

**Diagnostische Validität der manuellen Provokationstests für  
eine Handgelenksinstabilität: systematische  
Literaturübersicht und Meta-Analyse**

---

**VALÉRIE GOY**

**Studentin HES – Studiengang Physiotherapie**

**VÉRONIQUE BLÉTRY**

**Studentin HES – Studiengang Physiotherapie**

**Unter Betreuung von: ROGER HILFIKER**

---

**BACHELORTHESIS ZUR ERREICHUNG DES FH-DIPLOMS ALS  
BACHELOR OF SCIENCE HES-SO IN PHYSIOTHERAPIE**

**JULI 2014**

# Abstrakt

**Einleitung:** Das Ziel dieser systematischen Review ist, verschiedene manuelle Provokationstests für die Diagnose einer Handgelenksinstabilität zu beschreiben und deren diagnostische Validität zu bestimmen.

**Methode:** Verschiedene Datenbanken wurden durchsucht und die Ergebnisse systematisch mit Hilfe von Ein- und Ausschlusskriterien sortiert. Die für unsere Review eingeschlossenen Studien wurden mit der QUADAS-2 auf ihre Bias-Risiken geprüft und die Daten in einer Datenextraktionstabelle zusammengefasst. Studien über identische Tests wurden zusammen gepoolt. Die Validität der Tests wurde mithilfe der Sensitivität und Spezifität ausgedrückt.

**Resultate:** Sechs Studien, die insgesamt zehn Provokationstests untersuchen, wurden in die Review eingeschlossen. Diese zehn Indextests wurden mit der Arthroskopie verglichen. Eine Meta-Analyse von vier der zehn Indextests wurde ausgearbeitet.

Folgende vier Tests konnten gepoolt werden und deren Sensitivität und Spezifität beträgt: Scaphoid Shift Test: 77%, respektive 64%; LT-Ballottement Test: 59% und 57%; TFCC Test: 69% und 60%; DRUJ Test: 84% und 69%. Folgende Tests konnten nicht gepoolt werden. Ihre Sensitivität und Spezifität beträgt: UMTDG: 64% und 47%; Midcarpal Test: 100% und 59%; Ulnar Fovea Sign: 95% und 87%; SL-Shear Test: 100% und 95%; Press Test: 100% und 94%; Testabfolge von Christodoulou: 77% und 95%.

**Analyse:** Der Scaphoid-Shift Test, der DRUJ-Test, der Midcarpal-Test sowie das Ulnar Fovea Sign haben eine gute sensitive Aussagekraft von 77%, 84%, 100%, respektive 95%. Der SL-Shear Test weist sowohl eine gute Sensitivität wie eine gute Spezifität von 100%, respektive 95% auf.

**Diskussion:** Praktiker können diese Tests für diagnostische Untersuchungen nutzen. Die Aussagen sollten jedoch mit Vorbehalt an der jeweiligen Zielgruppe angewendet werden.

# Abstract (Français)

**Introduction:** L'objectif de cette revue systématique diagnostique est d'analyser la validité des tests manuels provocatifs qui détectent une instabilité de poignet.

**Méthode :** Différentes banques de données ont été explorées à l'aide de nos critères d'inclusion et d'exclusion. Les études sélectionnées furent analysées avec l'échelle QUADAS-2 concernant leur risque de biais et les données importantes reportées dans un tableau d'extraction de données. Les études évaluant un même test ont été poolées. La validité des tests est exprimée par leur sensibilité et spécificité.

**Résultats :** Six études évaluant un total de dix tests de provocation manuelle ont été inclus. Ces tests ont été comparés avec l'arthroscopie. Une méta-analyse de quatre de ces tests a pu être conduite. Il s'agit des tests suivants avec leur sensibilité et spécificité respective : Scaphoid Shift Test: 77% et 64%; LT-Ballotement Test: 59% et 57%; TFCC Test: 69% et 60%; DRUJ Test: 84% et 69%. Les tests suivants accompagnés de leurs sensibilité et spécificité respectivement n'ont pas pu être poolés : UMTDG: 64% et 47%; Midcarpal Test: 100% et 59%; Ulnar fovea sign: 95% et 87%; SL-Shear Test: 100% et 95%; Press Test: 100% et 94%; Suite de tests de Christodoulou: 77% et 95%.

**Analyse :** Le Scaphoid-Shift Test, le DRUJ-Test, le Midcarpal-Test ainsi que l'Ulna Fovea sign présentent une bonne sensibilité de respectivement 77%, 84%, 100% et 95%. Le SL-Shear test démontre une bonne sensibilité de 100% ainsi qu'une bonne spécificité de 95%.

**Discussion :** Les praticiens peuvent faire usage de ces tests dans un but diagnostique. Les résultats présentés dans cette étude sont tout de même à considérer avec prudence et à adapter au public cible.

# Danksagung

Zuerst möchten wir herzlich unserem Bachelorarbeitsbetreuer, Roger Hilfiker danken, für seine Geduld und Unterstützung und der bereitwillig unsere zahlreichen Fragen beantwortete. Ausserdem danken wir den Dozenten der HES-SO Wallis, Anne-Gabrielle Mittaz Hager und Martin Sattelmayer.

Des Weiteren möchten wir uns bei Loïc Bassang und Jonathan Cibran bedanken, die uns mit ihren Informatikkenntnissen und Geduld zur Seite standen.

Herzlich bedanken möchten wir uns auch bei den fleissigen Lesern, die unsere Bachelorarbeit nachgelesen und korrigiert haben: Ursula Goy und Cosette Blétry.

## Hinweis/Eigenständigkeitserklärung:

Die Verantwortung für den Inhalt, die Argumentationen und die Schlussfolgerung dieser Arbeit liegt ausschliesslich bei den Autoren und in keinem Fall bei der Fachhochschule für Gesundheit Wallis, der Jury oder dem Betreuer der Bachelorarbeit. Wir bezeugen, die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen benutzt zu haben.

Datum und Name der Autoren:

05. Juni 2014	Goy Valérie	..... Valérie Goy
05. Juni 2014	Blétry Véronique	..... V. Blétry

# Inhaltsverzeichnis

<b>1. Einführung</b> .....	<b>1</b>
<b>1.1 Problematik</b> .....	<b>1</b>
<b>1.2 Anatomie und Pathophysiologie</b> .....	<b>1</b>
<b>1.3 Instabilitätstypen</b> .....	<b>3</b>
<b>1.4 Risikofaktoren</b> .....	<b>4</b>
<b>1.5 Diagnostik</b> .....	<b>4</b>
1.5.1 Subjektiver und objektiver Befund .....	4
1.5.2 Indextests.....	5
1.5.2.1 Der Scaphoid Shift Test (SST) .....	5
1.5.2.2 Der Lunotriquetral Test (LT-Test).....	6
1.5.2.3 Der Triangular Fibrocartilage Complex Stress Test (TFCC Stress Test) (with or without compression) .....	6
1.5.2.4 Der Distal Radio-Ulnar Joint Test (DRUJ Test) .....	7
1.5.2.5 Der Ulnomeniscotriquetral Dorsal Glide Test (UMTDG) .....	7
1.5.2.6 Der Midcarpal Test.....	8
1.5.2.7 Das Ulnar fovea sign.....	8
1.5.2.8 Der Scapholunate Shear Test (SL-Shear Test).....	9
1.5.2.9 Der Press-Test.....	9
1.5.2.10 Die Christodoulou Testabfolge .....	9
1.5.3 Referenztests.....	10
<b>1.6 Folgen einer unbehandelten Instabilität</b> .....	<b>11</b>
<b>1.7 Behandlung</b> .....	<b>11</b>
<b>1.8 Relevanz der Review</b> .....	<b>12</b>
<b>1.9 Ziel der Review</b> .....	<b>12</b>
<b>2. Methode</b> .....	<b>13</b>
<b>2.1 Einschlusskriterien</b> .....	<b>13</b>
<b>2.2 Ausschlusskriterien</b> .....	<b>14</b>
<b>2.3 Suchstrategie</b> .....	<b>14</b>
<b>2.4 Suche auf Datenbanken</b> .....	<b>16</b>
<b>2.5 Datensammlung und Analyse</b> .....	<b>16</b>
2.5.1 Sortieren nach Titel und Abstrakt .....	16
2.5.2 Sortieren nach Volltext.....	16
2.5.3 Lesetabelle.....	16
<b>2.6 Quersuche</b> .....	<b>17</b>

<b>2.7</b>	<b>Suche der Volltexte .....</b>	<b>17</b>
<b>2.8</b>	<b>Datenextraktionstabelle.....</b>	<b>17</b>
<b>2.9</b>	<b>Studienqualität mit der QUADAS-2.....</b>	<b>18</b>
<b>2.10</b>	<b>Analyse .....</b>	<b>18</b>
2.10.1	Sensitivität und Spezifität.....	18
2.10.2	Validität.....	20
2.10.3	Begriffe zur diagnostischen Analyse.....	20
2.10.3.1	Die Prävalenz.....	20
2.10.3.2	Das Konfidenzintervall (CI).....	20
2.10.3.3	Die Heterogenität und der I <sup>2</sup> -Wert: .....	20
2.10.3.4	Der p-Wert.....	21
2.10.3.5	Die Likelihood Ratio .....	21
2.10.4	Auswertung.....	21
2.10.5	Arbeitsprogramme:.....	22
<b>3.</b>	<b>Resultate.....</b>	<b>23</b>
<b>3.1</b>	<b>Resultate der Suche auf Datenbanken.....</b>	<b>23</b>
<b>3.2</b>	<b>Resultate nach Ausschluss nach Titel .....</b>	<b>23</b>
<b>3.3</b>	<b>Resultate nach Ausschluss nach Abstrakte.....</b>	<b>24</b>
<b>3.4</b>	<b>Resultate nach Ausschluss nach Volltexte.....</b>	<b>24</b>
<b>3.5</b>	<b>Resultate der Quersuche .....</b>	<b>24</b>
<b>3.6</b>	<b>Flussdiagramm der Suche .....</b>	<b>25</b>
<b>3.7</b>	<b>Qualität der Studien mit der QUADAS-2.....</b>	<b>25</b>
3.7.1	Bias-Risiko.....	25
3.7.2	Anwendbarkeit.....	26
<b>3.8</b>	<b>Resultate Datenextraktion der Ergebnisse.....</b>	<b>28</b>
<b>3.9</b>	<b>Meta-Analyse.....</b>	<b>29</b>
3.9.1	Scaphoid Shift Test.....	29
3.9.2	LT-Ballottement Test.....	30
3.9.3	TFCC Test.....	31
3.9.4	DRUJ Test.....	32
3.9.5	Beschreibende Analysen der Studien, die nicht gepoolt werden konnten .....	33
<b>4.</b>	<b>Diskussion .....</b>	<b>35</b>
<b>4.1</b>	<b>Fazit.....</b>	<b>35</b>
<b>4.2</b>	<b>Auszeichnung dieser Studie .....</b>	<b>35</b>
<b>4.3</b>	<b>Limitationen dieser Studie.....</b>	<b>37</b>
<b>4.4</b>	<b>Vergleich mir ähnliche Studien .....</b>	<b>38</b>

4.5	Auswirkungen für die Forschung .....	39
4.6	Auswirkungen für die Praxis.....	40
5.	Schlussfolgerung .....	42
6.	Literaturverzeichnis .....	43
7.	Abbildungsverzeichnis.....	47
8.	Tabellenverzeichnis.....	47
9.	Anhang.....	I
9.1	Detaillierte QUADAS-2 Tabellen .....	I-V

# 1. Einführung

## 1.1 Problematik

Eine eindeutige Diagnostik einer Handgelenkinstabilität ist in der heutigen Zeit sehr schwierig: Einerseits weil eine Instabilität sowohl isoliert wie auch in Kombination mit anderen Pathologien vorkommen kann und dadurch oft übersehen wird (Trail, Stanley, & Hayton, 2007). Andererseits auch, weil eine Instabilität über Jahre hinweg asymptomatisch bleiben kann (Van Schoonhoven, Prommersberger, Krimmer, & Lanz, 1999). Hinzu kommt, dass der Auslösemechanismus, der zu einer Instabilität führt, nicht immer derselbe ist, und die Anatomie des Handgelenkes sowie dessen Kinematik sehr komplex sind (Carlsen & Shin, 2008). Daher liegen keine absoluten Zahlen über die Prävalenz der Handgelenkinstabilitäten vor.

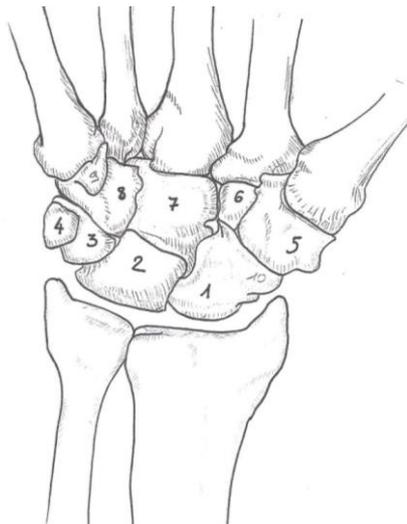
Diagnostik ist ein wichtiger Bestandteil in der Praxis der modernen Physiotherapie, auch wenn offiziell die Aufgabe der Diagnose bei den Medizinern liegt; dies, weil wir immer öfters Patienten sehen, die nicht zum Arzt gehen, weil sie z.B. eine hohe Franchise bei ihrer Krankenkasse haben. Eine richtige, frühzeitige Diagnose ist grundlegend für eine angepasste Behandlung. Frische Instabilitäten lassen sich meist durch einfache 4-6-wöchige Ruhigstellung ausheilen. Eine chronische Instabilität hingegen muss meistens operiert werden (siehe Kapitel 1.7 Behandlung) (Van Schoonhoven et al., 1999).

## 1.2 Anatomie und Pathophysiologie

Das Handgelenk verbindet den Vorderarm mit der Hand und erlaubt mit seinem komplexen Aufbau ein Maximum an Bewegungsmöglichkeiten (Carlsen & Shin, 2008). In der Bezeichnung "Handgelenk" werden 15 Knochen (Radius, Ulna, proximale (4) und distale (4) karpale Reihe, 5 Metakarpalknochen) und über 20 Gelenke mit einbezogen. Letztere werden in vier Hauptgelenke eingeteilt: Das distale radioulnare Gelenk, das radiokarpale Gelenk, das mediokarpale Gelenk und das karpometakarpale Gelenk. Das radiokarpale Gelenk erlaubt dem Handgelenk trotz des eingeschränkten Bewegungsraumes drei Bewegungskomponente: die Dorsalextension - Palmarflexion, die Pronation - Supination sowie die radiale und ulnare Deviation (Carlsen & Shin, 2008; Dumontier, 1996).

Die relative Stabilität, die dennoch sehr viel Mobilität zulässt, erfordert eine extreme Koordination zwischen Bändern, Muskeln und Sehnen (Forman, Forman, & Rose, 2005). Da sich, mit Ausnahme vom Muskel flexor carpi ulnaris keine Sehne an den Karpuskomplex ansetzt, hängt die Stabilität des Handgelenkes nur von der Knochenform und der Integrität seiner Bänder ab.

Sobald das Handgelenk belastet wird und muskuläre Spannung den Karpus durchläuft, schränken die Bänder die Bewegungen zwischen den einzelnen Knochen, und Knochengruppen durch eine wechselseitige Spannung ein (Carlsen & Shin, 2008).



**Abbildung 1: Karpus von palmar (Bild: Autoren)**

**Legende:**

1. Scaphoidknochen
2. Lunatumknochen
3. Triquetralknochen
4. Pisiformisknochen
5. Trapeziumknochen
6. Trapezoideumknochen
7. Capitatumknochen
8. Hamatumknochen

Die Bänder des Handgelenkes werden in zwei Gruppen eingeteilt: die extrinsischen und intrinsischen Bänder. Die Extrinsischen verbinden den Karpus mit dem Vorderarm, d.h. mit dem Radius und der Ulna, die intrinsischen verbinden die Karpalknochen untereinander.

Alle extrinsischen Bänder befinden sich auf der palmaren Seite: das ligamentum radioscaphocapitate, radiolunate longus und brevis, ulnolunate, ulnocapitate und ulnotriquetral-capitate.

Auf der dorsalen Seite befindet sich nur ein extrinsisches Band: das lig. dorsale radiocarpal oder auch lig. radiolunotriquetral genannt.

