens-Kappa-Werten wurden weitere statistische Kennwerte wie die Gesamt- und Zufallsübereinstimmung sowie die Konfidenzintervalle angegeben, um die Ergebnisse transparent darzustellen.

5.4 Einschränkungen im Bereich der angewandten Methodik

Neben möglichen Einflüssen des Studiendesigns, wie zum Beispiel der Stichprobengröße, der Auswahl der Testerinnen oder auch der Gestaltung der Trainingsphase, können auch Einschränkungen in Bezug auf die Literaturrecherche zu einer Verzerrung der Ergebnisse beigetragen haben. Aus Gründen der Übersichtlichkeit werden diese Aspekte in den folgenden Unterkapiteln einzeln dargestellt.

5.4.1 Limitierende Faktoren bezüglich der Literaturrecherche

Die Literaturrecherche wurde mit Sorgfalt durchgeführt, dennoch können bestimmte Verzerrungen aufgrund der Auswahl und Verfügbarkeit der Literatur nicht ausgeschlossen werden. Es könnten verschiedene Bias-Effekte aufgetreten sein.

Ein möglicher Einfluss ist der sogenannte „Language Bias“. Da die Recherche auf deutsch- und englischsprachige Quellen beschränkt war, könnte es sein, dass relevante Arbeiten in anderen Sprachen nicht erfasst wurden. Dadurch besteht die Möglichkeit, dass Teile des internationalen Forschungsstandes vernachlässigt worden sind (Scherfer et al., 2011, S. 138).

Auch ein „Publication Bias“ ist möglich, da Studien mit unauffälligen bzw. nicht signifikanten Ergebnissen seltener veröffentlicht werden oder sie erscheinen nur in sogenannten „grauen“ Quellen wie Abschlussarbeiten oder auch internen Berichten. In dieser Arbeit wurden solche Studien nicht berücksichtigt, was zu einer verzerrten Darstellung der Literatur geführt haben könnte (Hussy et al., 2013, S. 161; Scherfer et al., 2011, S. 138).

Trotz der Verwendung möglichst vieler Suchbegriffe und ihrer Synonyme in mehreren bekannten Datenbanken ist es darüber hinaus möglich, dass einzelne relevante Studien nicht gefunden wurden. Dies wird als „Retrieval Bias“ bezeichnet. Selbst bei einer sorgfältig geplanten Recherche lässt sich nicht ausschließen, dass bestimmte Arbeiten übersehen werden (Scherfer et al., 2011, S. 138).

5.4.2 Kritische Reflexion bezüglich des Studiendesigns

Um eine Standardisierung zu gewährleisten, wurde eine Limitierung auf ein einmaliges Testen pro Untersuchungsgang gewählt. Diese Entscheidung wurde getroffen, um die

gleichen Bedingungen für alle Proband_innen sowie für beide Testerinnen zu schaffen. Gleichzeitig kann dies jedoch als praxisfern angesehen werden, da es in der osteopathischen Praxis üblich ist, einen Test etwa bei auftretenden Unsicherheiten oder generell zur Absicherung mehrmals durchzuführen. Außerdem könnte ein mehrmaliges Testen einen mobilisierenden Effekt auf die Proband_innen haben. Dies bedeutet, dass durch den ersten Test bereits eine Veränderung im Gewebe des_der Proband_in oder auch in Bezug auf die Wahrnehmung der Testerin entstehen könnte, was dementsprechend das Ergebnis der zweiten Testerin beeinflusst hätte.

5.4.3 Kritische Reflexion bezüglich der Stichprobe

Auch wenn die Studie grundsätzlich auf eine gewisse Übereinstimmung zwischen den beiden Testerinnen hinweist, gibt es dennoch mehrere Punkte, die die Aussagekraft der Ergebnisse einschränken. Einer dieser Kritikpunkte ist die Größe der Stichprobe, die laut Patijn (2019) auf 40 Proband_innen festgelegt wurde. Insgesamt nahmen an der Studie 42 Proband_innen teil. Jedoch war die Anzahl der Proband_innen in den einzelnen Subgruppen (z. B. männlich, weiblich oder symptomatisch) deutlich kleiner.

Das 95%-Konfidenzintervall war in allen sieben Berechnungen ziemlich breit. Dies liegt vor allem darin begründet, dass die Stichprobe insgesamt eher klein war. Je kleiner eine Stichprobe ist, desto größer ist meistens auch das Konfidenzintervall (Scherfer et al., 2011, S. 209). Eine Möglichkeit wäre gewesen, vorab eine konkrete Fallzahl zu errechnen, wie es auch Patijn (2019) optional vorschlägt. Darauf wurde jedoch verzichtet, da Patijn (2019) zugleich eine Stichprobengröße von 40 Proband_innen für Interrater-Reliabilitätsstudien als ausreichend ansieht. Eine solche Fallzahlberechnung hätte vielleicht zu einer größeren Stichprobe geführt, wodurch das Konfidenzintervall entsprechend kleiner gewesen wäre und die Aussagekraft der Ergebnisse vermutlich höher ausgefallen wäre.

Die verwendete Stichprobe war in Bezug auf das Geschlecht der Proband_innen nicht homogen (26 Frauen, 16 Männer). Auch bei der Aufteilung in symptomatische und asymptomatische Proband_innen gab es Unterschiede hinsichtlich der Anzahl (19 symptomatisch, 23 asymptomatisch). Um die externe Validität der Studie zu erhalten,

wurde jedoch bewusst keine Homogenität angestrebt, da dies auch in der klinischen Praxis nicht realistisch ist.

5.4.4 Limitierende Faktoren der Testerinnen

In der mit 20 Proband_innen durchgeführten Pilotstudie wurde die von Patijn (2019) empfohlene Übereinstimmung (P_0) von mindestens 80 % erreicht, während in der Trainingsphase eine Gesamtübereinstimmung (P_0) von 81 % erzielt werden konnte. In der Hauptstudie lag der höchste Wert diesbezüglich jedoch bei 73,68 %, der niedrigste hingegen lediglich bei 50 %.

Ein möglicher Einflussfaktor könnte die unterschiedliche Zusammensetzung der Proband_innen zwischen der Pilot- und der Hauptstudie gewesen sein. In der Pilotstudie kannte die Studienleitung einen Großteil der Proband_innen persönlich, was zu einer entspannteren Atmosphäre beigetragen haben könnte. Dagegen nahmen an der Hauptstudie viele unbekannte Proband_innen teil, was für alle Beteiligten, insbesondere für die Proband_innen, mit einer höheren Zurückhaltung oder auch Nervosität verbunden gewesen sein könnte. Auch wenn die beiden Testerinnen selbst keine Proband_innen kannten, weder in der Pilot- noch in der Hauptstudie, hat sich die veränderte Gruppendynamik möglicherweise indirekt auf die Testatmosphäre ausgewirkt (Döring, 2023, S. 425–430).

Ein weiterer möglicher limitierender Faktor war der organisatorische Rahmen. Die Testerinnen reisten für die Studie aus ca. 600 km Entfernung an, was möglicherweise mit einer gewissen Anspannung und Erschöpfung verbunden war und eine ungewohnte Umgebung mit sich brachte. Solche Umstände können einen Einfluss auf die Konzentration und die Feinwahrnehmung von Tester_innen haben, vor allem bei manuellen Tests wie dem Induktionstest Os mandibulare.

Des Weiteren kommt hinzu, dass die Hauptstudie unter strengeren Bedingungen stattfand als die Pilotstudie. Es war kein Austausch zwischen den beiden Testerinnen erlaubt, um eine potenzielle Beeinflussung zu vermeiden. In der Trainingsphase dagegen war dieser Austausch noch möglich, wodurch auftretende Unsicherheiten direkt besprochen werden konnten. Ohne eine solche Rücksprache kann es allerdings

passieren, dass die Kriterien unterschiedlich interpretiert oder Entscheidungen weniger sicher getroffen werden, vor allem bei unklaren oder auch grenzwertigen Ergebnissen.

Ein weiterer Aspekt, der in diesem Zusammenhang berücksichtigt werden sollte, betrifft die körperlichen Voraussetzungen der beiden Testerinnen. Beide wiesen eine ähnliche Körpergröße und einen vergleichbaren Körperbau auf und waren zudem Rechtshänderinnen. Dank dieser Gemeinsamkeiten konnte die Arbeitshöhe der Therapieliege während der gesamten Studie auf die gleiche Höhe eingestellt werden. Unterschiede hinsichtlich der Körpergröße oder der Händigkeit, die sich möglicherweise auf die Kraftdosierung, das Feingefühl oder die Bewegungsausführung auswirken könnten, spielten in diesem Fall somit keine Rolle. Dadurch hatten diese Faktoren wahrscheinlich keinen großen Einfluss auf die Testergebnisse.

5.4.5 Durchführung und Interpretation des Tests

Eine weitere Limitation betrifft die Durchführung des Tests selbst. Auch wenn vorab eine Schulung und eine gemeinsame Testphase stattfanden, kam es zwischen den beiden Testerinnen im Nachhinein zu Diskussionen dahingehend, wann ein Testergebnis als auffällig gilt und wann nicht. In der Übungsphase war die Übereinstimmung diesbezüglich hoch (über 80 %), doch während der eigentlichen Studie zeigte sich, dass es doch unterschiedliche Testinterpretationen gab. Dies könnte erklären, warum die beiden Testerinnen nicht häufiger die gleichen Ergebnisse erzielten. Insbesondere bei Tests, die auf dem Tastsinn der oder des Testenden beruhen, gestaltet es sich oft schwierig für alle Tester_innen einheitliche und verständliche Bewertungskriterien festzulegen.

5.4.6 Einfluss äußerer Faktoren auf die Untersuchung

Die körperliche Verfassung der Proband_innen während der Testung stellt einen weiteren limitierenden Aspekt dar. So berichten die Testerinnen nach der Studie, dass einige Proband_innen während der Untersuchung angespannt wirkten. Diese Anspannung könnte zugleich zu einer erhöhten Muskelspannung im Bereich des Kiefers geführt haben und somit die Wahrnehmung und Beurteilung seitens der Testerinnen beeinflusst haben. Diesbezüglich hätte sich eine entspannte Ausgangssituation vermutlich besser erwiesen, lässt sich im Zuge einer solchen Studie allerdings nicht

immer vollständig sicherstellen. Obwohl für alle Proband_innen genug Zeit eingeplant wurde, ist eine gewisse Grundanspannung unter diesen Testbedingungen nicht gänzlich auszuschließen.

5.5 Vergleich mit anderen Studien

In der vorliegenden Arbeit wurden überwiegend „mäßige bis mittelmäßige“ Cohens-Kappa-Werte gemessen, was nach der Klassifikation von Landis und Koch (1977, S. 165) einem Bereich von 0,21 bis 0,60 entspricht. Dies weist auf eine teilweise Übereinstimmung zwischen den Testerinnen hin. Diese Werte stellen in der Literatur keine Ausnahme dar, so wurden auch in anderen Studien zur Interrater-Reliabilität in der Osteopathie ähnliche Ergebnisse erzielt. Beispielsweise zeigte sich in einer Übersichtsarbeit von Guillaud et al. (2016), dass die diagnostischen Techniken der kranialen Osteopathie in den meisten eingeschlossenen Studien sehr niedrige Cohens-Kappa-Werte aufweisen. Teilweise lagen diese sogar im negativen Bereich. Daher kamen die Autor_innen zu dem Schluss, dass es bisher keine ausreichenden Belege für die zuverlässige Anwendung dieser osteopathischen Techniken gibt.

Auch die Studie von Consorti et al. (2018) zur Beurteilung von osteopathischen Tests am Sakrum brachte insgesamt niedrige Werte hervor. Hierbei lagen die Cohens-Kappa-Werte bei verschiedenen Parametern wie der Gewebespannung, der Asymmetrie und der Bewegungseinschränkung zwischen 0,06 und 0,34. Damit liegen diese Werte sogar unterhalb der Werte, die in dieser Arbeit gemessen wurden.

Ein anderer Ansatz wurde von Colonna und Mazzanti (2024) beschrieben. In ihrer Übersichtsarbeit betonten sie, dass palpatorische Tests nicht allein auf dem Tastsinn, sondern auch auf visuellen Eindrücken basieren. Daher schlugen die Autor_innen vor, stattdessen von „visuell-palpatorischen“ Tests zu sprechen. Außerdem wird betont, dass ein gezieltes Training und ein standardisiertes Vorgehen notwendig sind, um die Reproduzierbarkeit solcher Tests zu verbessern.

Auch die Studie von Chang et al. (2021) verdeutlicht, dass manualtherapeutische Techniken eine niedrige Interrater-Reliabilität aufweisen können. In dieser Untersuchung wurde die Palpation an der Brustwirbelsäule mit einer

Ultraschallmessung verglichen. Hierbei wurde die Rotationsrichtung der Querfortsätze (T2–T5) zunächst osteopathisch palpatorisch bestimmt. Anschließend wurden diese Ergebnisse mit Ultraschallmessungen der Querfortsätze im Verhältnis zum Dornfortsatz des darüberliegenden Wirbels verglichen. Dabei ergaben sich Cohens-Kappa-Werte zwischen 0,02 und 0,14. Die Übereinstimmung zwischen zwei osteopathischen Tester_innen lag insgesamt nur bei etwa 30 %.