Bei den intrinsischen Bändern sind die zwei wichtigsten das lig. scapholunate (SL) und das lig. lunotriquetral (LT). Beide haben dorsale und palmare Anteile. Dabei ist das SL-Band das Stärkste auf der dorsalen, das LT-Band auf der palmaren Seite (Ankarath, 2006; Carlsen & Shin, 2008; Schmitt, Froehner, Coblenz, & Christopoulos, 2006; Trail et al., 2007).

Ein instabiles Handgelenk kann verschiedene Ursachen haben: Traumen (überwiegend Sturz auf das Handgelenk in Extension), lifting-and-twisting-Manöver, Handgelenksfrakturen, Überbeanspruchung des Handgelenkes, Osteoarthritis,

degenerative Veränderungen mit dem Älterwerden, repetitive mikro-Traumen oder Hyperlaxität des Gelenkes. Die Ursachen sind jedoch in ca. 27% der Fälle unbekannt (Pliefke et al., 2008; Prosser, Herbert, & LaStayo, 2007; Trail et al., 2007).

Ein kleiner Einriss des Scapholunatum-Bandes resultiert nicht zwangsläufig in einer Instabilität, bewirkt aber oft chronische Handgelenksschmerzen, begleitet von einer Synovitis und Ganglien. Eine grössere Bandläsion oder kompletter Bandriss führt zu einer Instabilität (Parmelee-Peters & Eathorne, 2005). Die verschiedenen auftretenden Instabilitäten werden im nächsten Kapitel beschrieben.

Das "Anatomy and Biomechanics Committee of the International Federation of Societies for Surgery of the Hand" kam zum Schluss, dass "ein Handgelenk nur als klinisch instabil erachtet werden sollte, wenn es eine symptomatische Dysfunktion aufweist, keine Belastung ertragen kann und keine normale Kinematik während einer Bewegung in der normalen Bewegungsamplitude aufweist" (Garcia-Elias, 1999).

Für das weitere Verständnis unserer Studie möchten wir in diesem Abschnitt zwei anatomische Ausdrücke, den Triangular Fibrocartilage Komplex (TFCC) und das distale radioulnar Gelenk (DRUJ) genauer definieren.

Der TFCC ist ein Komplex bestehend aus dem semizirkulären bikonkaven Fibrocartilage (TFC), sowie dem palmaren und dorsalen distalen radioulnar Band, dem ulnolunate Band, dem ulnotriquetral Band und dem ulnocarpal Meniscus homologue.

Das DRUJ wird auch durch mehrere Strukturen stabilisiert. Diese sind der TFC, der ulnokarpal Bandkomplex, die membrana interossea, der Muskel pronator quadratus und das Retinakulum infratendinosus extensor (Bednar et al., 1991 und Palmer et al., 1989, zitiert nach Parmelee-Peters & Eathorne, 2005).

## 1.3 Instabilitätstypen

Instabilitäten wurden schon seit den 1980er Jahren in statische und dynamische Instabilitäten eingeordnet. Die statische Instabilität ist ein endgültiger Zustand, da sich dieser nicht mehr zurückbildet und deren Deformationen auf Röntgenaufnahmen gut sichtbar sind. Die dynamische Instabilität hingegen muss unter Bewegung nachgewiesen werden, mit Hilfe einer dynamischen Röntgenaufnahme oder einer dynamischen Arthroskopie (Trail et al., 2007).

Die derzeit bekannteste Einteilung zur Beschreibung der verschiedenen Instabilitätsmuster ist die "*Mayo Clinical Classification*". Sie teilt die Handgelenkinstabilitäten in vier Muster auf: Dissoziative Instabilität (carpal instability

dissociative, CID), non-dissoziative Instabilität (carpal instability non-dissociative, CIND), kombiniertes Muster (carpal instability complex, CIC), sowie adaptiver Karpus. Die dissoziative Instabilität beschreibt ein betroffenes Band innerhalb einer karpalen Reihe, entweder der Proximalen oder Distalen, wo hingegen die Non-Dissoziative (CIND) ein Band zwischen zwei karpalen Reihen betrifft. Das kombinierte Muster (CIC) ist, wie der Name besagt, eine Kombination aus CID und CIND. Der adaptive Karpus ist eine sekundäre Fehlstellung der Karpalgelenke, verursacht durch extrakarpale Pathologien (Carlsen & Shin, 2008; Schmitt et al., 2006; Trail et al., 2007).

## **1.4 Risikofaktoren**

Bislang wurde vor allem das isolierte Trauma als Ursache erwähnt. Doch eine Instabilität tritt auch oft als Folge der rheumatischen Arthritis auf - bis zu 36% gemäss der Studie von Kushner et al. (1993) (zitiert nach Trail et al., 2007). Sie kann auch durch wiederholte kleinere Traumen hervorgerufen werden, wie im Fall von Paraplegikern, die sich viel mit den Händen abstützen müssen (Schrorer et al., 1996, zitiert nach Trail et al., 2007). Allerdings haben wir zu diesem Thema nur wenige Studien gefunden. Bei Patienten mit erhöhter Gelenkbeweglichkeit, d.h. Laxität, wird das Handgelenk schon bei kleinen Traumen problematisch (Garcia-Elias et al., 1995 und Saffar et al., 1998, zitiert nach Trail et al., 2007). Aus diesem Grund haben wir Studien mit Kindern ausgeschlossen, da sie über eine sehr grosse Laxität verfügen (Cameriere, Ferrante, Mirtella, & Cingolani, 2006).

## **1.5 Diagnostik**

### **1.5.1 Subjektiver und objektiver Befund**

Wenn ein Patient mit einem problematischen Handgelenk zur Behandlung kommt, wird dieses rundum angeschaut, abgetastet und durchbewegt. Dem Patienten werden Fragen gestellt, die zur Anamnese nötig sind. Dies bezeichnen wir als den subjektiven und objektiven Befund. Zum ersten Befund gehört eindeutig die ganze Patientengeschichte: Wann und wie die Beschwerden aufgetreten sind; falls diese durch einen Sturz verursacht wurden, möchten wir wissen, in welcher Stellung das Handgelenk beim Aufprall war; bei Schmerzen fragen wir nach dessen Ortung, Auslöser, Dauer, Beschreibung des Empfindens, Ausmass und Verlauf seit der Erscheinung, ob dem

Patienten ein spezielles Geräusch aufgefallen ist und ob ihn sein Handgelenk in einer Alltagsaktivität einschränkt. Diese Art der Befragung kommt bei den meisten Autoren vor (Dumontier, 1996; Prosser et al., 2007; Schmitt et al., 2006; Trail et al., 2007).

Einig sind sich die Autoren ebenfalls was den objektiven Befund angeht. Er soll die folgenden Anhaltspunkte möglichst einbeziehen: kutane Inspektion, neurovaskuläre Veränderungen, örtliche Schwellung, aktive und passive Mobilisation, Greifkraft, lokaler Schmerzpunkt, ungewöhnliche interkarpale Bewegungen, ungewöhnliche Geräusche (wie "Klick", "Klunk" oder Knistern) beim Mobilisieren. Dabei wird das problematische Handgelenk mit dem gesunden, kontrolateralen verglichen (Dumontier, 1996; Prosser et al., 2007; Trail et al., 2007).

## 1.5.2 Indextests

### 1.5.2.1 Der Scaphoid Shift Test (SST)



Abbildung 2: Der Scaphoid Shift Test (Bild: Autoren)

Der Test wird allgemein auch **Watson's Test** genannt.

Bei diesem Test sitzt der Therapeut dem Patienten gegenüber und hält die betroffene Hand mit seiner eigenen, gleichseitigen Hand wie beim Armdrücken. Beide Ellbogen ruhen auf dem Tisch. Der Vorderarm des Patienten wird in leichter Pronation positioniert und vom Therapeut von der radialen Seite her mit dem Daumen auf dem palmaren "Tuberkulum scaphoideum" umfasst, die Finger greifen um den distalen Radius. Diese Stellung erlaubt es, einen fixen Punkt (Radius) zu haben, um den Scaphoidknochen nach

dorsal zu mobilisieren. Die zweite Hand des Therapeuten wird um die metacarpale Reihe gelegt, um die Handgelenksposition zu kontrollieren.

Das Manöver wird folgendermassen durchgeführt: Mit konstantem Druck des Daumens auf den Scaphoidknochen wird das Handgelenk des Patienten mit der kontrollierenden Hand passiv erst in leichte Extension und ulnare Deviation gestellt und danach in leichte Flexion mit radialer Deviation geführt. In dieser Endstellung wird der Druck des Daumens langsam vermindert. Es ist wichtig, beide Handgelenke auf die gleiche Weise zu testen um die Zeichen besser interpretieren zu können. Der Scaphoidknochen dreht sich normalerweise bei dieser Bewegungsabfolge und steht an der palmaren Seite weiter vor. Der Druck der vom Daumen ausgeübt wird, gleicht diese Drehung aus und bewirkt

so einen Subluxationsstress nach dorsal, der das mehr oder weniger starke Heraustreten des Knochens bewirkt.

Dieser Test dient dazu, eine Läsion des scapholunatum-Bandes zu erkennen und wird als positiv bezeichnet, wenn ein asymmetrischer Befund links und rechts beim Patient festgestellt wird, der keine sonstigen einseitigen Probleme oder traumatische Ereignisse erkennen lässt. Ein Klicken beim Nachlassen des Daumendruckes und die Reproduktion des bekannten Schmerzes deuten ebenfalls auf ein positives Testergebnis hin (LaStayo & Howell, 1995; Watson, Ashmead 4th, & Makhlof, 1988).

### 1.5.2.2 Der Lunotriquetral Test (LT-Test)



Abbildung 3: Der Lunotriquetral Test (Bild: Autoren)

Dieser Test wird in der Literatur auch **Ballotement Test** genannt. Der Therapeut fixiert mit der gleichseitigen Hand wie das problematische Handgelenk des Patienten mit dem Daumen von dorsal und den Zeigefinger von palmar den Lunatumknochen. Mit seiner anderen Hand greift der Therapeut den Triquetrum-Pisiform-Komplex und kann diesen somit in dorsaler und palmarer

Richtung bewegen, wobei er das lunotriquetral-Band testet. Der Test ist positiv, wenn die Palpation des lunotriquetral-Gelenkes schmerzhaft ist, es ein knisterndes Geräusch ergibt oder das Bewegungsausmaß abnormal zur anderen Seite erhöht oder vermindert ist (LaStayo & Howell, 1995; Watson et al., 1988).

### 1.5.2.3 Der Triangular Fibrocartilage Complex Stress Test (TFCC Stress Test) (with or without compression)



Abbildung 4: Der Triangular Fibrocartilage Complex Stress Test (Bild: Autoren)

Das Handgelenk wird in ulnare Deviation gestellt und repetitive Flexions-Extensions-Bewegungen werden auf den Ulnarkomplex appliziert. Dasselbe wird mit zusätzlichem axialem Druck durchgeführt, um eine Kompression zu bewirken. Wenn eines der beiden Manöver positiv ist wird der Test für den TFCC positiv bewertet. Positiv gilt, wenn die Schmerzen reproduziert werden oder ein "Klick" ertönt (Hodgson, Royle, & Stanley, 1995; Parmelee-Peters & Eathorne, 2005; Prosser et al., 2011).

#### 1.5.2.4 Der Distal Radio-Ulnar Joint Test (DRUJ Test)



Abbildung 5: Der Distal Radio-Ulnar Joint Test (Bild: Autoren)

Der DRUJ-Test, dem man in der Literatur auch unter dem Namen **Kleinman's Test** begegnet, untersucht die Integrität des dorsalen und palmaren DRUJ-Bandes. Dabei wird die Ulna in je der neutralen, Supinations- und Pronations-Stellung in die maximale dorsale und palmare Position geleitet. Der Test wird als positiv angesehen, wenn die Schmerzen des Patienten wiederhergestellt werden (mit oder ohne Hyperlaxität, verglichen mit der gesunden Seite) (Hodgson et al., 1995; Prosser et al., 2011; Spies, Müller, Oppermann, Hahn, & Unglaub, 2014; Valdes & LaStayo, 2013).

#### 1.5.2.5 Der Ulnomeniscotriquetral Dorsal Glide Test (UMTDG)



Abbildung 6: Der Ulnomeniscotriquetral Dorsal Glide Test (Bild: Autoren)

Der Test wird in der Literatur auch **TFCC Shear Test** genannt. Dabei sitzt oder liegt der Patient, sein Ellbogen ist auf dem Tisch aufgestützt und der Vorderarm wird in neutraler Pro-/Supinations-Position vertikal gehalten. Der Therapeut legt mit der gegenseitigen Hand wie das problematische Handgelenk des Patienten seinen Daumen von dorsal auf das Ulnaköpfchen und greift mit dem Zeigefinger von palmar den Pisotriquetral-Komplex. Mit einer zusammendrückenden Bewegung von diesen beiden Fingern übt der Therapeut eine dorsale Gleitkraft des Pisotriquetral-Komplexes im Bezug auf das Ulnaköpfchen und des dazwischenliegenden Meniskus aus. Der Test ist positiv, wenn er die bekannten Symptome wieder hervorruft oder das Bewegungsausmass abnormal zur anderen Seite erhöht ist (LaStayo & Howell, 1995; Spies et al., 2014).

### 1.5.2.6 Der Midcarpal Test



Abbildung 7: Der Midcarpal Test (Bild: Autoren)

Dieser Test, der auch **Lichtman's Test** genannt wird, prüft die Integrität des Arcuate-Ligamentum (auch Deltoïd-Band oder V-Band genannt). Der Therapeut legt seinen Daumen der gleichseitigen Hand über die dorsale distale karpale Reihe des zu testenden Handgelenkes. Er positioniert das Handgelenk zuerst in Pronation und leichte ulnare Deviation ( $15^\circ$ ). Dann drückt er mit dem Daumen die distale karpale Reihe nach palmar. Diesen Druck haltend bringt der Therapeut das Handgelenk weiter in ulnare Deviation. Der Test wird positiv gedeutet, wenn ein "Catch-up Clunk" (ein Repositionsgeräusch) die bekannten Schmerzen des Patienten begleitet. Dies mit oder ohne Hypermobilität im Vergleich zur Gegenseite (Dumontier, 1996; Prosser et al., 2011; Valdes & LaStayo, 2013).

### 1.5.2.7 Das Ulnar fovea sign



Abbildung 8: Das Ulnar fovea sign (Bild: Autoren)

Der Patient sitzt dem Therapeuten gegenüber, den Ellbogen des pathologischen Handgelenkes auf dem Tisch abgestützt, das Handgelenk in einer neutralen Position. Der Therapeut drückt mit seinem Daumen in die Lücke zwischen dem Pisiformis, dem Ulnarköpfchen und der Sehne des Flexor Carpi Ulnaris. Der Test wird als positiv angesehen, wenn die Schmerzen des Patienten wieder hervorgerufen werden (im Seitenvergleich) (Tay, Tomita, & Berger, 2007; Valdes & LaStayo, 2013).

### 1.5.2.8 Der Scapholunate Shear Test (SL-Shear Test)



Abbildung 9: Der Scapholunate Shear Test (Bild: Autoren)

erhöhte Beweglichkeit vorhanden ist oder die Schmerzen des Patienten reproduziert werden (Hodgson et al., 1995).

Das zu testende Handgelenk bleibt in Neutralstellung. Der Therapeut greift mit Daumen und Zeigefinger den Scaphoid-Knochen und den Lunatum-Knochen. Er drückt dann von dorsal auf den Lunatum-Knochen und von palmar her auf das Tuberculum des Scaphoid. Dabei entstehen Scherkräfte, die das SL-Band spannen. Auch hier wird der Test als positiv angesehen, wenn eine

### 1.5.2.9 Der Press-Test



Abbildung 10: Der Press-Test (Bild: Autoren)

Der Patient sitzt auf einem Stuhl, mit oder ohne Armlehne. Er greift die Armlehnen oder die Sitzfläche beidseitig und versucht, sich auf seinen Händen aufzustützen und somit sein Körpergewicht hochzustemmen. Der Test gilt als positiv, wenn lokal Schmerzen, die dem Patient vertraut sind, auf der ulnaren Seite des Handgelenkes hervorgerufen werden (Lester, Halbrecht, Levy, & Gaudinez, 1995; Spies et al., 2014; Valdes & LaStayo, 2013).