Eine weitere Studie stammt von Demers et al. (2021), die die Interrater-Reliabilität von Tests an kranialen Knochen untersuchten. Die Tester_innen hatten zuvor ein gemeinsames Schulungstraining durchlaufen. Je nach untersuchtem Knochen lagen die Cohens-Kappa-Werte zwischen 0,48 und 0,96, wobei besonders hohe Werte bei den parietalen und frontalen Knochen gemessen wurden. Diese Ergebnisse zeigen, dass durch gezielte Schulung zwischen verschiedenen Tester_innen bessere Übereinstimmungen erreicht werden können.

In der vorliegenden Arbeit lagen die gemessenen Cohens-Kappa-Werte je nach Auswertung und Subgruppe zwischen 0,25 und 0,48. Die höchste Übereinstimmung wurde bei der symptomatischen Gruppe mit zwei Ergebniskategorien gemessen ($\kappa = 0,48$), die niedrigste hingegen bei den männlichen Probanden mit drei Kategorien ($\kappa = 0,25$). Die Ergebnisse dieser Arbeit bestätigen die bereits mehrfach erwähnten Schwierigkeiten, palpatorische osteopathische Tests mit einer hohen Reliabilität durchzuführen (Chang et al., 2021; Consorti et al., 2018; Guillaud et al., 2016, 2018). Gleichzeitig zeigen sie, dass sich durch Training, klare Kriterien und ein einheitliches Vorgehen die Interrater-Reliabilität verbessern lässt (Demers et al., 2021; Patijn, 2019). Damit fügen sich die Ergebnisse gut in die bestehende Literatur ein.

5.6 Übertragung der Ergebnisse auf die Osteopathie

Die Ergebnisse der vorliegenden Arbeit verdeutlichen, dass der Induktionstest Os mandibulare zwar eine gewisse, jedoch keine hohe Reliabilität zwischen den Ergebnissen zweier Testerinnen aufweist. In der osteopathischen Praxis bedeutet dies, dass der Test nicht als einziges Diagnoseinstrument verwendet werden sollte, sondern immer im Gesamtkontext der Untersuchung gesehen werden muss. Deswegen reicht das Ergebnis eines einzelnen Tests in der Regel nicht aus, um eine sichere

osteopathische Diagnose zu stellen, sondern sollte immer im Zusammenhang mit anderen Untersuchungsergebnissen betrachtet werden (Patijn, 2019).

Insbesondere bei symptomatischen Proband_innen zeigte sich im Rahmen dieser Untersuchung eine bessere Übereinstimmung. Jedoch kann die Stichprobe von $n = 19$ nicht als aussagekräftig eingestuft werden. Dies könnte einen Hinweis darauf liefern, dass die Durchführung des Tests vor allem dann sinnvoll ist, wenn bereits ein klinischer Anfangsverdacht besteht. In der osteopathischen Praxis wird in einem ersten Schritt meist ein ausführliches Anamnesegespräch geführt. Daraufhin wird eine gezielte Untersuchung je nach Klinik vorgenommen, wobei der Induktionstest Os mandibulare als ergänzendes Instrument zur Hypothesenabsicherung dienen könnte. Für asymptomatische Proband_innen oder auch bei unklaren Ergebnissen könnte der Test hingegen weniger aussagekräftig sein.

Für die osteopathische Ausbildung bedeutet dies, dass die Standardisierung und Schulung solcher Tests geprüft werden sollten. Demnach verdeutlichen die vorliegenden Ergebnisse, dass selbst bei einem vergleichbaren Ausbildungsstand der Tester_innen Unterschiede hinsichtlich der Beurteilung auftreten können. Folglich könnte eine bessere Schulung, wiederholte Übungseinheiten und klar definierte Bewertungskriterien dabei helfen, die Vergleichbarkeit zwischen den Ergebnissen verschiedener Tester_innen zu verbessern.

Des Weiteren zeigt die Studie auf, dass bei manuellen Tests grundsätzlich Grenzen im Hinblick auf die Reproduzierbarkeit bestehen. Dies ist insbesondere für die tägliche Praxis wichtig, in der Entscheidungen basierend auf der Palpation bzw. dem Tastsinn getroffen werden. Aus diesem Grund sollten die Testergebnisse stets im Kontext betrachtet und bei eventuellen Unsicherheiten durch weitere Untersuchungsmethoden ergänzt werden.

Insgesamt zeigt die vorliegende Arbeit auf, dass der Induktionstest Os mandibulare als Teil einer osteopathischen Untersuchung durchaus sinnvoll eingesetzt werden kann. Voraussetzung dafür ist, dass er richtig angewendet, regelmäßig geübt und nicht isoliert

bewertet wird. Hierbei sollten in zukünftigen Studien eine größere Stichprobe und klarere Definitionskriterien verwendet werden, um die Aussagekraft weiter zu erhöhen.

6 Zusammenfassung und Schlussbetrachtung

Die vorliegende Forschungsarbeit befasst sich mit der Frage, wie hoch die Übereinstimmung der Ergebnisse zweier Osteopathinnen mit ähnlichem Ausbildungsstand bei der Durchführung des Induktionstests Os mandibulare ist. Zu diesem Zweck wurden sowohl symptomatische als auch asymptomatische Proband_innen untersucht.

Bei dem Induktionstest Os mandibulare handelt es sich um ein palpatorisches Verfahren der osteopathischen Diagnostik. Er dient der qualitativen und quantitativen Beurteilung der Beweglichkeit des Os mandibulare in Außen- und Innenrotation. In der bisherigen Forschung wurde dieser Test weder hinsichtlich seiner Reliabilität noch hinsichtlich seiner Validität untersucht.

Die vorliegende Studie wurde mit 42 Proband_innen durchgeführt, von denen 19 symptomatisch und 23 asymptomatisch waren. Die Reihenfolge der Testdurchführung wurde randomisiert festgelegt. Eine Verblindung war gegeben. Ebenfalls war jegliche verbale Kommunikation zwischen den beiden Testerinnen und den Proband_innen untersagt. Auch ein Austausch zwischen den beiden Testerinnen selbst war während der Datenerhebung nicht erlaubt.

Vorab wurde die gesamte Testdurchführung standardisiert, diese wurde ein paar Wochen zuvor in einer Pilotstudie getestet. Im Anschluss an diese Trainingsphase wurde die Hauptstudie durchgeführt. Während der Testung lagen die Proband_innen in Rückenlage, wohingegen die jeweilige Testerin an der Kopfseite Platz nahm. Es standen drei Bewertungskategorien zur Verfügung, nämlich „unauffällig“, „auffällig rechts“ und „auffällig links“.

Die Auswertung erfolgte anhand des Cohens-Kappa-Koeffizienten für acht definierte Untergruppen. Die Ergebnisse zeigen, dass die Interrater-Reliabilität insgesamt als „mäßig“ bis „mittelmäßig“ einzustufen ist. So lag der höchste Cohens-Kappa-Wert bei $\kappa = 0,48$, und der niedrigste bei $\kappa = 0,25$. Damit befand sich die Übereinstimmung zwar über dem Zufall, war jedoch nicht hoch genug, um den Test als zuverlässig einzustufen.

Für die gesamte Stichprobe ($n = 42$) wurde bei der Auswertung in drei Ergebniskategorien ein Cohens-Kappa-Wert von $\kappa = 0,33$ erzielt. Die weibliche Stichprobe ($n = 26$) erreichte $\kappa = 0,39$, die männliche ($n = 16$) hingegen $\kappa = 0,25$. Werden die Auswertungen von zwei Ergebniskategorien betrachtet, so ergeben sich leicht höhere Werte. Die gesamte Stichprobe erreichte $\kappa = 0,41$, wobei die weibliche Gruppe $\kappa = 0,40$ und die männliche Gruppe $\kappa = 0,43$ erzielten. Der höchste Wert von $\kappa = 0,48$ wurde wiederum bei den symptomatischen Proband_innen ermittelt. Im Gegensatz dazu erreichten asymptomatische Proband_innen einen Wert von $\kappa = 0,35$. Allerdings ist die Aussagekraft dieser Subgruppen aufgrund der geringen Größe eingeschränkt.

In der osteopathischen Praxis bedeuten diese Resultate, dass der Induktionstest Os mandibulare nicht als alleiniges Diagnoseinstrument verwendet, sondern vielmehr stets im Zusammenspiel mit anderen Untersuchungsergebnissen betrachtet werden sollte. Demnach deuten die Ergebnisse darauf hin, dass der Test vor allem in gezielten klinischen Situationen sowie insbesondere bei symptomatischen Patient_innen eine unterstützende Rolle einnehmen kann. Für die osteopathische Ausbildung ergibt sich hieraus, dass klar definierte Bewertungskriterien sowie eine intensive und standardisierte Schulung notwendig sind, um vergleichbare ebenso wie verlässlichere Ergebnisse zu erzielen.

6.1 Empfehlungen für weitere Studien

Die in Kapitel 6 dargestellten Ergebnisse und Überlegungen führen zu den nachfolgenden Empfehlungen für die weiterführende Forschung.

6.1.1 Fallzahlberechnung

Für weitere Studien oder auch eine Wiederholung der Untersuchung empfiehlt die Autorin eine Fallzahlberechnung. Auf diese Weise könnte bereits im Vorfeld festgelegt werden, wie viele Proband_innen zur Erzielung aussagekräftiger Ergebnisse benötigt werden. Zugleich würde dies dabei helfen, die weiten Konfidenzintervalle zu verkleinern.

6.1.2 Stichprobengröße

Die Stichprobengröße in der vorliegenden Forschung lag bei insgesamt 42 Proband_innen, wobei die Anzahl der Proband_innen in den einzelnen Subgruppen

(z. B. männlich, weiblich, symptomatisch) jedoch teilweise recht klein war. Für zukünftige Studien wäre es daher sinnvoll, größere Gruppen zu untersuchen, um genauere Aussagen zu machen und die Ergebnisse besser beurteilen zu können.

6.1.3 Intrarater-Reliabilität

In dieser Arbeit wurde ausschließlich die Interrater-Reliabilität untersucht, das heißt die Übereinstimmung zwischen den Ergebnissen zweier verschiedener Testerinnen. Um die Aussagekraft des Induktionstest Os mandibulare besser beurteilen zu können, wäre es in zukünftigen Studien sinnvoll, zusätzlich die Intrarater-Reliabilität zu untersuchen. Dabei wird wiederum überprüft, ob der_die Tester_in bei wiederholter Durchführung eines Tests zu denselben Ergebnissen gelangt (Scherfer et al., 2011, S. 250). Somit könnte die Erhebung der Intrarater-Reliabilität weitere Erkenntnisse dahingehend liefern, wie konstant der Test ist und ob die Durchführung auch bei derselben Testperson zuverlässig bleibt.

Darüber hinaus könnte der Test auch auf das Gütekriterium der Validität hin überprüft werden, also darauf, ob er tatsächlich das misst, was er zu messen vorgibt. Allerdings weisen Scherfer et al. (2011, S. 250) darauf hin, dass Messverfahren ohne ausreichende Reliabilität ebenfalls keine hohe Validität erreichen können.

6.1.4 Weiterentwicklung der Testmethode

Im Rahmen zukünftiger Studien könnte untersucht werden, ob gezielte Anpassungen der Testdurchführung wie visuelle Ergänzungen, alternative Handpositionen oder auch standardisierte Bewertungsrichtlinien die Reliabilität des Induktionstest Os mandibulare verbessern können.

6.2 Fazit

Der Induktionstest Os mandibulare zeigte in der vorliegenden Forschungsarbeit lediglich eine „mäßige“ bis „mittelmäßige“ Übereinstimmung zwischen den beiden Testerinnen. Daher kann er nicht als verlässliches Diagnoseinstrument für die osteopathische Praxis empfohlen werden. Zwar kann der Test Hinweise auf mögliche Einschränkungen oder auch Spannungen im Bereich der Mandibula liefern, allerdings ist er nicht zuverlässig

genug, um als alleiniges Kriterium für eine osteopathische Diagnose herangezogen zu werden.

Trotz dieser Einschränkung weist der Induktionstest Os mandibulare einige Vorteile auf. Er ist schnell durchführbar, für die Patient_innen nicht mit unangenehmen Empfindungen verbunden und lässt sich auf unkomplizierte Weise in die Untersuchung integrieren. Demnach besitzt er eine hohe Praktikabilität. Somit kann der Induktionstest Os mandibulare eine sinnvolle Ergänzung der osteopathischen Befundung darstellen.

Zusammenfassend kann auf Basis der vorliegenden Ergebnisse allerdings keine Empfehlung für die alleinige Anwendung des Induktionstests Os mandibulare im klinischen Alltag gegeben werden.