### 1.5.2.10 Die Christodoulou Testabfolge



Abbildung 11: Erster Test der Christodoulou Testabfolge (Bild: Autoren)

Die Testreihe von drei aufeinanderfolgenden Provokationstests soll eine peritriquetrale Instabilität testen. Der **erste Test** soll den Triquetrumknochen wieder an seinen Platz einsetzen. Der Ellbogen des Patienten wird auf dem Tisch abgestützt und der Vorderarm in maximale Pronationsstellung gebracht. Das Handgelenk wird in die maximale Dorsalflexion und radiale Deviation gestellt. Der Therapeut drückt mit seinem Daumen auf die palmare Seite des Pisiformis und gibt gleichzeitig Gegendruck am distalen Ende der Ulna. Jetzt wird das Handgelenk wieder in die

neutrale Position gebracht. Oft ist dabei in Klicken zu hören. Der Patient soll danach eine Faust machen. Hat er mehr Kraft als vor dem Test und hat sein Instabilitätsgefühl nachgelassen, so wird der Test als positiv gewertet.



Abbildung 12: Zweiter Test der Christodoulou Testabfolge (Bild: Autoren)

Der **zweite Test** wird durchgeführt, sobald die Instabilität wieder vorhanden ist: Das Handgelenk wird in Pronation und radiale Deviation gebracht. Der Therapeut drückt mit seinem Daumen den Triquetrumknochen von dorsal nach palmar und bringt dabei das Handgelenk in ulnare Deviation. Der Test wird als positiv angesehen, wenn sich der Patient über Schmerzen beklagt.



Abbildung 13: Dritter Test der Christodoulou Testabfolge (Bild: Autoren)

Der **dritte Test** ist wie der Zweite, nur umgekehrt: Das Handgelenk wird von der ulnaren in die radiale Deviation gebracht während der Therapeut den Triquetrumknochen (über den Pisiformis) nach dorsal drückt. Auch hier sind die Schmerzen des Patienten für einen positiven Test entscheidend (Christodoulou & Bainbridge, 1999; Valdes & LaStayo, 2013).

### 1.5.3 Referenztests

Als Referenztest - auch Goldstandard bezeichnet - ist die Arthroskopie allgemein anerkannt (Carlsen & Shin, 2008; Schmitt et al., 2006; Trail et al., 2007; Van Schoonhoven et al., 1999). Das arthroskopische Verfahren besteht aus direkter Visualisierung der Strukturen innerhalb des Handgelenkes. Hierfür wird mit Hilfe von Mädchenfingern an den Fingern und der Schwerkraft eine Traktion auf das Handgelenk erzeugt. So ist es dem Chirurg möglich, Mikrokameras und -instrumente durch verschiedene, gut definierte Eingangspunkte eindringen zu lassen (Herron & Craigen, 2001).

## **1.6 Folgen einer unbehandelten Instabilität**

Eine unbehandelte Instabilität des Handgelenks entwickelt sich meistens mit dem Auftreten von Arthrose. Oft wenden sich die Patienten erst dann an einen Arzt, wenn die durch die Instabilität verursachten Arthroseschmerzen auftreten (Engelhardt & Krimmer, 2006; Van Schoonhoven et al., 1999).

Eine Scapholunäre-Instabilität kann zu einem kompletten karpalen Kollaps und Arthrose führen. Damit verliert der Patient seine karpale Höhe und die knöchernen Kongruenz die zur Kraftübertragung unersetzlich ist (Trail et al., 2007; Van Schoonhoven et al., 1999).

## **1.7 Behandlung**

Die Behandlung hängt von vielen Faktoren ab: Von der Qualität des zu reparierenden Gewebes, von der Präsenz arthritischer Veränderungen, von der Chronizität der Verletzung und der Reduzierbarkeit der Deformitäten.

Sofern die Qualität des Gewebes in Ordnung ist und die Veränderungen im Gelenk nicht gravierend sind, sollte eine Ruhigstellung und Wiederherstellung der normalen Anatomie des Handgelenkes ausreichen.

Die Wundheilung eines Bandes beginnt mit der Entzündungsphase, die ungefähr fünf Tage dauert. Danach beginnt die Proliferationsphase, die bis zu sechs Wochen nach der Läsion anhält. Ab der sechsten Woche beginnt die Umbauphase. Bis zur vollständigen Regeneration eines Bandes dauert es 10-16 Monate (Roduner, Baeriswyl, & Stahl, 2013).

Falls die Verletzung jedoch gravierender ist, oder die Läsion nicht ausheilt, sollte eine Operation mit einer Rekonstruktion des Bandes in Betracht gezogen werden. Eine solche Operation wird entweder offen oder geschlossen (Arthroskopie) durchgeführt. Bei grossen Deformationen wird zum Teil sogar eine Arthrodesse gemacht (Carlsen & Shin, 2008; Trail et al., 2007).

## **1.8 Relevanz der Review**

Diagnostik ist heute ein wichtiger Bestandteil in der physiotherapeutischen Praxis: Um dem Patienten eine angemessene Behandlung bieten zu können, muss der Therapeut das vorliegende Problem korrekt beurteilen. Hierfür werden in der Praxis oft manuelle Tests zur Hilfe gezogen, wie der Watson's Test oder Lunotriquetral Ballottement Test, die im Gegensatz zur Arthroskopie oder Radiographie weniger aufwendig und nicht invasiv sind, überall und schnell durchgeführt werden können, sofortige Resultate liefern und vor allem viel weniger kosten (Fritz & Wainner, 2001).

Diagnostik wird auch in Zukunft immer wichtiger werden, da die Physiotherapeuten einen direkten Zugang zu ihren Dienstleistungen anstreben, wobei es ihre Aufgabe sein wird, evidenz-basiert zu praktizieren. Um auf dieser Basis arbeiten zu können, muss der Physiotherapeut die Tests, die Symptome und klinischen Zeichen möglichst genau kennen und interpretieren können, um seinem Patienten eine angepasste Betreuung bieten zu können. Um sich über die Validität seiner Tests im Klaren zu sein, sollte der Physiotherapeut die Sensitivität und Spezifität des jeweiligen Tests kennen.

## **1.9 Ziel der Review**

Das Ziel unserer Review ist, uns eine Übersicht der verschiedenen existierenden nicht-invasiven Tests der Handgelenksinstabilität zu verschaffen. Die diagnostischen Studien darüber werden wir untereinander vergleichen, um die statistisch besten Tests ausfindig zu machen.

Wir suchen die Zuverlässigkeit der manuellen Provokationstests, um im Rahmen der Evidenz-basierten-Wissenschaft eine karpale Instabilität physiotherapeutisch diagnostizieren zu können. Dies im Interesse der frühzeitigen Erkennung und richtigen Diagnostizierung eines problematischen Handgelenkes, damit eine sofortige, angepasste Betreuung geboten und sekundäre Folgen vermieden werden können.

# 2.Methode

## 2.1 Einschlusskriterien

Um unsere systematische Review zu verfassen haben wir Studien eingeschlossen, die ein basic test accuracy- oder ein diagnostic case control-Design haben. Beide erlauben uns die Validität eines Indextests im Vergleich zu einem Referenztest zu bestimmen. Bei basic test accuracy-Studien werden bei gesunden und kranken Patienten zuerst ein Indextest durchgeführt und danach mit einem Referenztest verglichen.

Beim diagnostic case control werden die gesunden und kranken Patienten bereits vor der Durchführung des Indextests identifiziert (Rutjes, Reitsma, Vandenbroucke, Glas, & Bossuty, 2005).

Weiter wurden Studien eingeschlossen, die in englischer, deutscher oder französischer Sprache verfasst wurden. Die jeweiligen Studien mussten auch einen oder mehrere der Indextests analysieren und mit einem der Referenztests vergleichen. Der Artikel musste Daten enthalten mit denen die Sensitivität und die Spezifität berechnet werden kann.

Die Indextests mussten manuell und ohne Hilfsmittel durchgeführt worden sein und eine mögliche Instabilität in einem oder mehreren der folgenden Gelenke testen: distales radioulnares Gelenk, radiokarpales Gelenk, interkarpale Gelenke, sowie Triangular-Fibrocartilage-Komplex (TFCC).

Der Referenztest war im Idealfall die Arthroskopie die als Goldstandart gilt. Falls wir Studien fanden, die einen anderen Referenztest (Radiographie, MRI, Arthrographie, Kinematographie, usw.) verwendeten, haben wir diese trotzdem eingeschlossen. Allerdings akzeptierten wir diese nur unter der Bedingung, dass wir die Zuverlässigkeit der angewandten Methode im Vergleich zur Arthroskopie in einem Artikel fanden, um die Resultate mit diesem Bias trotzdem verarbeiten zu können.

## 2.2 Ausschlusskriterien

Studien, die in anderen Sprachen geschrieben wurden, als in den drei obengenannten, wurden ausgeschlossen. Auch Studien, die an Kadavern durchgeführt worden sind haben wir ausgeschlossen, da die Resultate in der physiotherapeutischen Praxis nicht interpretiert werden können und deshalb nicht relevant sind. Ausserdem haben wir keine Studien eingeschlossen die an Kindern durchgeführt worden sind, da es sich um eine spezifische Population handelt, die nicht mit den anderen Patienten verglichen werden kann.

## 2.3 Suchstrategie

Wir haben die **PICM**-Elemente benutzt, um unsere Suchstrategie zusammen zu stellen. (P=Population, I=Indextest, C=Comparaison = Referenztest zum Vergleich, M=Methodologischer Filter)

Innerhalb der einzelnen PICM-Kategorien wurden die Begriffe mit einem "OR" und die Kategorien untereinander mit einem "AND" verbunden. Einzige Ausnahme war die Kategorie Population, in der wir 2 untergeordnete Kategorien (1. & 2.) ebenfalls mit einem "AND" verbunden haben.

### **Population:**

1. Triangular Fibrocartilage Complex, TFCC, carpal, carpus, carpal joint, wrist joint, wrist, midcarpal, mid-carpal, mid carpal, perilunate, peri-lunate, peri lunate, central midcarpal, central mid-carpal, central mid carpal, ulnar midcarpal, ulnar mid-carpal, ulnar mid carpal, peritrapezoid, peri-trapezoid, peri trapezoid, perizapezium, peri-trapezium, peri trapezium, perihamate, peri-hamate, peri hamate, peripisiforme, peri-pisiforme, peri pisiforme, pisiforme, pisotriquetral, piso-triquetral, scapholunate, scapho-lunate, lunotriquetral, luno-triquetral, radiocarpal, radio-carpal, radiocarpal joint, radio-carpal joint, ulnotriquetral, ulno-triquetral, capitulunate, capito-lunate, distal radioulnar joint, distal radio-ulnar joint, DRUJ, VISI, DISI, CID, CIND
2. instability, nonunion, subluxation, dissociation, pain, painful, ligament, trauma, injury, dysfunction, disruption, collapsus, collaps

**Indextest:**

scaphoid shift test, Watson test, Watson's test, scaphoid shear test, scaphoid thrust test, scaphoid lift test, synovial irritation test, scaphoid bell sign, scapholunate ballottement test, scapho-lunate ballottement test, wrist flexion finger extension maneuver, wrist-flexion finger-extension maneuver, wrist flexion finger extension manoeuvre, wrist-flexion finger-extension manoeuvre, lunotriquetral ballottement test, luno-triquetral ballottement test, LT ballottement test, Regan test, Regan's test, lunotriquetral compression test, luno-triquetral compression test, LT compression test, lunotriquetral shear test, luno-triquetral shear test, LT shear test, Kleinman test, Kleinmann test, Kleinman's test, Kleinmann's test, shuck test, triquetral lift maneuver, triquetral lift manoeuvre, Linscheid test, Linscheid's test, ulnar snuff box compression test, raised triquetrum test, Lichtman test, Lichtmann test, Lichtman's test, Lichtmann's test, catch up clunk test, midcarpal shift test, mid-carpal shift test, midcarpal test, mid-carpal test, pivot shift test, dorsal capitate displacement apprehension test, dorsal capitate displacement apprehension test, wrist hanging test, finger extension test, FET, anterior posterior drawer test, AP drawer test, A-P drawer test, radiocarpal glide test, radio-carpal glide test, mid-carpal snap sign, midcarpal snap sign, triquetrohamate shear test, triquetro-hamate shear test, pisiforme grinding test, triangular fibrocartilage complex test, GRIT, gripping rotatory impaction test, Triangular Fibrocartilage Complex Stress test, Triangular Fibrocartilage Complex Stress test with compression, TFCC Stress test, TFCC Stress test with compression

**Comparaison:**

arthroscopy, arthrography, MRI, magnetic resonance imaging, radiography, magnetic resonance, CT, computer tomography, medical imaging

**Methodologischer Filter:**

diagnostic test accuracy, sensitivity, specificity, validity, reliability, Clinical examination, evaluation, diagnostic, diagnosis, sensitivity and specificity[MeSH], sensitivity[tw], specificity[tw], predictive value\*[tw], false positiv\*[tw], false negativ\*[tw], observer variation\*[tw], roc curve\*[tw], likelihood ratio\*[tw], Likelihood Functions[MeSH] (The Cochrane Collaboration, 1996)

## **2.4 Suche auf Datenbanken**

Folgende Datenbanken haben wir gemeinsam systematisch mit unserer Suchstrategie durchsucht: Pubmed, Web of Science, Cinahl, Cochrane, PeDro, Otseeker und Embase. Die erste Suche haben wir in der 44. Kalenderwoche des Jahres 2013 durchgeführt, die zweite in der 09. Woche im Jahr 2014.

## **2.5 Datensammlung und Analyse**

### **2.5.1 Sortieren nach Titel und Abstrakt**

Zuerst haben wir alle gefundenen Artikel in eine Excel-Tabelle eingefügt um eine bessere Übersicht zu haben und uns das Sortieren zu vereinfachen. Danach haben wir unabhängig voneinander die Artikel nach ihrem Titel aussortiert, indem wir sie mit einem "Ja", einem "Nein", einem "?" oder "Background" markiert haben. Anschliessend verglichen wir unsere Resultate, einigten uns bei Meinungsverschiedenheiten und sortierten die übriggebliebenen Artikel nach den Abstrakten. Wir bewerteten sie wie beim Aussortieren nach Titel ebenfalls unabhängig voneinander mit einem "Ja", einem "Nein", einem "?" oder "Background". Und auch hier verglichen wir unsere Resultate, lasen die Abstrakte, bei denen wir nicht einig waren, nochmals gemeinsam durch und bestellten die Volltexte der Artikel, die wir mit einem "Ja" oder "?" markiert hatten.

### **2.5.2 Sortieren nach Volltext**

Wenn der Volltext nicht auf den Datenbanken zur Verfügung stand, suchten wir sie auf Google Scholar; sofern er auch da nicht erhältlich war, haben wir dem Autor eine Mail geschrieben und ihn um Zusendung des Artikels gebeten. Falls wir dabei keinen Erfolg hatten, haben wir den Artikel über die Schule HES-SO Wallis erworben.

Auch die Volltexte wurden unabhängig voneinander von uns beurteilt und danach die Resultate verglichen. Bei unterschiedlicher Meinung zogen wir die Hilfe unseres Bachelorarbeitsbetreuers hinzu.

### **2.5.3 Lesetabelle**

Die wichtigsten Resultate der gelesenen Artikel wurden in eine Lesetabelle eingetragen. Sie beinhaltet Rubriken wie Autor, Datum, Titel, Land, Ziel der Studie, Population (Anzahl, Ein- und Ausschlusskriterien, Läsionsmechanismus, Zeit bevor Arztvisite),

Diagnostische Tests (Blindiert, Beschreibung vorhanden, Durchführperson), Zeit zwischen Indextest und Referenztest, Referenztest (Blindiert, Beschreibung vorhanden, Durchführperson), Studiendesign, Resultate, Klassifikation der Instabilität, Interpret der Resultate, Ethik, Bias und Kommentare.

## **2.6 Quersuche**

Damit wir keine wichtige Studie übersahen haben wir auch alle Bibliographien der von uns eingeschlossenen Artikel durchgelesen; bei für uns wichtig erscheinendem Titel haben wir den Abstrakt und eventuell sogar den Volltext durchgelesen. Diese Artikel wurden ebenfalls nach unseren obengenannten Kriterien ein- oder ausgeschlossen.

## **2.7 Suche der Volltexte**

Sofern wir den Volltext zu lesen hatten versuchten wir ihn direkt von der Datenbank, auf der wir ihn gefunden hatten, herunterzuladen. War er auf der Datenbank nicht zugänglich, versuchten wir es über den Switch-AAI-Login der Fachhochschule. Weiter suchten wir sie auch auf der Datenbank der Konferenz der Fachhochschulen (KFH). Konnten wir den Artikel immer noch nicht finden, schrieben wir dem Autor eine E-Mail. Bekamen wir auch hier keine Antwort, erwarben wir den Volltext über die Mediathek der HES-SO Wallis.