LITERATURVERZEICHNIS

- Asiri, S. N., Tadlock, L. P., & Buschang, P. H. (2019). The prevalence of clinically meaningful malocclusion among US adults. *Orthodontics & Craniofacial Research*, 22(4), 321–328. <https://doi.org/10.1111/ocr.12328>
- Basile, F., Scionti, R., & Petracca, M. (2017). Diagnostic reliability of osteopathic tests: A systematic review. *International Journal of Osteopathic Medicine*, 25, 21–29. <https://doi.org/10.1016/j.ijosm.2017.03.004>
- Bordoni, B. (2020). The Five Diaphragms in Osteopathic Manipulative Medicine: Myofascial Relationships, Part 1. *Cureus*. <https://doi.org/10.7759/cureus.7794>
- Bordoni, B., Morabito, B., Mitrano, R., Simonelli, M., & Toccafondi, A. (2018). The Anatomical Relationships of the Tongue with the Body System. *Cureus*. <https://doi.org/10.7759/cureus.3695>
- Brinkmann, K. M. (2022). *Vergleichende Auswertung von Realtime(Echtzeit)-MRT-Aufnahmen des Kiefergelenks bei Bildraten von 3 und 15 Bildern pro Sekunde mit der klinischen Untersuchung anhand der RDC/TMD* [Dissertation/Thesis]. Georg-August-Universität Göttingen, Medizinische Fakultät.
- Chang, S., Maddox, J., Berg, E., Kim, K., Messier, S., Swanson, L., Dobrusin, R., Stein, A. B., Nakken, G. N., Noble, J., & Nydam, R. (2021). Evaluating for a correlation between osteopathic examination and ultrasonography on thoracic spine asymmetry. *Journal of Osteopathic Medicine*, 122(1), 31–43. <https://doi.org/10.1515/jom-2021-0020>
- Coffey, D. (2005). C1 Somatic Dysfunction and Unilateral Retroorbital Cephalalgia. *The AAO Journal*, 15(2), 11–18.

- Colonna, S., & Mazzanti, M. (2024). Pelvic Palpatory Tests in Manual Therapy and Osteopathy: A Critical Review of the Literature and Suggestions for New Research. *Cureus*. <https://doi.org/10.7759/cureus.64066>
- Consorti, G., Basile, F., Pugliese, L., & Petracca, M. (2018). Interrater Reliability of Osteopathic Sacral Palpatory Diagnostic Tests Among Osteopathy Students. *Journal of Osteopathic Medicine*, 118(10), 637–644. <https://doi.org/10.7556/jaoa.2018.132>
- Corts, M. (2020). *Diagnoseleitfaden Osteopathie* (3. Aufl.). Thieme.
- Corts, M. (2023). *Anatomie für Osteopathen—Kopf* (1. Aufl.). Thieme.
- 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.
- Demers, K., Morin, C., Collette, L., & DeMont, R. (2021). Moderate to Substantial Interrater Reliability in the Assessment of Cranial Bone Mobility Restrictions. *The Journal of Alternative and Complementary Medicine*, 27(3), 263–272. <https://doi.org/10.1089/acm.2020.0325>
- Detoni, R., Hartz, C. S., Fusatto, E. L., Bicalho, E., Nascimento-Moraes, K. S. G., Rizzatti-Barbosa, C. M., & Lopes, F. O. T. (2022). Relationship between osteopathic manipulative treatment of the temporomandibular joint, molar shim and the orthostatic position: A randomized, controlled and double blinded study. *Journal of Bodywork and Movement Therapies*, 29, 187–197. <https://doi.org/10.1016/j.jbmt.2021.09.021>

- Donker, D. K., Hasman, A., & Van Geijn, H. P. (1993). Interpretation of low kappa values. *International Journal of Bio-Medical Computing*, 33(1), 55–64. [https://doi.org/10.1016/0020-7101\(93\)90059-F](https://doi.org/10.1016/0020-7101(93)90059-F)
- Döring, N. (2023). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (6. Aufl. 2022). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-64762-2>
- Emshoff, R., & Rudisch, A. (2001). Validity of clinical diagnostic criteria for temporomandibular disorders. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, 91(1), 50–55. <https://doi.org/10.1067/moe.2001.111129>
- Fernandes, P. R. B., De Vasconcellos, H. A., Okeson, J. P., Bastos, R. L., & Maia, M. L. T. (2003). The Anatomical Relationship Between the Position of the Auriculotemporal Nerve and Mandibular Condyle. *CRANIO®*, 21(3), 165–171. <https://doi.org/10.1080/08869634.2003.11746246>
- Goh, B. T., Lee, S., Tideman, H., & Stoelinga, P. J. W. (2008). Mandibular reconstruction in adults: A review. *International Journal of Oral and Maxillofacial Surgery*, 37(7), 597–605. <https://doi.org/10.1016/j.ijom.2008.03.002>
- Grouven, U., Bender, R., Ziegler, A., & Lange, S. (2007). Der Kappa-Koeffizient. *DMW - Deutsche Medizinische Wochenschrift*, 132(S 01), e65–e68. <https://doi.org/10.1055/s-2007-959046>
- Guillaud, A., Darbois, N., Monvoisin, R., & Pinsault, N. (2016). Reliability of Diagnosis and Clinical Efficacy of Cranial Osteopathy: A Systematic Review. *PLOS ONE*, 11(12), e0167823. <https://doi.org/10.1371/journal.pone.0167823>
- Guillaud, A., Darbois, N., Monvoisin, R., & Pinsault, N. (2018). Reliability of diagnosis and clinical efficacy of visceral osteopathy: A systematic review. *BMC*

- Complementary and Alternative Medicine*, 18(1), 65.
<https://doi.org/10.1186/s12906-018-2098-8>
- Hinkelthein, E., & Weitendorff, A. (Hrsg.). (2022). *Biomechanik in Osteopathischer und Manueller Medizin: Physiologie - Dysfunktion - Behandlung* (2., aktualisierte Auflage). Thieme.
- Hofmann, E., Ebker, T., & Koerdt, S. (2023). Hyperplastische Fehlbildungen des Gesichtsschädels. *Die MKG-Chirurgie*, 16(1), 17–26.
<https://doi.org/10.1007/s12285-022-00396-9>
- Hülse, M., & Losert-Bruggner, B. (2002). Der Einfluss der Kopfgelenke und/oder der Kiefergelenke auf die Hüftabduktion Ein einfacher Test zur Frage, ob eine CMD durch eine HWS-Manipulation beeinflusst werden konnte. *Manuelle Medizin*, 40(2), 97–100. <https://doi.org/10.1007/s00337-002-0131-x>
- Hussy, W., Schreier, M., & Echterhoff, G. (2013). *Forschungsmethoden in Psychologie und Sozialwissenschaften: Für Bachelor* (2. Aufl.). Springer.
- Jo, J. H., Park, J. W., Jang, J. H., & Chung, J. W. (2022). Hyoid bone position as an indicator of severe obstructive sleep apnea. *BMC Pulmonary Medicine*, 22(1), 349. <https://doi.org/10.1186/s12890-022-02146-0>
- Kopp, M., Wiesmueller, M., Buchbender, M., Kesting, M., Nagel, A. M., May, M. S., Uder, M., Roemer, F. W., & Heiss, R. (2024). MRI of Temporomandibular Joint Disorders: A Comparative Study of 0.55 T and 1.5 T MRI. *Investigative Radiology*, 59(3), 223–229. <https://doi.org/10.1097/RLI.0000000000001008>
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159. <https://doi.org/10.2307/2529310>
- Liem, T. (Hrsg.). (2019). *Praxis der Kraniosakralen Osteopathie* (4. Aufl., S. b-006-161640). Georg Thieme Verlag. <https://doi.org/10.1055/b-006-161640>

- Liem, T., & Dobler, T. K. (Hrsg.). (2018). *Checkliste Kraniosakrale Osteopathie* (3. unveränderte Auflage). Haug Fachbuch.
- Limchaichana, N., Petersson, A., & Rohlin, M. (2006). The efficacy of magnetic resonance imaging in the diagnosis of degenerative and inflammatory temporomandibular joint disorders: A systematic literature review. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, *102*(4), 521–536. <https://doi.org/10.1016/j.tripleo.2006.02.001>
- Lövgren, A., Häggman-Henrikson, B., Visscher, C. M., Lobbezoo, F., Marklund, S., & Wänman, A. (2016). Temporomandibular pain and jaw dysfunction at different ages covering the lifespan—A population based study. *European Journal of Pain (London, England)*, *20*(4), 532–540. <https://doi.org/10.1002/ejp.755>
- Masoudi, P., Karimi, N., Abdollahi, I., Bakhshi, E., Moravej, S., & Aghazadeh, A. (2024). Applicability of using dynamic MRI to evaluate alleged cranial rhythmic impulse (CRI). *BMC Musculoskeletal Disorders*, *25*(1), 941. <https://doi.org/10.1186/s12891-024-08064-y>
- Mayer, H., Nonn, C., Osterbrink, J., & Evers, G. (2004). Qualitätskriterien von Assessmentinstrumenten – Cohen’s Kappa als Maß der Interrater-Reliabilität (Teil 1). *Pflege*, *17*(1), 36–46. <https://doi.org/10.1024/1012-5302.17.1.36>
- Minelli, L., Van Der Lei, B., & Mendelson, B. C. (2024). The Deep Fascia of the Head and Neck Revisited: Relationship with the Facial Nerve and Implications for Rhytidectomy. *Plastic & Reconstructive Surgery*, *153*(6), 1273–1288. <https://doi.org/10.1097/PRS.00000000000010556>
- Nolet, P. S., Yu, H., Côté, P., Meyer, A.-L., Kristman, V. L., Sutton, D., Murnaghan, K., & Lemeunier, N. (2021). Reliability and validity of manual palpation for the assessment of patients with low back pain: A systematic and critical review.

- Chiropractic & Manual Therapies*, 29(1), 33. <https://doi.org/10.1186/s12998-021-00384-3>
- Ohmure, H., Miyawaki, S., Nagata, J., Ikeda, K., Yamasaki, K., & Al-Kalaly, A. (2008). Influence of forward head posture on condylar position. *Journal of Oral Rehabilitation*, 35(11), 795–800. <https://doi.org/10.1111/j.1365-2842.2007.01834.x>
- Orhan, K., Seki, U., & Rozylo-Kalinowska, I. (2017). Diagnostic accuracy of magnetic resonance imaging and clinical signs of temporomandibular joint disorders: A 10-year research update review. *Oral Radiology*, 33(2), 81–91. <https://doi.org/10.1007/s11282-017-0278-8>
- Patijn, J. (2019). Reproducibility protocol for diagnostic procedures in Manual/Musculoskeletal Medicine: Edition 2019. *Manuelle Medizin*, 57(6), 451–479. <https://doi.org/10.1007/s00337-019-00581-5>
- Peters, K. (2024). Kraniosakrale Osteopathie. In W. Langer & E. Hebgen (Hrsg.), *Lehrbuch Osteopathie* (3. Auflage, S. 295–391). Thieme.
- Pirri, C., Fede, C., Fan, C., Guidolin, D., Macchi, V., De Caro, R., & Stecco, C. (2021). Ultrasound Imaging of Head/Neck Muscles and Their Fasciae: An Observational Study. *Frontiers in Rehabilitation Sciences*, 2, 743553. <https://doi.org/10.3389/fresc.2021.743553>
- Plato, G. (2001). Gesichtsschmerz aus manualmedizinischer und kieferorthopädischer Sicht. *Manuelle Medizin*, 39(5), 254–258. <https://doi.org/10.1007/s003370170022>
- Prodinger-Glöckl, D. (2013). *CMD in der Osteopathie: Interdisziplinäre Zusammenarbeit mit der Kieferorthopädie* (1. Auflage). Haug Fachbuch.

- Reißmann, D. (2017). Therapie von kranio-mandibulären Dysfunktionen. *Zahnmedizin up2date*, 11(02), 179–202. <https://doi.org/10.1055/s-0042-116618>
- Sakaguchi, K., Mehta, N. R., Abdallah, E. F., Forgione, A. G., Hirayama, H., Kawasaki, T., & Yokoyama, A. (2007). Examination of the Relationship Between Mandibular Position and Body Posture. *CRANIO®*, 25(4), 237–249. <https://doi.org/10.1179/crn.2007.037>
- Schäfer, A., & Schöttker-Königer, T. (2015). *Statistik und quantitative Methoden für Gesundheitsfachberufe*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-45519-7>
- Scherfer, E., Boßmann, T., & Herbert, R. D. (2011). *Forschung verstehen: Ein Grundkurs in evidenzbasierter Praxis* (2., überarbeitete und erweiterte Auflage). Pflaum.
- Schünke, M., Schulte, E., Schumacher, U., Voll, M., & Wesker, K. (2022). *Prometheus LernAtlas der Anatomie—Kopf, Hals und Neuroanatomie* (6., vollständig überarbeitete Auflage). Georg Thieme Verlag. <https://doi.org/10.1055/b0000000615>
- Sim, J., & Wright, C. C. (2005). The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Physical Therapy*, 85(3), 257–268. <https://doi.org/10.1093/ptj/85.3.257>
- Thiesen, G., Gribel, B. F., Kim, K. B., Pereira, K. C. R., & Freitas, M. P. M. (2016). Prevalence and Associated Factors of Mandibular Asymmetry in an Adult Population: *Journal of Craniofacial Surgery*, 28(3), e199–e203. <https://doi.org/10.1097/SCS.00000000000003371>
- Tramontano, M., Tamburella, F., Dal Farra, F., Bergna, A., Lunghi, C., Innocenti, M., Cavera, F., Savini, F., Manzo, V., & D'Alessandro, G. (2021). International