## **2.8 Datenextraktionstabelle**

Alle eingeschlossenen Artikel haben wir zusätzlich noch in unsere Datenextraktionstabelle eingefügt um danach einfacher mit den Daten arbeiten zu können. Die Tabelle enthält Informationen wie Autor, Jahr, Indextest, getestete Struktur durch Indextest, Seitenvergleich, Indextestbeschreibung, Anzahl Teilnehmer, richtig positive Testergebnisse (TP), falsch positive Testergebnisse (FP), richtig negative Testergebnisse (TN), falsch negative Testergebnisse (FN), Sensitivität, Spezifität, positiver und negativer Likelihood Ratio, positive und negative prädiktiver Wert, Kriterium für positiver Test (threshold), Referenzstandart, Ort, Population und Design. Siehe Zusammenfassung Extraktionstabelle im Kapitel 3.8 Resultate Datenextraktion der Ergebnisse.

## **2.9 Studienqualität mit der QUADAS-2**

QUADAS-2 ist ein Analysewerkzeug, ähnlich einem Fragebogen, das gezielt auf diagnostische Studien bezogen ist. Sie erlaubt dem Leser mögliche Bias zu erkennen und diese kritisch zu bewerten. Wir haben die Quadas-2 benutzt, um die Bias-Risiken unserer eingeschlossenen Studien zu bestimmen.

QUADAS-2 besteht aus vier Abschnitten. Im Ersten wird die Fragestellung der Studie präsentiert, mit allen genauen Daten. Der Zweite ermöglicht es, Fragen der QUADAS-2 unserer Suchfrage anzupassen, um sie bei den eingeschlossenen Studien unserer Meta-Analyse anwenden zu können. Für unsere Studie konnten wir die Fragen in der Originalfassung belassen. Im dritten Abschnitt wird durch ein Flussdiagramm das Verständnis der verschiedenen Etappen der Studie besser sichtbar gemacht und so eventuelle Bias aufgezeigt. Der letzte Abschnitt, der nochmals unterteilt ist, besteht darin, die Bias und den Anwendungsgrad für unsere Review zu bestimmen. Jeder Abschnitt entspricht einem Fachgebiet, wird durch Fragen analysiert und mit einem "Yes", "No" oder "Unclear" beantwortet. Ein "Yes" bedeutet ein tiefes Risiko für einen Bias und je mehr "Yes" der Artikel erhält, desto besser erachten wir die Qualität der Studie. Diese Fachgebiete sind: die Population, der Indextest, der Referenztest und Zeitabfolge und Ablauf (Whiting et al., 2011).

## **2.10 Analyse**

### **2.10.1 Sensitivität und Spezifität**

Mit einem diagnostischen Test versucht man herauszufinden, ob eine Krankheit bei einem Patienten vorliegt oder nicht. Die diagnostische Testgenauigkeit, auch diagnostische Validität oder Kriteriumsvalidität genannt, wird durch die Sensitivität und Spezifität berechnet (Fischer, Bachmann, & Jaeschke, 2003; Fritz & Wainner, 2001):

- Die Sensitivität beschreibt die Fähigkeit eines Tests, tatsächlich Kranke als krank zu erkennen. Will man eine Krankheit ausschliessen, benötigt man eine hohe Sensitivität. (Eselsbrücke: SnOut: Eine hohe Sensitivität kann bei einem negativen Testergebnis eine Krankheit ausschliessen (out).)
- Die Spezifität beschreibt die Fähigkeit eines Tests, gesunde Patienten tatsächlich als gesund zu identifizieren. Will man eine Krankheit einschliessen, benötigt man eine

hohe Spezifität. (Eselsbrücke: SpIn: Eine hohe Spezifität kann bei einem positiven Testergebnis, die Krankheit einschliessen (in).)

Die Werte für die Sensitivität und Spezifität lassen sich anhand der untenstehenden Formeln berechnen. Die 4-Felder-Tabelle trägt zum Verständnis bei:

	Referenztest $\oplus$	Referenztest $\ominus$
Indextest $\oplus$	A (= richtig positiv, RP)	C (= falsch positiv, FP)
Indextest $\ominus$	B (= falsch negativ, FN)	D (= richtig negativ, RN)

$$\text{Sensitivität} = A / (A+B)$$

$$\text{Spezifität} = D / (D+C)$$

Je mehr der Wert der Sensitivität und Spezifität gegen 1 (oder 100%) strebt, desto aussagekräftiger ist er.

Wenn mehrere Studien die Sensitivität und/oder Spezifität eines selben Tests prüfen, ist es interessant, die verschiedenen Ergebnisse zusammen zu poolen. Dieser Test ergibt eine Sensitivität und Spezifität, als ob die Gesamtpopulation aller gepoolten Studien in einer selben Studie untersucht worden wären.

Die Werte der 4-Felder-Tabelle haben wir in unsere Extraktionstabelle eingetragen und damit die Sensitivität und Spezifität nochmals ausgerechnet und mit den Studienergebnissen verglichen. Wenn die von uns ausgerechnete Sensitivität und Spezifität nicht mit den Studienergebnissen übereinstimmten, haben wir unsere ausgerechneten Werte zur Analyse verwendet.

In der Studie von Prosser et al. (2011) wurden zudem Testpersonen beim Indextest mit einem "Unklar" markiert und weder zu den positiven, noch zu den negativen gezählt. Für die Berechnung der Sensitivität und Spezifität haben wir diese Personen zu den positiv Getesteten hinzugefügt, wobei die Sensitivität etwas höher ausfällt. Diesen Entschluss haben wir im Bezug auf die Anwendung der diagnostischen Tests in der Praxis gefasst, wo die Sensitivität wichtiger ist. Mehr dazu erläutern wir in der Diskussion, Kapitel 4.3 und 4.6.

## **2.10.2 Validität**

Die Kriteriumsvalidität eines Tests ist seine diagnostische Testgenauigkeit, das heisst, sie bestimmt, inwiefern ein Test das misst, was er messen soll. Um die Validität zu bestimmen brauchen wir einen Test, der möglichst genau die Wahrheit wiedergibt (Goldstandard oder Referenztest genannt). Der zu untersuchende Test wird mit diesem Referenztest verglichen. Es ist die Eignung eines Tests in Anbetracht seines Zwecks. In unserer Studie wird die Validität durch die Sensitivität und Spezifität aufgezeigt (Fritz & Wainner, 2001).

## **2.10.3 Begriffe zur diagnostischen Analyse**

### **2.10.3.1 Die Prävalenz**

Die Prävalenz wird im Bezug auf die eingeschlossene Population in den verschiedenen Studien gerechnet. Sie gibt Auskunft über die Anzahl betroffener Patienten mit der Pathologie (RP+FN) in der Gesamtpopulation der Studie. Sie wird in Prozent ausgerechnet und wird ebenfalls als Vortestwahrscheinlichkeit bezeichnet. Die Formel lautet;  $A+D/(A+B+C+D) * 100$  (Fischer et al., 2003).

### **2.10.3.2 Das Konfidenzintervall (CI)**

Wenn eine Studie mehrmals durchgeführt wird und eine Stichprobe mit denselben Variablen wie die angegebene Population (in unserem Fall Personen mit Handgelenkschmerzen) untersucht, decken sich die Resultate in 95% der Fälle mit dem angegebenen Intervall. Bei unseren Resultaten haben wir einen Konfidenzintervall von 95% angegeben. Dieser Wert ist aussagekräftig, solange er über der Grenze von 0,5 oder 50% bleibt (Fischer et al., 2003).

### **2.10.3.3 Die Heterogenität und der $I^2$ -Wert:**

Der  $I^2$ -Wert wird bei der Analyse von mehreren gepoolten Studien verwendet, um herauszufinden, ob sich die Ergebnisse im selben Rahmen im Bezug auf den CI befinden oder stark abweichen. In diagnostischen Studien erwarten wir generell höhere Heterogenität (Higgins, Thompson, Deeks, & Altman, 2003).

- 0% bis 40%: vermutlich unbedeutsam;
- 30% bis 60%: Anzeichen von moderater Heterogenität;
- 50% bis 90%: Anzeichen von bedeutsamer Heterogenität;
- 75% bis 100%: hohe Heterogenität.

### 2.10.3.4 Der p-Wert

Der p-Wert ist die statistische Signifikanz und untersucht, ob die unterschiedlichen Ergebnisse von gepoolten Studien dem Zufall entspringen. Dabei werden zwei Hypothesen gestellt:

Hypothese H<sub>0</sub>: Test ist nicht aussagekräftig

Alternative Hypothese H<sub>1</sub>: der Test ist oder ist ein bisschen aussagekräftig

Wenn  $p \leq 0.05$ : H<sub>0</sub> wird verworfen und H<sub>1</sub> wird angenommen → statistisch signifikant

Wenn  $p \geq 0.05$ : keine Aussage (Higgins et al., 2003)

### 2.10.3.5 Die Likelihood Ratio

Die Likelihood Ratio wird mit Hilfe der Vor- und Nachtestwahrscheinlichkeit berechnet und zeigt auf, wie viel mal häufiger ein Testresultat bei tatsächlich kranken Personen im Vergleich zu Gesunden positiv ausfällt (Pewsnar et al., 2001a).

Es ist nicht empfehlenswert diesen diagnostischen Wert zu poolen. Besser ist, die Sensitivität und Spezifität zusammen zu poolen und danach die LR<sub>+/-</sub> auszurechnen (Zwinderman & Bossuyt, 2008).

## 2.10.4 Auswertung

Mit den ganzen Ergebnissen mussten wir entscheiden, in welcher Weise wir diese zum Alltagsgebrauch auswerten könnten.

Wir haben arbiträre Grenzen gesetzt, um die Testergebnisse auszuwerten und zu entscheiden, ob die Resultate vertrauenswürdig sind, oder ob diese mit Vorsicht gehandhabt werden müssen. Um die Studie als "gut" einzustufen, durfte diese keine roten und maximal drei gelbe Punkte auf der QUADAS-2 aufweisen, die sich nicht in der Applikabilität befanden. Wenn ein oder mehrere gelbe Punkte in der Applikabilität lagen, aber kein Roter anzutreffen war, haben wir die Studie als "mittelwertig" bezeichnet. Als "schlecht" wurden die Studien eingestuft, die mit einem oder mehreren roten Punkten benotet wurden und sich weniger als zwei grüne Punkte in dem Bias-Risiko befanden.

Um einen Indextest als empfehlenswert einzustufen haben wir zusätzlich bei den gepoolten Studien die Grenzwerte für die Sensitivität und Spezifität bei 75% angelegt. Bei den nicht gepoolten Studien bei 95%.

## 2.10.5 Arbeitsprogramme:

Zum Erstellen verschiedener Tabellen, Statistiken und Literaturverzeichnisse haben wir folgende Computerprogramme benutzt:

- *Review Manager (RevMan)*. Version 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012 für die Strukturierung der Kapitel, sowie die Qualitätsbewertung mit Quadas-2
- *Meta-DiSc* um die Forest-Plots zu erstellen. Die Software ist frei erhältlich. (Zamora J., Abairra V., Muriel A., Khan K.S. & Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Medical Research Methodology, 2006)
- *Zotero Standalone* für das Literaturverzeichnis und das Artikelmanagement (A production of the Roy Rosenzweig Center for History and New Media at George Mason University and the Corporation for Digital Scholarship).

# 3. Resultate

## 3.1 Resultate der Suche auf Datenbanken

Wir haben folgende Datenbanken systematisch durchsucht: Pubmed, Web of Science, Cinahl, Cochrane, PeDro, Otseeker, Embase und respektive 480, 5, 18, 168, 85, 344, 21 Artikel bei der ersten Suche gefunden. Dies ergibt eine Summe von insgesamt 1121 Studien.

Bei der zweiten Suche fanden wir 16 Neuerscheinungen auf Pubmed, 0 auf Web of Science, 31 auf Cinahl, 12 auf Cochrane, 7 auf PeDro, 18 auf Otseeker und 1 auf Embase. Davon konnten wir jedoch keine in diese Studie einschliessen, da sie unseren Kriterien nicht entsprachen.

Die Zusammenfassung unserer Suche und Auslese wird im Kapitel 3.6 als Flussdiagramm dargestellt.

## 3.2 Resultate nach Ausschluss nach Titel

Zuerst schlossen wir 23 Studien aus, weil sie Duplikate aus unterschiedlichen Datenbanken waren.

Die Studien, die mit einem "ja" oder einem "unklar" markiert waren und nicht genug Informationen im Titel enthielten, um sie auszuschliessen, wurden für den Ausschluss nach Abstrakt einbezogen. So schlossen wir 883 Artikel wegen der nicht zutreffenden Population aus, 39 weil sie keinen Indextest oder Referenztest enthielten oder die Methode nicht mit unserer Suche übereinstimmte. Weiter wurden 32 Artikel ausgeschlossen, weil sie weder auf Deutsch, Französisch oder Englisch verfasst worden waren.

Nach dem Sortieren blieben uns schlussendlich noch 144 Artikel, bei denen wir den Abstrakt lasen.

### **3.3 Resultate nach Ausschluss nach Abstrakte**

Wir schlossen insgesamt 111 Artikel nach der Auslese nach Abstrakt aus: 61 auf Grund der Population, 44 weil sie keinen zutreffenden Indextest enthielten und 6 weil sie keinen Referenztest beinhalteten oder die Methode nicht stimmte. Auch hier behielten wir die Studien, die wir als "unklar" markierten für die weitere Auslese nach Volltext. So blieben uns 33 Artikel für das Sortieren nach Volltext.

### **3.4 Resultate nach Ausschluss nach Volltexte**

Die 33 Artikel wurden wieder unabhängig voneinander gelesen. Am Schluss schlossen wir 7 aus, weil die Population nicht mit unserer Suche übereinstimmte, 7 weitere weil sie keinen Indextest enthielten und 14 weil sie nicht das gewünschte Design aufwiesen. Somit schlossen wir 5 Artikel in unsere Studie ein.

### **3.5 Resultate der Quersuche**

Bei der Quersuche der Bibliographien unserer bereits eingeschlossenen Artikel stiessen wir auf viele interessante Artikel, die wir für die Hintergrundinformation behielten. Wir stiessen aber auch auf einen Artikel, der unseren Einschlusskriterien entsprach und den wir daher in unsere Studie einschlossen (Christodoulou & Bainbridge, 1999).

## 3.6 Flussdiagramm der Suche

Wir haben eine übersichtliche Zusammenfassung unserer Suchresultate in einem Flussdiagramm wie folgt dargestellt:

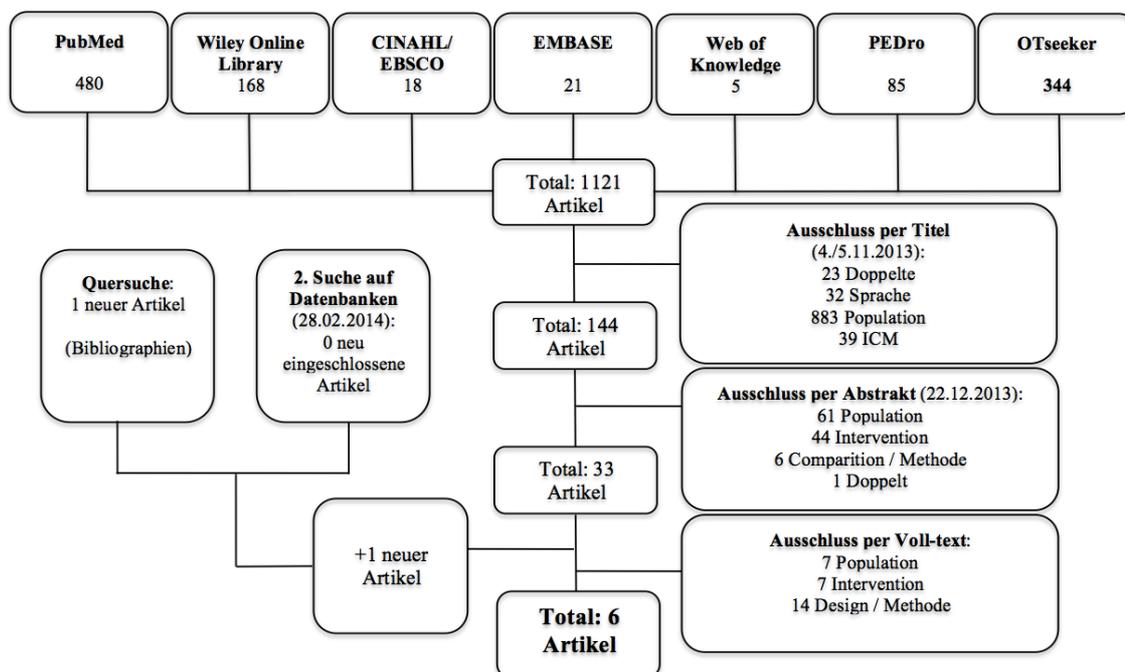


Tabelle 1: Zusammenfassung unserer Suche und Sortieren der Artikel (Tabelle: Autoren)

## 3.7 Qualität der Studien mit der QUADAS-2

Mit der QUADAS-2 haben wir das Bias-Risiko unserer Studien analysiert, sowie die Anwendbarkeit auf unsere Suchfrage bestimmt.

Die Patientenauswahl, der ausgewertete Indextest, der Referenztest, sowie die Zeitabfolge und der Ablauf können einen Einfluss auf die Ergebnisse haben und werden deshalb im Bias-Risiko analysiert. Die ersten drei Punkte werden auch für die Anwendbarkeit ausgewertet. Nachfolgend werden die vier Thematiken in Bezug auf das Bias-Risiko zusammengefasst.

### 3.7.1 Bias-Risiko

**Patientenauswahl:** In unseren Studien sind vorwiegend Patienten mit einem problematischen Handgelenk vertreten, ausser in der Studie von Lester et al. (1995), der eine zusätzliche Kontrollgruppe ohne Beschwerden einschloss.

**Indextest:** In allen Studien wurden im Vornherein die Kriterien zur Aussage über ein positives oder negatives Testergebnis definiert. Weiter wurde bei allen Studien der

Indextest ohne das Wissen über die Referenztestresultate durchgeführt und interpretiert, ausser bei Lester et al. (1995), bei dem die Angaben unklar waren, da alle Patienten schon eine Diagnose gestellt bekommen hatten, bevor die Tests bei ihnen durchgeführt wurden.

**Referenztest:** Alle Studien, die wir eingeschlossen haben, haben die Arthroskopie als Referenztest gewählt, was, wie obengenannt, als Goldstandart für die Diagnosestellung einer Handgelenkinstabilität gilt. Dennoch haben die meisten Studien (Christodoulou & Bainbridge, 1999; LaStayo & Howell, 1995; Lester et al., 1995; Tay et al., 2007) nicht angegeben, ob dem Chirurg während der Arthroskopie die Indextestresultate oder andere Diagnosestellungen (z.B. MRI) bekannt waren.

**Zeitabfolge und Ablauf:** In der Zeitabfolge und Ablauf wurden folgende Elemente angeschaut: Unangemessener Ausschluss von Patienten für einen oder mehrere Tests oder in den Ergebnissen, ob die Zeit zwischen dem Indextest und dem Referenztest angemessen war (keine Zeit für Veränderungen der getesteten Strukturen) und ob alle Patienten mit demselben Referenztest bewertet wurden. Nur bei Prosser et al. (2011) wurden diese Punkte klar ausgedrückt. Bei den anderen waren einer oder mehrere dieser Aspekte unklar und bei Lester et al. (1995) wurde den Patienten zwischen den Tests einer Behandlung unterzogen und er liess auch Patienten bei der arthroskopischen Diagnosestellung aus.

Weiter haben wir die drei Punkte im Bezug auf die Anwendbarkeit zusammengefasst.

### 3.7.2 Anwendbarkeit

Bei der Anwendbarkeit werden die Patientenauswahl, der Indextest und der Referenztest zu unserer Suchfrage ausgewertet.

**Patientenauswahl:** Die Patientenauswahl war in allen Studien klar beschrieben und entsprach unserer Suchfrage. Das Durchschnittsalter lag in den verschiedenen Studien zwischen 31 und 37 Jahren. Nur bei Tay et al. (2007) entsprach das tiefste Alter von 12,6 Jahren nicht unseren Einschlusskriterien, da wir uns eine Altersgrenze von mindestens 16 Jahren gesetzt hatten. In der Studie wurden eins oder mehrere Kinder unter 16 Jahren getestet. Das Durchschnittsalter dieser Studie liegt dennoch bei 33,7 Jahren, daher sollten die Ergebnisse davon nicht allzu stark beeinflusst worden sein.

**Indextest:** Die Indextests der Studien sind alles manuelle Provokationstests, die eine Instabilität prüften. Nur in der Studie von Prosser et al. (2011) mussten wir einen der Indextests, den "Gripping Rotary Impaction Test" (GRIT), ausschliessen, da es kein

manuell durchführbarer Test ist. Und die Testabfolge von Christodoulou & Bainbridge (1999) bewerteten wir mit einem "unklar", da der Ursprung der einzelnen Tests unklar ist.

**Referenztest:** Der Referenztest ist bei allen Studien die Arthroskopie, die als Goldstandart für unsere Suchfrage gilt.

Die bestbewertete Studie war die von Prosser et al. (2011), welche auch die jüngste Studie ist. Die am schlechtesten abschneidende Studie war die von Lester et al. (1995).

Der nächste Schritt bestand darin, die jeweiligen Tests und Ergebnisse der verschiedenen Studien zu entnehmen, um übereinstimmende Indextests poolen zu können.

Im Anhang befinden sich die Tabellen, in denen wir die einzelnen Items der QUADAS-2 detailliert haben.

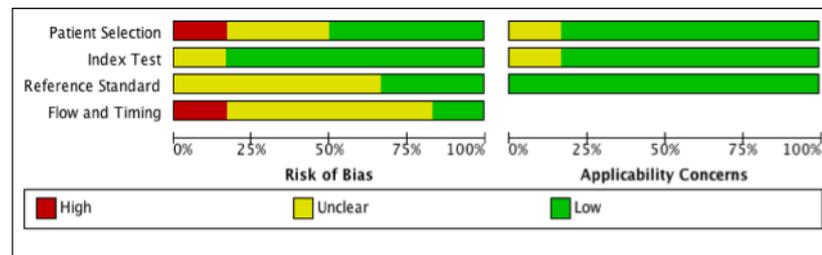


Tabelle 2: Bias-Risiko und Anwendbarkeit der QUADAS2-Punkte (Tabelle: Autoren)

	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Christodoulou 1999	+	+	?	?	+	?	+
Hodgson 1995	?	+	+	?	+	+	+
LaStayo 1995	?	+	?	?	+	+	+
Lester 1995	-	?	?	-	+	+	+
Prosser 2011	+	+	+	+	+	+	+
Tay 2007	+	+	?	?	?	+	+

Tabelle 3: Bias-Risiko und Anwendbarkeit der einzelnen Studien (Tabelle: Autoren)

### 3.8 Resultate Datenextraktion der Ergebnisse

Unseren Studien konnten wir zehn verschiedene Tests entnehmen, deren Validität ausgewertet wurde. Von den zehn Tests wurden vier in je zwei Studien untersucht, die wir nachfolgend poolen konnten.

In der folgenden Extraktionstabelle werden die zur Analyse notwendigen Daten der eingeschlossenen Studien nach Indextest sortiert dargestellt. Es zeigt sich dabei, ob die Studien homogen genug sind, um miteinander gepoolt werden zu können.

Test Autor, Jahr	Anzahl Teilnehmer	TP	FP	TN	FN	Sensi	Spezi	LR +	LR -	Referenz Standard	Population	Design
<b>Scaphoid Shift Test</b>												
LaStayo, 1995	50	11	12	21	6	65	64	1.81	0.55	Arthroskopie	26 Männer, 24 Frauen, Durchschnittsalter : 38 (16-67)	Diagnostic test accuracy
Prosser, 2011	105	36	22	39	8	82	64	2.28	0.29	Arthroskopie	Durchschnittsalter 37 +/- 12	Cross-sectional Study mit Prospektiver diagnostic accuracy
<b>Lunotriquetral Ballottement Test</b>												
LaStayo, 1995	50	7	22	17	4	64	44	1.15	0.82	Arthroskopie	26 Männer, 24 Frauen, Durchschnittsalter : 38 (16-67)	Diagnostic test accuracy
Prosser, 2011	105	3	37	62	3	50	63	1.36	0.8	Arthroskopie	Durchschnittsalter 37 +/- 12	Cross-sectional Study mit Prospektiver diagnostic accuracy
<b>Triangular Fibrocartilage Complex Test (with and without compression)</b>												
Prosser, 2011	105	43	21	24	17	72	53	1.54	0.53	Arthroskopie	Durchschnittsalter 37 +/- 12	Cross-sectional Study mit Prospektiver diagnostic accuracy
Hodgson, 1995	40	7	8	20	5	58	71	2	0.6	Arthroskopie	17 Männer, 23 Frauen, Durchschnittsalter 31.7 (16-65)	Diagnostic test accuracy
<b>Distal Radioulnar Joint Test</b>												
Prosser, 2011	105	14	37	51	3	82	58	1.96	0.32	Arthroskopie	Durchschnittsalter 37 +/- 12	Cross-sectional Study mit Prospektiver diagnostic accuracy
Hodgson, 1995	40	2	2	36	0	100	95	20	0	Arthroskopie	17 Männer, 23 Frauen, Durchschnittsalter 31.7 (16-65)	Diagnostic test accuracy
<b>Ulnomeniscotriquetral Dorsal Glide Test (UMTDG)</b>												
LaStayo, 1995	50	14	20	18	8	64	47	1.21	0.77	Arthroskopie	26 Männer, 24 Frauen, Durchschnittsalter : 38 (16-67)	Diagnostic test accuracy
<b>Midcarpal Test (MC Test)</b>												
Prosser, 2011	105	5	41	59	0	100	59	2.44	0	Arthroskopie	Durchschnittsalter 37 +/- 12	Cross-sectional Study mit Prospektiver diagnostic accuracy
<b>Ulnar Fovea Sign</b>												
Tay, 2007	272	139	17	109	7	95	87	7.31	0.06	Arthroskopie	146 Männer, 126 Frauen, Durchschnittsalter 33.7 (12.6-74.7)	Diagnostic test accuracy
<b>Scapholunate Shear Test (SL-Shear Test)</b>												
Hodgson, 1995	40	19	1	20	0	100	95	20	0	Arthroskopie	17 Männer, 23 Frauen, Durchschnittsalter 31.7 (16-65)	Diagnostic test accuracy
<b>Press Test</b>												
Lester, 1995	27 100 100	14	13	200	0	100	94	16.7	0	Arthroskopie (nicht alle Testpersonen)	27 mit diagnostizierter TFCC-Läsion (16 Männer, 9 Frauen, Durchschnittsalter 31 Jahre (20-49). 100 beschwerdefreie Handgelenke. 100 schmerzende Handgelenke ohne TFCC-Läsion.	Case-control Design
<b>Christodoulou Testabfolge (3 Testes)</b>												
Christodoulou, 1999	62	17	2	38	5	77	95	15.4	0.25	Arthroskopie	27 Männer, 35 Frauen, Durchschnittsalter 31.8 Jahre	Diagnostic test accuracy

Tabelle 4: Zusammenfassung der Datenextraktionstabelle (Tabelle: Autoren)

Abkürzungen: (alle Zahlen wurden aufgerundet)

TP: true positive in [%]

FP: false positive in [%]

TN: true negative in [%]

FN: false negative in [%]

Sensi: Sensitivität in [%]

Spezi: Spezifität in [%]

LR+: positiver Likelihood-Ratio

LR-: negativer Likelihood-Ratio

## 3.9 Meta-Analyse

### 3.9.1 Scaphoid Shift Test

Für den Scaphoid Shift Test konnten wir zwei Studien mit insgesamt 155 Patienten zusammen poolen. Die gepoolte Sensitivität ergab 77% (65–87%) mit einem Konfidenzintervall (CI) von 95%. Der  $I^2$ -Wert (Heterogenitätsbestimmung) ist bei nur zwei gepoolten Studien nicht interpretierbar.

Die gepoolte Spezifität ergab 64% (53-73%) mit einem Konfidenzintervall von 95%. Der  $I^2$ -Wert ist bei nur zwei gepoolten Studien nicht interpretierbar.

Die Prävalenz der Population, die eine Läsion des Scapholunatum-Bandes aufwies, betrug 34% in der Studie von LaStayo und 42% in der von Prosser (gepoolt: 39%).

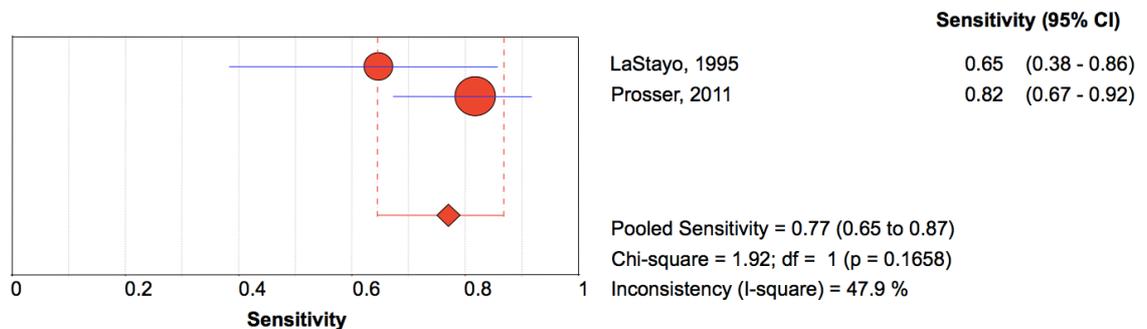


Tabelle 5: Forest Plot Sensitivität Scaphoid Shift Test (Tabelle: Autoren)

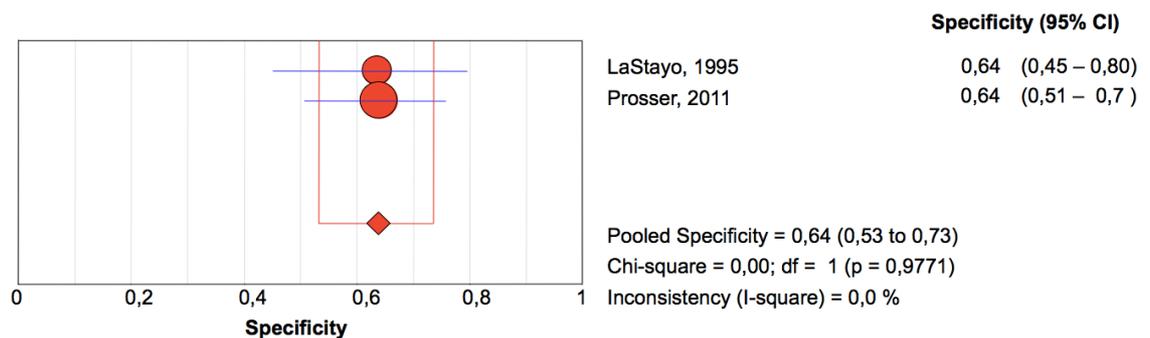


Tabelle 6: Forest Plot Spezifität Scaphoid Shift Test (Tabelle: Autoren)

### 3.9.2 LT-Ballottement Test

Für den Lunotriquetral Ballottement Test konnten wir zwei Studien mit insgesamt 155 Patienten einschliessen. Die gepoolte Sensitivität ergab 59% (33-82%), mit einem Konfidenzintervall von 95%. Der  $I^2$ -Wert ist bei nur zwei gepoolten Studien nicht interpretierbar.

Die gepoolte Spezifität ergab 57% (49-66%) mit einem Konfidenzintervall von 95%. Der  $I^2$ -Wert ist bei nur zwei gepoolten Studien nicht interpretierbar.

Die Prävalenz der Population, die eine Läsion des lunotriquetralen Bandes aufwies, betrug 22% in der Studie von LaStayo und 6% in der von Prosser (gepoolt: 11%).