Overview of Somatic Dysfunction Assessment and Treatment in Osteopathic Research: A Scoping Review. *Healthcare*, 10(1), 28. <https://doi.org/10.3390/healthcare10010028>

- Valesan, L. F., Da-Cas, C. D., Réus, J. C., Denardin, A. C. S., Garanhani, R. R., Bonotto, D., Januzzi, E., & De Souza, B. D. M. (2021). Prevalence of temporomandibular joint disorders: A systematic review and meta-analysis. *Clinical Oral Investigations*, 25(2), 441–453. <https://doi.org/10.1007/s00784-020-03710-w>
- Weber, P., Castilhos Rodrigues Corrêa, E., de Paula Bolzan, G., dos Santos Ferreira, F., Corrêa Soares, J., & Toniolo da Silva, A. M. (2012). Relationship between craniocervical posture, mandible and hyoid bone and influence on alimentary functions. *Brazilian Journal of Oral Sciences*, 11(2). <https://doi-10.130z5a7d70668.han3.donau-uni.ac.at/10.20396/bjos.v11i2.8641445>
- Wong, R. C. W., Tideman, H., Kin, L., & Merkx, M. A. W. (2010). Biomechanics of mandibular reconstruction: A review. *International Journal of Oral and Maxillofacial Surgery*, 39(4), 313–319. <https://doi.org/10.1016/j.ijom.2009.11.003>

ABBILDUNGSVERZEICHNIS

Abbildung 1: Anteil der asymptomatischen und symptomatischen Proband_innen ...	31
Abbildung 2: Handposition für den Induktionstest Os mandibulare	34

TABELLENVERZEICHNIS

Tabelle 1: Übersicht bildgebender Verfahren	22
Tabelle 2: Übersicht osteopathischer Testverfahren	25
Tabelle 3: Altersverteilung der teilnehmenden Proband_innen.....	31
Tabelle 4: Übersicht der Symptome der Proband_innen	32
Tabelle 5: Vierfeldtafel (Kreuztabelle) - Quelle: adaptiert nach Schäfer & Schöttker-Königer (2015, S. 137).....	38
Tabelle 6: Mehrfeldertafel – Quelle: adaptiert nach Schäfer und Schöttker-Königer (2015, S. 136)	38
Tabelle 7: Klassifikation der Cohens-Kappa-Koeffizienten – Quelle: adaptiert nach Landis und Koch (1977, S. 165)	40
Tabelle 8: Kreuztabelle – drei Ergebniskategorien der gesamten Stichprobe	42
Tabelle 9: Kreuztabelle – drei Ergebniskategorien der weiblichen Stichprobe	42
Tabelle 10: Kreuztabelle – drei Ergebniskategorien der männlichen Stichprobe.....	43
Tabelle 11: Kreuztabelle – zwei Ergebniskategorien der gesamten Stichprobe	44
Tabelle 12: Kreuztabelle – zwei Ergebniskategorien der weiblichen Stichprobe	44
Tabelle 13: Kreuztabelle – zwei Ergebniskategorien der männlichen Stichprobe	45
Tabelle 14: Kreuztabelle – zwei Ergebniskategorien der asymptomatischen Stichprobe	46
Tabelle 15: Kreuztabelle – zwei Ergebniskategorien der symptomatischen Stichprobe	46
Tabelle 16: Übersicht über alle errechneten Cohens-Kappa-Koeffizienten mit zugehörigem 95%-Konfidenzintervall.....	47
Tabelle 17: Verteilung der Bewertungen durch Testerin 1 und Testerin 2.....	48
Tabelle 18: Geschlechtsverteilung der auffälligen Testergebnisse	48

ABKÜRZUNGSVERZEICHNIS

Abkürzung	Bedeutung
CMD	Craniomandibuläre Dysfunktion
CRI	Cranial Rhythmic Impulse
DVT	Digitale Volumentomographie
FHP	Forward Head Posture
Lig.	Ligamentum
MRT	Magnetresonanztomographie
OPT	Orthopantomogramm
PRM	Primär respiratorischer Mechanismus
SD	Standardabweichung
SPSS	Statistical Package for the Social Sciences
SSB	Synchondrosis sphenobasilaris
TMG	Temporomandibuläres Gelenk
p-Wert	Statistischer Wahrscheinlichkeitswert
pc	Zufallsübereinstimmung
p ₀	Gesamtübereinstimmung
95% KI	95% Konfidenzintervall

ANHANG ÜBERSICHT

Anhang A	Literaturrecherche
Anhang B	Rohdaten
Anhang C	Aushang zur Rekrutierung der Proband_innen
Anhang D	Fragebogen bezüglich Ausschlusskriterien und Einverständniserklärung
Anhang E	Studieninformation
Anhang F	Fragebogen zu bestehenden Symptomatiken
Anhang G	Beurteilungsbogen

ANHANG A

Literaturrecherche

Von Februar 2023 bis Oktober 2025 wurde eine systematische Literaturrecherche in folgenden Datenbanken durchgeführt: PubMed, Osteopathic Research Web, PEDro, Google Scholar, Cochrane Library, biber.

Gesucht wurde nach den unten angeführten Begriffen in Verbindung mit verschiedenen Booleschen Operatoren („AND“ und „OR“):

- Mandibula/ os mandibulare/ TMG/ CMD
- reliability/ inter-rater reliability/ inter-observer reliability/ intertester reliability
- motion/ mobility/ motion test/ manual examination/ palpation
- cranial bone/ craniosacral/ cranio*/ osteopath*

Des Weiteren wurde mit dem Schneeballsystem recherchiert.

ANHANG B

Rohdaten der Testung vom 14. und 15. Juni 2025

(Testung mit 3 Ergebniskategorien; 0 = unauffällig, 1 = auffällig rechts, 2 = auffällig links)

Probandennummer	Testperson 1	Testperson 2
1	1	2
2	0	0
3	0	0
4	2	2
5	2	0
6	0	0
7	0	0
8	2	2
9	1	2
10	2	0
11	0	1
12	0	0
13	2	0
14	0	0
15	0	1
16	1	1
17	2	2
18	0	0
19	0	0
20	0	0
21	2	2
22	2	2
23	1	2
24	0	0
25	2	0
26	0	0
27	2	0
28	1	0
29	2	0
30	1	2
31	0	0
32	1	0
33	2	2
34	1	0
35	2	0
36	0	0
37	1	0
38	0	0
39	1	2
40	1	1
41	2	2
42	2	2

ANHANG C

StudienteilnehmerInnen gesucht

Sehr geehrte Damen und Herren,

ich befinde mich gerade im **Masterstudium Osteopathie** an der Universität für Weiterbildung Krems.