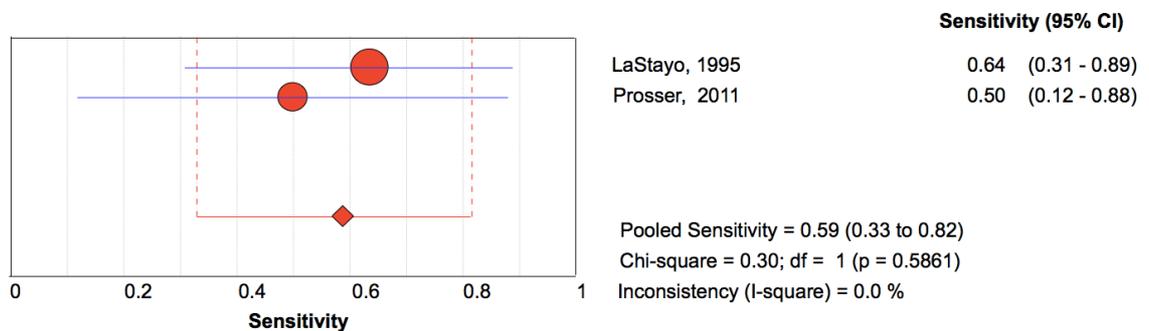


Tabelle 7: Forest Plot Sensitivität LT-Ballottement Test (Tabelle: Autoren)

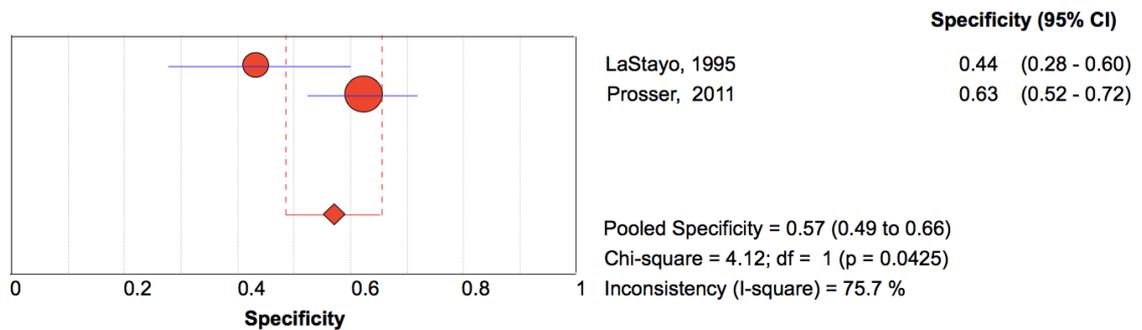


Tabelle 8: Forest Plot Spezifität LT-Ballottement Test (Tabelle: Autoren)

### 3.9.3 TFCC Test

Für den TFCC Test konnten wir zwei Studien mit insgesamt 145 Patienten zusammen poolen. Die gepoolte Sensitivität ergab 69% (57–80%) mit einem Konfidenzintervall (CI) von 95%. Der  $I^2$ -Wert ist bei nur zwei gepoolten Studien nicht interpretierbar.

Die gepoolte Spezifität ergab 60% (48-72%) mit einem Konfidenzintervall von 95%. Der  $I^2$ -Wert ist bei nur zwei gepoolten Studien nicht interpretierbar.

Die Prävalenz der Population, die eine Läsion des Triangular-Fibrokarilage-Komplexes aufwies, betrug 57% in der Studie von Prosser und 30% in der von Hodgson (gepoolt: 50%).

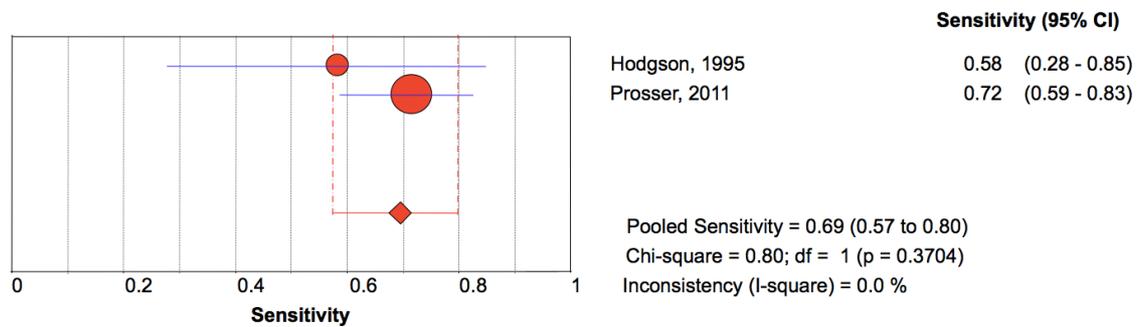


Tabelle 9: Forest Plot Sensitivität TFCC Test (Tabelle: Autoren)

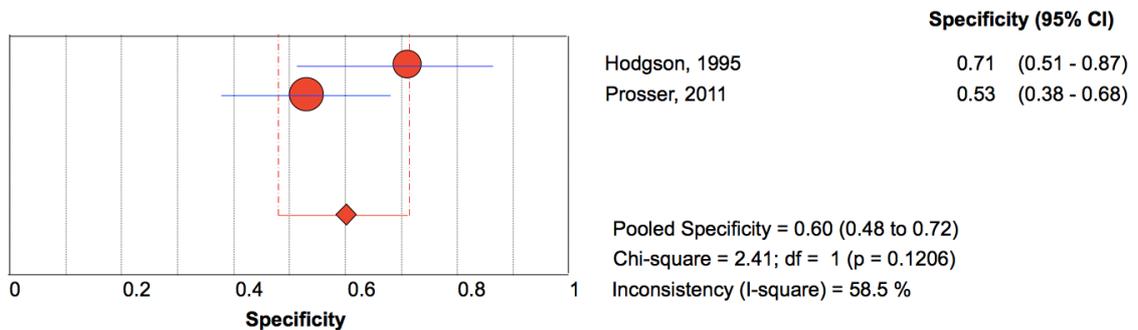


Tabelle 10: Forest Plot Spezifität TFCC Test (Tabelle: Autoren)

### 3.9.4 DRUJ Test

Für den DRUJ Test konnten wir zwei Studien mit insgesamt 145 Patienten zusammen poolen. Die gepoolte Sensitivität ergab 84% (60–97%) mit einem Konfidenzintervall (CI) von 95%. Der  $I^2$ -Wert ist bei nur zwei gepoolten Studien nicht interpretierbar.

Die gepoolte Spezifität ergab 69% (60-77%) mit einem Konfidenzintervall von 95%. Der  $I^2$ -Wert ist bei nur zwei gepoolten Studien nicht interpretierbar.

Die Prävalenz der Population, die eine Läsion der Bänder des distalen radioulnaren Gelenkes aufwies, betrug 16% in der Studie von Prosser und 5% in der von Hodgson (gepoolt: 13%).

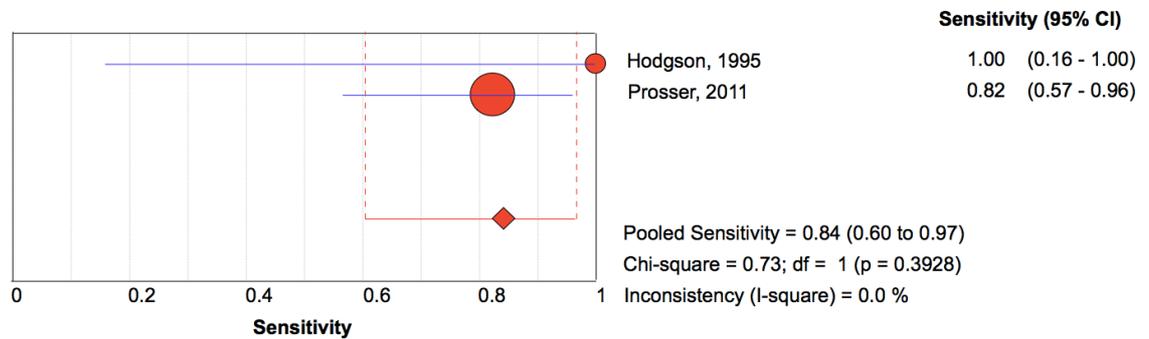


Tabelle 11: Forest Plot Sensitivität DRUJ Test (Tabelle: Autoren)

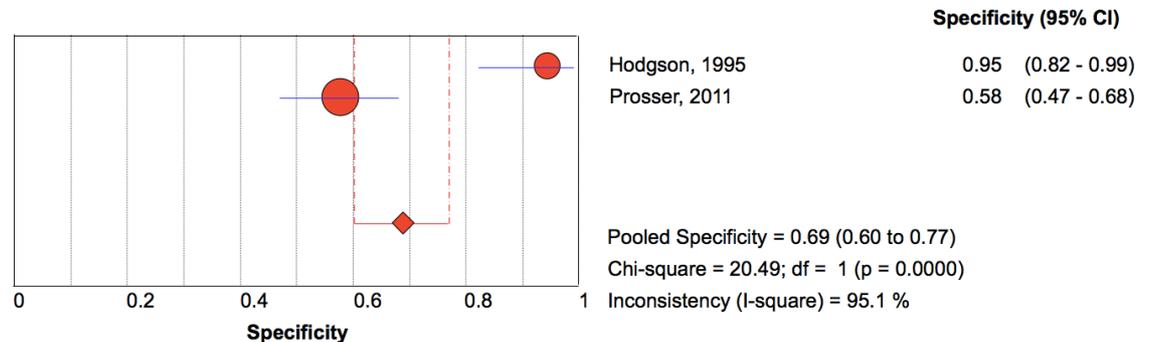


Tabelle 12: Forest Plot Spezifität DRUJ Test (Tabelle: Autoren)

### 3.9.5 Beschreibende Analysen der Studien, die nicht gepoolt werden konnten

Die Sensitivität und Spezifität der folgenden Tests konnte nicht gepoolt werden, da sie nur von einer Studie untersucht wurden. Dennoch haben wir die Prävalenz ausgerechnet und sie zusammen in einem Forest Plot dargestellt:

Test	Sensitivität	Spezifität	Prävalenz
UMTDG	64%	47%	44%
Midcarpal Test	100%	59%	5%
Ulnar fovea sign	95%	87%	54%
SL-Shear Test	100%	95%	48%
Press Test	100%	94%	52%
Testabfolge Christodoulou	77%	95%	35%

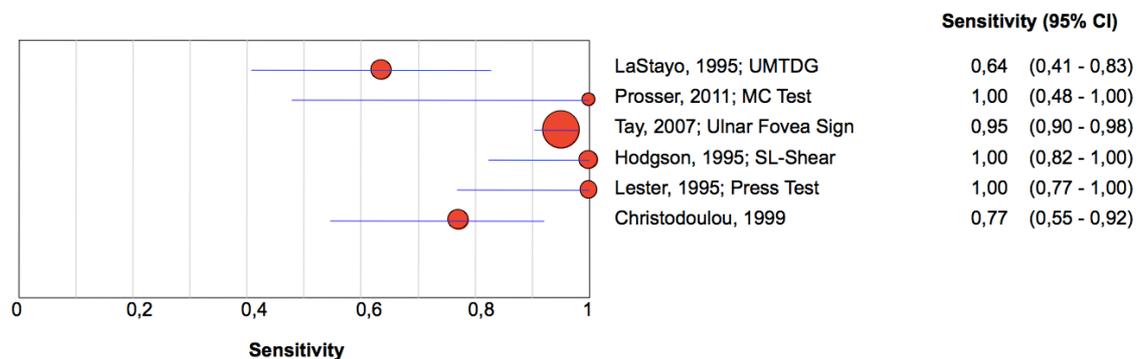


Tabelle 13: Forest Plot Sensitivität verschiedener Tests (Tabelle: Autoren)

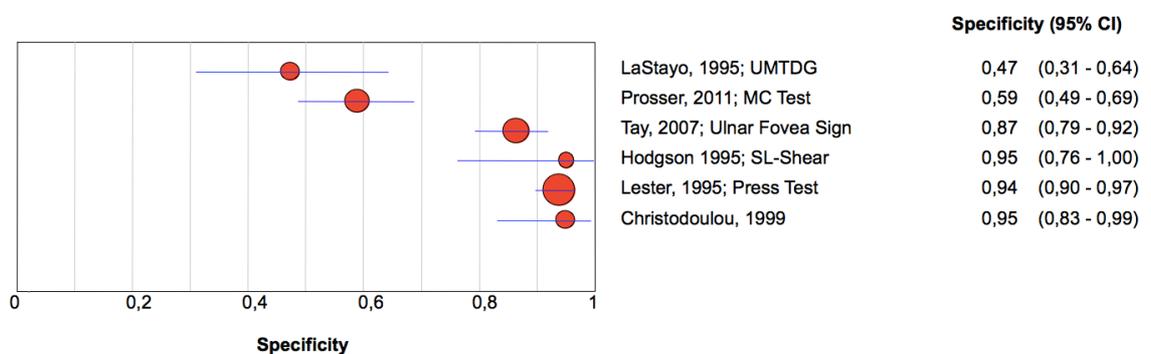


Tabelle 14: Forest Plot Spezifität verschiedener Tests (Tabelle: Autoren)

In der Literatur konnten wir mehrere Beschreibungen von manuellen Provokationstests finden, doch nicht jedes Mal eine dazugehörige Studie, die deren diagnostische Evidenz untersucht. Wie zum Beispiel das Piano Key Sign, das in der Literatur zwar öfters beschrieben, jedoch nie mit einem Referenztest verglichen wird. Auch mussten wir Indextests ausschliessen, die zwar mit einem Referenztest verglichen, jedoch an Kadavern durchgeführt wurden, was zu unseren Ausschlusskriterien gehörte. Hinzu kommt, dass nicht alle Instabilitätstypen über einen zugewiesenen Test verfügen, wie die pisotriquetrale Instabilität.

Die Qualität der von uns eingeschlossenen Studien ist unterschiedlich und deshalb können wir nicht allen die selbe Aussagekraft zuweisen. Infolgedessen empfehlen wir die im nächsten Kapitel zusammengefassten Ergebnisse für den physiotherapeutischen Gebrauch. Uns ist jedoch klar geworden, dass es einfacher ist, eine Instabilität auszuschliessen als diese nachzuweisen. Diesen Aspekt werden wir in der Diskussion noch vertiefen.

# 4. Diskussion

## 4.1 Fazit

Schlussendlich haben wir insgesamt sechs Studien in unsere Review eingeschlossen, die zusammen zehn verschiedene manuelle Provokationstests mit einer Gesamtzahl von 756 Patienten untersuchen.

Wir können folgende Hauptaussagen machen: Der Scaphoid-Shift Test, der DRUJ-Test, der MC-Test sowie das Ulnar Fovea Sign haben eine gute sensitive Aussagekraft, respektive 77%, 84%, 100% und 95%. Der SL-Shear Test weist sowohl eine gute Sensitivität wie eine gute Spezifität auf, respektive 100% und 95%. Bei den restlichen diagnostischen Tests kann man keine weiteren Aussagen machen, da die Studien entweder nicht genügend vertrauenswürdig sind oder die Testwerte unter der von uns fixierten Grenze liegen.

Eine Instabilität war bei den getesteten Personen in 5 bis 54% der Fälle vorhanden (Prävalenz). Dieser Unterschied spiegelt das Vorkommen der verschiedenen Instabilitätstypen wider, wobei Läsionen des SL-Bandes am häufigsten vorkommen.

## 4.2 Auszeichnung dieser Studie

Die starken Punkte unserer Review sind folgende:

Die Erarbeitung unserer Review haben wir zu zweit gemacht. Wir sind systematisch vorgegangen, gemäss den Empfehlungen des Handbuchs der Cochrane Collaboration (Stepnicka, 2008). Dabei wurde die Artikelsuche und Auslese von beiden Autoren unabhängig durchgeführt.

Alle von uns eingeschlossenen Studien können als Referenztest den Goldstandard Arthroskopie aufweisen. Dies steigert die Qualität der Ergebnisse der Studien und trägt zur Homogenität unter den Studien bei. Ein weiterer starker Punkt zur Homogenität ist die Population, deren Durchschnittsalter in allen Studien recht nahe beieinander liegt (zwischen 31 und 38 Jahren). Weiter weisen alle Patienten, die in den Studien untersucht wurden, ein schmerzhaftes Handgelenk auf. Daher sind unsere Werte verlässlich (d.h. gute externe Validität), wenn diese auf eine vergleichbare Population angewendet wird. Auch haben alle Studien, ausser jene von Lester et al. (1995), Patienten studiert, wie man sie in der Praxis antrifft: Patienten mit ähnlichen

Beschwerden wurden in die verschiedenen Studien eingeschlossen. Dabei handelt es sich um ein bei Bewegung auftretendes Geräusch, Kraftverlust, anhaltende, diffuse oder punktuelle Schmerzen im Handgelenk oder Schmerzen nach einem Fallereignis. Nur in der Studie von Lester et al. (1995) wurden 2 Kontrollgruppen von jeweils 100 Personen zu der Zielgruppe hinzugefügt. Dabei wurde einerseits die Population im Vergleich zu den anderen Studien geändert, andererseits stellt gemäss Fischer et al. (2003) das case control-Design, die Studie von Lester et al. (1995), ein grosses Bias Risiko dar. Deshalb möchten wir betonen, dass in dieser Studie das Ergebnis einer hundertprozentigen Sensitivität mit Vorsicht zu betrachten ist.

Die Daten (TP, FP, TN, FN) haben wir dank den angegebenen Zahlen jeder Studie mit Hilfe einer Extraktionstabelle dargestellt. Die einzelnen Angaben wurden in eine 4-Felder-Tabelle eingetragen und die Sensitivität und Spezifität nochmals ausgerechnet (Siehe Kapitel 2.10.1 Sensitivität und Spezifität). Die Ergebnisse konnten so mit den in den Studien angegebenen Ergebnissen verglichen werden. Wenn die von uns ausgerechnete Sensitivität und Spezifität nicht mit den Studienergebnissen übereinstimmten, haben wir unsere ausgerechneten Werte zur Analyse verwendet. Dieser Fall ist nur bei der Studie von LaStayo & Howell (1995) aufgetreten. Mögliche Ursache dieser Differenz könnte bei der Auf- und Abrundung der Werte liegen. Solche Differenzen fielen aber sehr klein aus.

Für die meisten Provokationstests werden beide Handgelenke des Patienten untersucht. Dabei zeigt der Seitenvergleich, ob das symptomatische Handgelenk mehr Bewegungsspiel aufweist, Schmerzen oder Geräusche erzeugt. Doch wird beim Referenztest immer nur das symptomatische Handgelenk diagnostiziert, wobei das gesunde Handgelenk als Kontrollgruppe gelten könnte. Dies könnte die zu testenden Werte verändern (LaStayo & Howell, 1995). Der Grund liegt darin, möglich falsch positive Ergebnisse ausfindig zu machen, wie es bei physiologischen strukturellen Veränderungen, beispielsweise einer Hyperlaxität zu finden wäre. Denn in den Studien wird bei einem positiven Indextest sowohl im symptomatischen als auch im asymptomatischen Handgelenk der positive Test im symptomatischen Handgelenk annulliert und zu den negativen gezählt.

Zusätzlich wurde bei der Testdurchführung in den meisten Studien eine logische Reihenfolge beachtet, vor allem wenn die Patienten mehreren Tests unterzogen wurden.

Dabei wurde vor allem die schmerzende Region als letztes getestet, damit die möglicherweise verletzte Struktur zuletzt gereizt wird. Dies soll verhindern dass die anderen Tests durch eventuell anhaltende, "falsche" Schmerzen beeinflusst werden.

### **4.3 Limitationen dieser Studie**

Die Limiten unserer Studie lassen sich wie folgt zusammenfassen:

Der erste Punkt ist unser begrenzter sprachlicher Studieneinschluss, da wir nur Studien in deutscher, französischer oder englischer Sprache berücksichtigt haben. Dadurch können uns verwendbare Artikel in anderen Sprachen entgangen sein. Zusätzlich hat sich unsere Suche auf Internetdatenbanken begrenzt, daher können uns Studien entgangen sein, die in Konferenzbroschüren vorgestellt wurden. Somit hat unsere Studie einen möglichen Publikations-Bias. Hinzu kommt noch der begrenzte Zugriff auf Studien, die bezahlt werden mussten oder nicht auf den Datenbanken vorhanden waren.

Als zweites wäre zu berücksichtigen, dass sich unsere Studie auf das Handgelenk und dessen mögliche Instabilitäten bezieht, die wir in der Literatur ausfindig machen konnten. Demnach beinhaltet sie keine Elemente zur Differenzialdiagnose, sei es für Nachbargelenke oder andere Pathologien, die den Patienten zur ärztlichen Untersuchung bewegen könnten. Autoren wie Hwang, Goldfarb, Gelberman, & Boyer (1999); Prosser et al. (2011) und Tay et al. (2007) haben diese in Betracht gezogen und Pathologien notiert, die Einflüsse auf unsere Indextests haben können. Für die Differenzialdiagnose zu betrachtende Pathologien sind, unter anderem: Frakturen, Synovitis, Ganglien, Knorpelschäden, Arthrosen, crystaline Arthropathien, Extensorcarpi-ulnarissehnen-Subluxationen, Knochennekrosen, Kapselrisse, andere Bandläsionen und das Ulna-impaction-Syndrom. Darauf sind wir in unserer Studie jedoch nicht weiter eingegangen.

Ein weiterer Punkt, der uns in unserer Studie eingeschränkt hat, ist unsere klinische Unerfahrenheit in der Handpraxis. Es könnte sein, dass Patienten mit akuten Schmerzen die Testdurchführung während den ersten 2-3 Tage nicht unbedingt zulassen würden oder keine adäquate Antwort auf die getestete Struktur gegeben hätten, da der Schmerz noch akut ist. Dieser Aspekt wird von Trail et al. (2007) angedeutet, aber von keinem anderen Autoren erwähnt.

Am Anfang unserer Arbeit haben wir verschiedene Tests in der Literatur gesucht, um jene ausfindig zu machen, die das Handgelenk untersuchten. Dabei haben wir alle möglichen eingeschlossen, um die Suchstrategie zusammenzustellen, wobei manche Tests jedoch nicht unbedingt die Instabilität testen oder manuelle Tests ohne Hilfsmittel sind. So kann man uns vorwerfen, möglicherweise nicht alle existierenden manuellen Instabilitätstests ausfindig gemacht zu haben. Mit der Quersuche wären wir jedoch vermutlich darauf gestossen.

Wie in den Resultaten bereits erwähnt, hat die Studie von Prosser et al. (2011) einige seiner Patienten mit dem Testergebnis "unklar" bewertet. Dies hat uns für die Berechnung der Sensitivität und Spezifität behindert. Deshalb haben wir entschieden, diese unterteilte Patientengruppe zu den positiven Testergebnissen hinzuzufügen, da für uns in der Physiotherapie die Spezifität in diesem Bereich wichtiger ist. Damit vermeiden wir, in der Praxis falsch negative Patienten zu verpassen und tatsächlich kranke Patienten können zu weiteren Untersuchungen geschickt werden. Indem wir die "Unklaren" zu den positiven zählen, wird die Spezifität nicht künstlich erhöht. Bei den Ergebnissen der Sensitivität vom Scaphoid Shift Test, vom MC-Test und vom DRUJ muss daher beachtet werden, dass sie durch unser Vorgehen leicht erhöht worden sind.

Gemäss dem Bayes'sche Theorem ist es möglich, verschiedene diagnostische Tests zusammen zu hängen in dem man die einzelnen Likelihood Werte und die Vortestwahrscheinlichkeit multipliziert, um die diagnostische Entscheidungskraft zu verbessern. Die Testsgrössen müssen dabei aber grundsätzlich unabhängig sein, das heisst, sie dürfen keinen direkten Bezug zueinander haben (wie z.B. Fieber und Leukozytenerhöhung im Blut zusammenhängen) (Pewsnar et al., 2001b). Doch dies ist mit unseren Studien nicht realisierbar, da die Tests entweder nicht dieselbe Struktur testen, oder aber in der Durchführung zu ähnlich sind.

## **4.4 Vergleich mit ähnliche Studien**

Nach unserem Wissen gibt es zurzeit nur eine andere Review, die die diagnostische Evidenz von Tests der Hand und des Arms untersucht. Es ist die Studie, die von Valdes & LaStayo (2013). publiziert wurde. Der Autor meint bei seiner Suche in der Literatur keine weiteren diagnostischen Studien gefunden zu haben, wobei wir eine zusätzliche entdeckten, die von 1995 stammt und von Hodgson et al. (1995) verfasst wurde. Der

Artikel wurde unsererseits mit der QUADAS-2 sogar gut bewertet. Dieser zusätzliche Artikel hat es uns ermöglicht, bei zwei unserer vier gepoolten Provokationstests den gewichteten Mittelwert zu berechnen.

Wir möchten darauf hinweisen, dass in der Studie von Valdes & LaStayo (2013) der Press Test von Lester et al. (1995) zu den empfehlenswerten Tests klassifiziert worden ist, ohne dessen Qualität, die wir als mittelmässig bezeichnet haben, wirklich in Frage zu stellen.

Valdes & LaStayo (2013) hat zudem weitere Studien eingeschlossen (Nakamura et al., 1997; Wolfe & Crisco, 1994; Wolfe, Gupta, & Crisco III, 1997), die auch wir für unsere Studie untersucht haben, aber ausschliessen mussten, da wir mit den im Artikel angegebenen Zahlen die 4-Felder-Tabelle nicht ausfüllen und somit die Sensitivität und Spezifität nicht ausrechnen konnten.

Weitere Unterschiede zwischen unserer Review und der von Valdes & LaStayo (2013) sind, dass er als Qualitätsbewertungswerkzeug die 12-Punkte MacDermid scale benutzt hat, und wir die QUADAS-2, wie sie vom Cochrane Handbuch (Stepnicka, 2008) empfohlen wird. Die QUADAS-2 stammt von 2011, die MacDermid scale ist schon älter, deshalb hätte auch Valdes die QUADAS-2, die modernere Qualitätsskala benutzen können (MacDermid & Wessel, 2004). Ausserdem haben wir für die Auswertung der Testvalidität die Sensitivität und Spezifität ausgewählt, Valdes & LaStayo (2013) hingegen präsentiert die Likelihood Ratio der verschiedenen Tests, was statistisch nicht empfohlen wird (Zwinderman & Bossuyt, 2008).

Weiter beschreibt Valdes & LaStayo (2013) sein ganzes Vorgehen nicht, zum Beispiel, wie er seine Suchstrategie zusammengestellt hat und wie viele Artikel er aussortiert hat. Ein guter Punkt der Studie von Valdes & LaStayo (2013) ist, dass sie die Tests in "empfehlenswert", "neutral/keine Aussage" und "nicht empfehlenswert" unterscheidet, was für die Praxis sehr hilfreich und sofort umsetzbar ist.

## **4.5 Auswirkungen für die Forschung**

Unsere Studie konnte zehn manuelle Provokationstests ausfindig machen, die in der Praxis anzutreffen sind. Die untersuchten Personen sind repräsentativ für die in der Praxis anzutreffenden Patienten. Bei der Testauswertung sollte man jedoch nicht vergessen, dass die Ergebnisse sich aus subjektiven Kriterien zusammensetzen und dass es noch keine Studie gibt, die die interexaminatorische Verlässlichkeit auswertet. Diese Verlässlichkeit soll aufzeigen, dass bei einem Test, der von verschiedenen

Therapeuten/Mediziner auf einer selben Testperson durchgeführt wird, bei allen ein einheitliches Ergebnis gefunden wird.

Für weitere Studien schlagen wir deshalb vor, die manuellen Provokationstest auf deren interexaminatorische Verlässlichkeit zu testen, um diese Testeigenschaft zu untersuchen. Ein gutes Ergebnis wäre dabei eine möglichst hohe Übereinstimmung zwischen den Testdurchführipersonen aufzuweisen.

Weiter schlagen wir vor, Studien zu machen, die Differenzialdiagnosen berücksichtigen, wie wir es im Kapitel 4.3 beschrieben haben, um deren Einfluss auf die Indextests ausfindig zu machen.

Eine dritte, von uns als nützlich eingeschätzte Studie wäre eine, die manuelle Tests mit der Aussagekraft von Symptomen (Outcomes wie Fallmechanismen, Schwellungen, Schmerzregionen, Druckschmerz, eingeschränkte Mobilität und Kraftverlust) und bildgebende Verfahren zu bewerten (multivariable Analyse, in der die Korrelationen zwischen den Tests berücksichtigt wird), wie es die Studie von Prosser et al. (2011) schon eingeleitet hat.

## **4.6 Auswirkungen für die Praxis**

Durch unsere Studien tragen wir dazu bei, die Datenbank der diagnostischen Validitätsstudien zu erweitern (Pewsnar et al., 2001a). Damit wird den Praktikern den Zugang zu evidenz-basierenden Information erleichtert. Aus unserer Studie können sowohl Handtherapeuten wie Physiotherapeuten oder Ärzte Nutzen ziehen. Dennoch sollten sie einen kritischen Blick beibehalten, wenn sie diese Tests in ihrer Praxis anwenden. All unsere Ergebnisse und Zahlen sind mit Vorsicht zu gebrauchen.

Wie gut ein Test sein muss hängt von seiner Anwendung ab. Wenn Patienten einen Arzt aufsuchen hat dieser die Aufgabe, eine Diagnose zu stellen. Seine Handlungsschwelle, ob nun eine Behandlung einzuleiten ist oder nicht, ist nicht dieselbe wie zum Beispiel die eines Physiotherapeuten. Der Therapeut nützt die Tests, um sich abzusichern. Er muss sicher gehen, dass seine Behandlung dem Patienten nicht schadet, was unter anderem die Kontraindikationen beinhaltet. Die Handlungsschwelle des Physiotherapeuten ist dann erreicht, wenn nichts Eigenartiges vorliegt, das nicht seiner Berufung entspricht (Fritz & Wainner, 2001). Bei Unsicherheiten hat der Physiotherapeut immer noch die Möglichkeit, Rücksprache mit dem Arzt zu nehmen und den Patienten mit der Behandlung warten zu lassen. Da wir in unserer Studie mehr Wert auf die Sensitivität gelegt haben, um so eine karpale Instabilität bei Patienten

ausschliessen zu können, beziehen sich unsere Resultate eher auf die physiotherapeutische Praxis als auf die ärztliche, da wir in der Physiotherapie keine Instabilität diagnostizieren müssen, sondern diese nur ausschliessen wollen, da wir physiotherapeutisch eine Instabilität nicht wirklich therapieren können. Deshalb haben wir in der Studie von Prosser et al. (2011) die Patienten, die in den Indextests als "unklar" bewertet wurden zu den positiven dazugezählt, um die Sensitivität und Spezifität zu berechnen, da wir in der Physiotherapie bei einem unklaren Testergebnis den Patienten lieber als positiv ansehen und zu weiteren Abklärungen schicken, als eine wichtige Diagnose zu verpassen.

Patienten, die einer Arthroskopie unterzogen werden, bekommen dadurch eine genaue Diagnosestellung und werden, bei Bedarf, während der selben Operation gleich behandelt (Herron & Craigen, 2001). Wir möchten jedoch darauf hinweisen, dass jeglicher Eingriff Nebenwirkungen und Komplikationen nach sich ziehen kann. Deshalb wäre es vorteilhaft, genügend zuverlässige manuelle diagnostische Tests zur Verfügung zu haben, um eine Handgelenkinstabilität ausschliessen zu können. Mit solchen Tests könnte man eine unnötige Arthroskopie und somit die Eingriffsrisiken, hohe Kosten und unnötigen Zeitverlust vermeiden.

## **5. Schlussfolgerung**

Während dieser Arbeit haben wir die diagnostische Validität verschiedener manueller Provokationstests untersucht, die eine Handgelenksinstabilität aufzeigen sollten. Mit den Provokationstests können Praktiker eine solche Instabilität ausschliessen oder den Patienten, falls nötig zu weiteren Untersuchungen schicken. Patienten, bei denen eine frühe Erkennung einer Instabilität gelingt, können mühsame Folgeerkrankungen erspart werden. Die verschiedenen Tests untersuchten jeweils einen Instabilitätstypen und wurden dank einem Vergleich mit der Arthroskopie auf ihre Validität geprüft.