Für meine Studie, die im Rahmen meiner Masterthesis in dieser Ausbildung stattfindet, suche ich **StudienteilnehmerInnen ab 18 Jahren**, die bereit sind, sich zirka 20 Minuten der Osteopathie zu widmen.

Die **Ausschlusskriterien** wären:

- Minderjährigkeit
- Akute Verletzungen/Traumata/ Operationen im Bereich des Unterkiefers und Schädelbereichs, die weniger als 6 Wochen zurück liegen oder nicht ärztlich abgeklärt wurden
- Zustände/Beschwerden, die eine Rückenlage nicht zulassen

Wenn keiner dieser oben genannten Ausschlusskriterien auf Sie zutrifft, würde ich mich sehr über Ihre verbindliche Anmeldung freuen.

Die Studiendurchführung wird am **14. und 15.07.2023** am Oberen Bahnweg 14 in 8523 Frauental zwischen 8:00 und 15:00 stattfinden.

Wichtig: Bei der Studie handelt es sich um eine reine Untersuchung und keine Behandlung bzw. therapeutische Intervention des Unterkiefers.

Anmeldung:

Bis 08.07.2023 per Mail an e.nauschnegg@gmx.at oder telefonisch/per SMS unter der Nummer 0680/32 72 138.

Nach der Anmeldung folgen weitere Information zur Studie, sowie die genauen Testzeiten. Sollte es noch Fragen oder Unklarheiten bezüglich der Teilnahme an dieser Studie geben, melden Sie sich bitte bei mir über die oben genannten Kontaktdaten.

Vielen Dank!

Liebe Grüße,

Elisabeth Nauschnegg

ANHANG D

Fragebogen bezüglich Ausschlusskriterien und Einverständniserklärung

Name: _____

Hatten Sie in den vergangenen 6 Wochen **Verletzungen/Operationen/Traumata** in Bereich des Unterkiefers oder Schädels?

ja nein

Wenn ja, welche?

Leiden Sie unter Beschwerden die eine **Rückenlage** nicht zulassen?

ja nein

Wenn ja, welche? _____

Falls Sie eine der Fragen mit JA beantwortet haben, nehmen Sie bitte kurz Kontakt mit
der Studienleitung auf

- Ich bestätige die Richtigkeit meiner angegebenen Daten.
- Ich nehme freiwillig an dieser Studie teil. Es besteht die Möglichkeit, zu jedem Zeitpunkt von der Teilnahme an der Studie, ohne Angaben von Gründen zurückzutreten.
- Ich bin damit einverstanden, dass meine Daten in anonymisierter Form gespeichert und ausgewertet/weiterverwendet werden.
- Ich bestätige, ausreichend über Inhalt, Ablauf, Risiken und Ziele der Studie aufgeklärt worden zu sein.

Ort, Datum: _____ Unterschrift: _____

ANHANG E

Studieninformation

Thema: Interrater-Reliabilität des Induktionstest Os mandibulare mit der Forschungsfrage: „Wie hoch ist die Übereinstimmung der Ergebnisse zweier Osteopathinnen mit gleichem Ausbildungsstand bei der passiven Untersuchung der Beweglichkeit der Mandibula in Außenrotation und Innenrotation mittels Induktionstest Os mandibulare?“

Studienleiterin: Elisabeth Nauschnegg

Zeitpunkt: 14.07.2023 und 15.07.2023

Ort: Oberer Bahnweg 14, 8523 Frauental an der Lassnitz

Kurzbeschreibung:

Im Zuge meiner Masterthesis möchte ich die Verlässlichkeit eines manualtherapeutischen Tests untersuchen. Ziel ist es herauszufinden, wie hoch die Übereinstimmung der Ergebnisse von zwei TesterInnen bei der Durchführung desselben Tests ist.

Bei der Untersuchung „**Testung der Mobilität des Unterkiefers**“ geht es darum, mögliche Immobilitäten des Unterkiefers im Seitenvergleich festzustellen.

Die Untersuchung stellt ein gängiges Diagnose-Verfahren in der Osteopathie dar. Bei der Studie handelt es sich um eine reine Untersuchung und keine Behandlung bzw. therapeutische Intervention.

Organisatorisches:

Die Untersuchung wird von den Testerinnen ausschließlich mit den Händen mittels Palpation („spüren“) durchgeführt. Die TeilnehmerInnen werden gebeten sich in Rückenlage auf der Behandlungsliege hinzulegen. Für die Untersuchung selbst ist es nicht notwendig, dass sich die TeilnehmerInnen ausziehen. Die Testung wird zweimal gleich hintereinander von den Testerinnen durchgeführt.

Um die Studienergebnisse nicht zu verfälschen, sollten die Gespräche während der Testung auf ein Minimum reduziert werden. Ich bitte um Ihr Verständnis.

Ich bitte Sie, am Testtag pünktlich zu erscheinen, da sie vorab einen Fragebogen und die Einverständniserklärung ausfüllen werden.

Nebenwirkungen:

Da es sich bei der Testung um keine osteopathische Behandlung handelt, sind keinerlei Nebenwirkungen zu erwarten.

Datenschutz:

Ihre Daten werden selbstverständlich vertraulich behandelt und in anonymisierter Form gespeichert und weiterverarbeitet.

Sonstiges:

Sollte es noch Fragen oder Unklarheiten bezüglich der Teilnahme an dieser Studie geben, melden Sie sich bitte per Mail an e.nauschnegg@gmx.at oder telefonisch unter 0680/ 32 72 138.

Vielen lieben Dank für die Bereitschaft zur Teilnahme an dieser Studie.

Elisabeth Nauschnegg

ANHANG F

Teilnehmerfragebogen bezüglich Symptomatik

Probandennummer: _____

Alter: _____

weiblich

männlich

Bei mir treten fallweise folgende Symptome auf:

➤ **Schmerzen im Bereich des Kiefergelenks?**

nein

ja

➤ **Hörbare bzw. fühlbare Geräusche im Kieferbereich?**

nein

ja

➤ **Zähneknirschen?**

nein

ja

➤ **Kopfschmerzen (1x Woche, häufiger)?**

nein

ja

➤ **Verspannung der Kaumuskulatur?**

nein

ja

➤ **Zahnschmerzen (ohne erkennbare zahnmedizinische Ursache)**

nein

ja

➤ **Schmerzen im Gesicht/Trigeminusneuralgie**

nein

ja

➤ **Bewegungseinschränkung oder Verspannungen im oberen Nackenbereich oder an der Schädelbasis?**

nein

ja

ANHANG G

Beurteilungsbogen

Probandennummer: _____

Tester: 2

Induktionstest Os mandibulare:

unauffällig: o

auffällig:

rechts o

links o