## 6. Literaturverzeichnis

- Ankarath, S. (2006). Chronic wrist pain: Diagnosis and management. *Current Orthopaedics*, 20(2), 141–151. doi:10.1016/j.cuor.2006.01.004
- Cameriere, R., Ferrante, L., Mirtella, D., & Cingolani, M. (2006). Carpals and epiphyses of radius and ulna as age indicators. *International Journal of Legal Medicine*, 120(3), 143–146. doi:10.1007/s00414-005-0040-3
- Carlsen, B. T., & Shin, A. Y. (2008). Wrist instability. *Scandinavian Journal of Surgery: SJS: Official Organ for the Finnish Surgical Society and the Scandinavian Surgical Society*, 97(4), 324–332.
- Christodoulou, L., & Bainbridge, L. C. (1999). CLINICAL DIAGNOSIS OF TRIQUETROLUNATE LIGAMENT INJURIES. *The Journal of Hand Surgery: Journal of the British Society for Surgery of the Hand*, 24(5), 598–600. doi:10.1054/jhsb.1999.0269
- Dumontier, C. (1996). Examen Clinique des Traumatismes Ligamentaires du Poignet. *Maîtrise Orthopédique*, 49.
- Engelhardt, T. O., & Krimmer, H. (2006). Dissoziative Instabilitäten der proximalen Handwurzelreihe. *Handchirurgie*.
- Fischer, J. E., Bachmann, L. M., & Jaeschke, R. (2003). A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Medicine*, 29(7), 1043–1051. doi:10.1007/s00134-003-1761-8
- Forman, T. A., Forman, S. K., & Rose, N. E. (2005). A clinical approach to diagnosing wrist pain. *American Family Physician*, 72(9).
- Fritz, J. M., & Wainner, R. S. (2001). Examining diagnostic tests: an evidence-based perspective. *Physical Therapy*, 81(9), 1546–1564.
- Garcia-Elias, M. (1999). Position Statement: Definition of Carpal Instability: The Anatomy and Biomechanics Committee of the International Federation of Societies for Surgery of the Hand<sup>2</sup>. *The Journal of Hand Surgery*, 24(4), 866–867.
- Herron, M. L., & Craigen, M. A. C. (2001). Arthroscopy of the wrist for trauma. *Trauma*, 3(1), 9–16. doi:10.1177/146040860100300102
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal*, 327(7414), 557.

- Hodgson, S. P., Royle, S. G., & Stanley, J. K. (1995). An approach to the diagnosis of chronic wrist pain. *Journal of the Royal College of Surgeons of Edinburgh*, 40(6), 407–410.
- Hwang, J. J., Goldfarb, C. A., Gelberman, R. H., & Boyer, M. I. (1999). The effect of dorsal carpal ganglion excision the scaphoid shift test. *The Journal of Hand Surgery: Journal of the British Society for Surgery of the Hand*, 24(1), 106–108. doi:10.1016/S0266-7681(99)90053-0
- LaStayo, P., & Howell, J. (1995). Clinical provocative tests used in evaluating wrist pain: a descriptive study. *Journal of Hand Therapy*, 8(1), 10–17.
- Lester, B., Halbrecht, J., Levy, I. M., & Gaudinez, R. (1995). “Press test” for office diagnosis of triangular fibrocartilage complex tears of the wrist. *Annals of Plastic Surgery*, 35(1), 41–45.
- MacDermid, J. C., & Wessel, J. (2004). Clinical diagnosis of carpal tunnel syndrome: a systematic review. *Journal of Hand Therapy: Official Journal of the American Society of Hand Therapists*, 17(2), 309–319. doi:10.1197/j.jht.2004.02.015
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine*, 151(4), 264–269.
- Nakamura, R., Horii, E., Imaeda, T., Nakao, E., Kato, H., & Watanabe, K. (1997). The ulnocarpal stress test in the diagnosis of ulnar-sided wrist pain. *The Journal of Hand Surgery: Journal of the British Society for Surgery of the Hand*, 22(6), 719–723. doi:10.1016/S0266-7681(97)80432-9
- Parmelee-Peters, K., & Eathorne, S. W. (2005). The Wrist: Common Injuries and Management. *Primary Care: Clinics in Office Practice*, 32(1), 35–70. doi:10.1016/j.pop.2004.11.015
- Pewsner, D., Bleuer, J. P., Bucher, H. C., Battaglia, M., Jüni, P., & Egger, M. (2001a). Der Intuition auf der Spur? Das Bayes'sche Theorem und die Diagnostik in der Grundversorgung, Teil 1. *PRAXIS; Schweiz Med Forum*, (Nr. 3), 41–45.
- Pewsner, D., Bleuer, J. P., Bucher, H. C., Battaglia, M., Jüni, P., & Egger, M. (2001b). Der Intuition auf der Spur? Das Bayes'sche Theorem und die Diagnostik in der Grundversorgung, Teil 2. *PRAXIS; Schweiz Med Forum*, (Nr. 3), 46–52.
- Pliefke, J., Stengel, D., Rademacher, G., Mutze, S., Ekkernkamp, A., & Eisenschenk, A. (2008). Diagnostic accuracy of plain radiographs and cineradiography in diagnosing traumatic scapholunate dissociation. *Skeletal Radiology*, 37(2), 139–

145. doi:10.1007/s00256-007-0410-7

- Prosser, R., Harvey, L., Lastayo, P., Hargreaves, I., Scougall, P., & Herbert, R. D. (2011). Provocative wrist tests and MRI are of limited diagnostic value for suspected wrist ligament injuries: a cross-sectional study. *Journal of Physiotherapy*, *57*(4), 247–253. doi:10.1016/S1836-9553(11)70055-8
- Prosser, R., Herbert, R., & LaStayo, P. C. (2007). Current Practice in the Diagnosis and Treatment of Carpal Instability—Results of a Survey of Australian Hand Therapists. *Journal of Hand Therapy*, *20*(3), 239–243. doi:10.1197/j.jht.2007.04.006
- Roduner, Y., Baeriswyl, E., & Stahl, H. (2013, April). Wundheilung des Bindegewebes - Physiotherapeutische Prinzipien und Behandlung in den Wundheilphasen des Bindegewebes.
- Rutjes, A. W. S., Reitsma, J. B., Vandenbroucke, J. P., Glas, A. S., & Bossuyt, A. M. M. (2005). Case-Control and Two-Gate Designs in Diagnostic Accuracy Studies. *Clinical Chemistry*, *51*(8), 1335–1341. doi:10.1373/clinchem.2005.048595
- Schmitt, R., Froehner, S., Coblenz, G., & Christopoulos, G. (2006). Carpal instability. *European Radiology*, *16*(10), 2161–2178. doi:10.1007/s00330-006-0161-1
- Schünemann, A. H. J., Oxman, A. D., Brozek, J., Glasziou, P., Jaeschke, R., Vist, G. E., ... Guyatt, G. H. (2008). GRADE: grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ: British Medical Journal (International Edition)*, *336*(7653).
- Spies, C. K., Müller, L. P., Oppermann, J., Hahn, P., & Unglaub, F. (2014). [Instability of the Distal Radioulnar Joint - an Overview of Clinical and Radiological Procedures Regarding their Efficacies.]. *Handchirurgie, Mikrochirurgie, plastische Chirurgie: Organ der Deutschsprachigen Arbeitsgemeinschaft für Handchirurgie: Organ der Deutschsprachigen Arbeitsgemeinschaft für Mikrochirurgie der Peripheren Nerven und Gefässe: Organ der V...* doi:10.1055/s-0033-1363662
- Stepnicka, P. (2008). *Ferrocenes: ligands, materials and biomolecules*. John Wiley & Sons.
- Tay, S. C., Tomita, K., & Berger, R. A. (2007). The “ulnar fovea sign” for defining ulnar wrist pain: an analysis of sensitivity and specificity. *The Journal of Hand Surgery*, *32*(4), 438–444.
- The Cochrane Collaboration (Ed.). (1996). *Cochrane Database of Systematic Reviews*:

*Reviews*. Chichester, UK: John Wiley & Sons, Ltd.

- Trail, I. A., Stanley, J. K., & Hayton, M. J. (2007). Twenty questions on carpal instability. *Journal of Hand Surgery: European Volume*, 32(3), 240–255. doi:10.1016/j.jhsb.2007.02.008
- Valdes, K., & LaStayo, P. (2013). The value of provocative tests for the wrist and elbow: A literature review. *Journal of Hand Therapy*, 26(1), 32–43. doi:10.1016/j.jht.2012.08.005
- Van Schoonhoven, J., Prommersberger, K.-J., Krimmer, H., & Lanz, U. (1999). Läsionen der skapholunären Bandverbindung. *Trauma Und Berufskrankheit*, 1(2), 173–181.
- Watson, H. K., Ashmead 4th, D., & Makhlof, M. V. (1988). Examination of the scaphoid. *The Journal of Hand Surgery*, 13(5), 657–660.
- Whiting, P. F., Rutjes, A. W. S., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., ... Bossuyt, P. M. M. (2011). QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 155(8), 529–536.
- Wolfe, S. W., & Crisco, J. J. (1994). Mechanical evaluation of the scaphoid shift test. *The Journal of Hand Surgery*, 19(5), 762–768. doi:10.1016/0363-5023(94)90180-5
- Wolfe, S. W., Gupta, A., & Crisco III, J. J. (1997). Kinematics of the scaphoid shift test. *The Journal of Hand Surgery*, 22(5), 801–806.
- Zwinderman, A. H., & Bossuyt, P. M. (2008). We should not pool diagnostic likelihood ratios in systematic reviews. *Statistics in Medicine*, 27(5), 687–697. doi:10.1002/sim.2992

## 7. Abbildungsverzeichnis

Abbildung 1: Karpus von palmar .....	2
Abbildung 2: Der Scaphoid Shift Test .....	5
Abbildung 3: Der Lunitriquetral Test .....	6
Abbildung 4: Der Triangular Fibrocartilage Complex Stress Test .....	6
Abbildung 5: Der Distal Radio-Ulnar Joint Test .....	7
Abbildung 6: Der Ulnomeniscotriquetral Dorsal Glide Test .....	7
Abbildung 7: Der Midcarpal Test .....	8
Abbildung 8: Das Ulnar fovea sign.....	8
Abbildung 9: Der Scapholunate Shear Test .....	9
Abbildung 10: Der Press-Test.....	9
Abbildung 11: Erster Test der Christodoulou Testabfolge .....	9
Abbildung 12: Zweiter Test der Christodoulou Testabfolge .....	10
Abbildung 13: Dritter Test der Christodoulou Testabfolge .....	10

## 8. Tabellenverzeichnis

Tabelle 1: Zusammenfassung unserer Suche und Sortieren der Artikel.....	25
Tabelle 2: Bias-Risiko und Anwendbarkeit der QUADAS2-Punkte.....	27
Tabelle 3: Bias-Risiko und Anwendbarkeit der einzelnen Studien .....	27
Tabelle 4: Zusammenfassung der Datenextraktionstabelle.....	28
Tabelle 5: Forest Plot Sensitivität Scaphoid Shift Test.....	29
Tabelle 6: Forest Plot Spezifität Scaphoid Shift Test .....	29
Tabelle 7: Forest Plot Sensitivität LT-Ballottement Test .....	30
Tabelle 8: Forest Plot Spezifität LT-Ballottement Test .....	30
Tabelle 9: Forest Plot Sensitivität TFCC Test .....	31
Tabelle 10: Forest Plot Spezifität TFCC Test .....	31
Tabelle 11: Forest Plot Sensitivität DRUJ Test .....	32
Tabelle 12: Forest Plot Spezifität DRUJ Test .....	32
Tabelle 13: Forest Plot Sensitivität verschiedener Tests.....	33
Tabelle 14: Forest Plot Spezifität verschiedener Tests .....	33

**Quelle aller Abbildungen und Tabellen: Autoren**

# 9. Anhang

## 9.1 Detaillierte QUADAS-2 Tabellen

Item 1 – Patient Selection .....	II und III
Item 2 – Index Test .....	II und III
Item 3 – Reference standart.....	IV und V
Item 4 – Flow and Timing.....	IV und V

		QUADAS-2	Christodoulou, 1999	Hodgson, 1995	LaStayo, 1995
Patient selection	A. Risk of Bias	Patient Sampling	consecutive selection	40 Patienten, mit chron. undiagnostizierten HG Sz.	unspez. HG-Sz.
		Was a consecutive or random sample of patients enrolled?	Yes	Unclear	Yes
		Was a case-control design avoided?	Yes	Yes	Yes
		Did the study avoid inappropriate exclusions?	Yes	Yes	Unclear
		<b>Could the selection of patients have introduced bias?</b>	Low risk	Unclear risk	Unclear risk
	B. Concerns regarding applicability	Patient characteristics and setting	Pat with wrist pain and had an arthroscopy between 1994 and 1998	23 Frauen, 17 Männer, Durchschnittsalter 31.7 Jahre	4 Wochen mindest Beschwerden
<b>Are there concerns that the included patients and setting do not match the review question?</b>		Low concern	Low concern	Low concern	
Index Test		Index tests	3 Tests, ev. gleich wie Scaph shift, Lt ballotment und Shear test	Algorithmus, SL shear, DRUJ, TFCC, Tenderness, AP glide, pseudo-stability	SST, Ballotement test, UMTDG Test
	A. Risk of Bias	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Yes	Yes
		If a threshold was used, was it pre-specified?	Yes	Yes	Yes
		<b>Could the conduct or interpretation of the index test have introduced bias?</b>	Low risk	Low risk	Low risk
B. Concerns regarding applicability	<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>	unclear concern	Low concern	Low concern	

		QUADAS-2	Lester, 1995	Prosser, 2011	Tay, 2007
Patient selection	A. Risk of Bias	Patient Sampling	prospective study. Nicht verbessernde Symptome nach Konservativem TTT	Undiagnostizierte Handgelenkschmerzen	Handgelenk-schmerzen
		Was a consecutive or random sample of patients enrolled?	Unclear	Yes	Yes
		Was a case-control design avoided?	No	Yes	Yes
		Did the study avoid inappropriate exclusions?	Unclear	Yes	Yes
		<b>Could the selection of patients have introduced bias?</b>	High risk	Low risk	Low risk
	B. Concerns regarding applicability	Patient characteristics and setting	27 Pat (16 M, 9 F), Durchschnittsalter 31Jahre. Physisch-Aktive Pat. Ulnarer Seiten Sz. HG wurde überstrapaziert.	4 Wochen mindest Beschwerden	12.6 Jahre gilt als Kind, dennoch alters mittelwert von 33.7
		<b>Are there concerns that the included patients and setting do not match the review question?</b>	Low concern	Low concern	Unclear concern
Index Test		Index tests	Press test	SS test, LT test, TFCC test, TFCC comp. test, MC test, DRUJ test, GRIT	Ulna fovea sign
	A. Risk of Bias	Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear	Yes	Yes
		If a threshold was used, was it pre-specified?	Yes	Yes	Yes
		<b>Could the conduct or interpretation of the index test have introduced bias?</b>	Unclear risk	Low risk	Low risk
	B. Concerns regarding applicability	<b>Are there concerns that the index test, its conduct, or interpretation differ from the review question?</b>	Low concern	Low concern	Low concern

		QUADAS - 2	Christodoulou, 1999	Hodgson, 1995	LaStayo, 1995
Reference standard	A. Risk of Bias	Target condition and reference standard(s)	Arthroscopie	Arthroscopie three/four radio-carpal portal and the central midcarpal portal	Arthroscopie (Chirurge, 4-6 Jahre Berufserfahrung)
		Is the reference standards likely to correctly classify the target condition?	Yes	Yes	Yes
		Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	Yes	Unclear
		<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	Unclear risk	Low risk	Unclear risk
	B. Concerns regarding applicability	<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	Low concern	Low concern	Low concern
Flow and Timing	A. Risk of Bias	Flow and timing	von 1994 bis 1998	Pat symptome von 7-72 Monaten (1 patient seit 45 Jahren)	Studie über 2 Jahre
		Was there an appropriate interval between index test and reference standard?	Unclear	Unclear	Unclear
		Did all patients receive the same reference standard?	Yes	Yes	Yes
		Were all patients included in the analysis?	Yes	Yes	Yes
		<b>Could the patient flow have introduced bias?</b>	Unclear risk	Unclear risk	Unclear risk
		Note			

△

		QUADAS - 2	Lester, 1995	Prosser, 2011	Tay, 2007
Reference standard	A. Risk of Bias	Target condition and reference standard(s)	Arthroskopie	Arthroskopie (Chirurg, 27 Jahre Berufserfahrung)	Arthroskopie
		Is the reference standards likely to correctly classify the target condition?	Yes	Yes	Yes
		Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear	Yes	Unclear
		<b>Could the reference standard, its conduct, or its interpretation have introduced bias?</b>	Unclear risk	Low risk	Unclear risk
	B. Concerns regarding applicability	<b>Are there concerns that the target condition as defined by the reference standard does not match the question?</b>	Low concern	Low concern	Low concern
Flow and Timing	A. Risk of Bias	Flow and timing	Pat wurden zuerst Behandelt	21 Tage max Intervall zwischen Index Test /Gold Standart	Retrospektive Studie über 7 Jahre
		Was there an appropriate interval between index test and reference standard?	Unclear	Yes	Unclear
		Did all patients receive the same reference standard?	Yes	Yes	Yes
		Were all patients included in the analysis?	No	Yes	Yes
		<b>Could the patient flow have introduced bias?</b>	High risk	Low risk	Unclear risk
		Note		Pat in 2 Gruppen geteilt	