

Bachelorthesis zur Erlangung des FH-Diploms als Bachelor of science HES-SO in Physiotherapie
HES-SO Valais Wallis Bereich Gesundheit & Soziale Arbeit

**Effekt des Interventionsvolumens bei Patienten mit chronischen
unspezifisch lumbalen Rückenschmerzen: eine systematische
Literaturübersicht.**

Erarbeitet von : Lolita Leiggener und Sylvie Ponnaz

Studiengang : BA06

Unter Betreuung von : Roger Hilfiker

Leukerbad, 7. Juli 2009

DANKSAGUNG

Ein herzliches Dankeschön geht an alle Leute, die in irgendeiner Form zu dieser Arbeit beigetragen haben.

Wir bedanken uns bei unserem Betreuer Roger Hilfiker für die kompetente Beratung während der gesamten Arbeit. Danke vielmals für die Hilfe und die nützlichen Tipps.

Vielen Dank an Dr. Tobias Lorenz, Oberarzt in der Rehabilitationsklinik und Leiter des Rückenschmerzentrums Leukerbad. Wir haben seine fachkundige Unterstützung sowie seine wertvollen Ratschläge sehr geschätzt und konnten viel von ihm lernen.

Einen weiteren Dank richten wir an unsere Korrekturleser: Christian Jäger, Franziska Schnitzer und Miriam Venetz. Für ihre konstruktive Kritik und die aufgeopferten Stunden sind wir sehr dankbar.

Wir bedanken uns bei Christophe Buchard für den Einsatz bei der Kopierarbeit.

Danke an alle Personen in unserer Umgebung, die uns in irgendeiner Art beigestanden sind.

Ein herzliches Vergelt's Gott!

INHALTSVERZEICHNIS

ABBILDUNGSVERZEICHNIS	III
TABELLENVERZEICHNIS.....	III
ZUSAMMENFASSUNG	IV
RÉSUMÉ.....	V
1 EINFÜHRUNG	1
1.1 Verbreitung von Rückenschmerzen	1
1.2 Persönliche Motivation	2
2 PROBLEMATIK	3
2.1 Unspezifisch lumbale Rückenschmerzen.....	3
2.2 Chronizität.....	4
2.3 Analytierte Outcomes in unserer Review	4
2.3.1 Schmerz und Disability	4
2.3.2 Anzahl arbeitsunfähige Tage.....	6
2.4 Begründung der Wahl unserer Outcomes	7
2.5 Biopsychosoziales Modell	7
2.6 Hypothese.....	9
3 METHODEN.....	10
3.1 Suchstrategie	10
3.1.1 Pubmedsuche	11
3.1.2 Cochrane central Suche.....	13
3.1.3 CINAHL Suche.....	14
3.1.4 PEDro Suche	14
3.2 Einschluss der Studien	15
3.3 Bewertung der Studien.....	18
3.4 Datenextraktion	20
3.5 Datenanalyse	21
3.6 Messinstrumente	21
3.6.1 Schmerz.....	21
3.6.2 Disability	22
3.6.3 Arbeitsunfähigkeitstage.....	24

4	RESULTATE	25
4.1	Outcome Schmerzintensität	26
4.1.1	Analyse direkt nach der Intervention	26
4.1.2	Analyse 1-8 Monate nach der Intervention	27
4.1.3	Analyse 9-16 Monate nach der Intervention	27
4.1.4	Analyse mehr als 16 Monate nach der Intervention.....	28
4.2	Outcome Disability	28
4.2.1	Analyse direkt nach der Intervention	29
4.2.2	Analyse 1-8 Monate nach der Intervention	29
4.2.3	Analyse 9-16 Monate nach der Intervention	30
4.2.4	Analyse mehr als 16 Monate nach der Intervention.....	31
4.3	Andere Studien.....	31
4.4	Zusammenfassung der Outcomes Schmerzintensität und Disability.....	33
4.5	Outcome Anzahl arbeitsunfähige Tage	35
5	DISKUSSION	38
5.1	Bewertung der Studien	38
5.2	Limitationen unserer Review	39
5.3	Interpretation der Resultate	41
5.3.1	Outcome Schmerzintensität	41
5.3.2	Outcome Disability	41
5.3.3	Outcome Anzahl arbeitsunfähige Tage.....	41
5.3.4	Interpretation der statistischen Analyse	42
5.3.5	Analyse des Verlaufs.....	44
5.4	Zusammenfassung der Diskussion.....	44
6	SCHLUSSFOLGERUNG.....	47
7	BIBLIOGRAPHIE.....	48
8	ANHANG	VI
	Anhang A: Red und Yellow Flags	VI
	Anhang B: Definitionen der elf Bewertungskriterien der PEDro-Skala.....	VII
	Anhang C: Oswestry Disability Questionnaire	VIII
	Anhang D: Quebec Back Pain Disability Scale	IX
	Anhang E: Roland-Morris Disability Questionnaire.....	X
	Anhang F: ADL-Assessment nach Manniche.....	XI
	Anhang G: Pain and Disability Million Index	XII

ABBILDUNGSVERZEICHNIS

Abbildung 1: Flussdiagramm der Studien.	16
Abbildung 2: Übersicht der Bewertung nach PEDro-Skala.	19
Abbildung 3: Forest Plot der Schmerzintensität direkt nach der Intervention.	26
Abbildung 4: Forest Plot mit Vergleich der Schmerzintensität 1-8 Monate nach der Intervention.	27
Abbildung 5: Forest Plot mit Vergleich der Schmerzintensität 9-16 Monate nach der Intervention.	27
Abbildung 6: Forest Plot der Schmerzintensität mehr als 16 Monate nach der Intervention.	28
Abbildung 7: Forest Plot der Disability direkt nach der Intervention.	29
Abbildung 8: Forest Plot mit Vergleich der Disability 1-8 Monate nach der Intervention.	30
Abbildung 9: Forest Plot mit Vergleich der Disability 9-16 Monate nach der Intervention.	30
Abbildung 10: Forest Plot der Disability mehr als 16 Monate nach der Intervention...	31

TABELLENVERZEICHNIS

Tabelle 1: Suchstrategie vom 29. September 2008 im Pubmed.....	12
Tabelle 2: Suchstrategie vom 29. September 2008 in Cochrane central.....	13
Tabelle 3: Suchstrategie vom 29. September 2008 in CINAHL.....	14
Tabelle 4: Ausgeschlossene Studien mit Begründung.	17
Tabelle 5: Bewertung auf der PEDro-Skala.....	18
Tabelle 6: Übersichtstabelle der eingeschlossenen Studien.....	20
Tabelle 7: Berechnungsformeln zum Poolen von Gruppen.	25
Tabelle 8: Anfangswerte der Messungen der Schmerzintensität.	26
Tabelle 9: Anfangswerte der Messungen der Disability.	29
Tabelle 10: Volumenübersichtstabelle mit Effektgrößen	33

ZUSAMMENFASSUNG

Hintergrund: Rückenschmerzen sind weit verbreitet und ziehen grosse sozioökonomische Folgen nach sich. Es herrscht Unklarheit über die Therapie bei chronischen unspezifisch lumbalen Rückenschmerzen. Wir stellen die Hypothese auf, dass mehr Behandlungsvolumen eine bessere Effektivität als niedriges Volumen aufweist, was aber noch nie im Rahmen einer systematischen Literaturübersicht untersucht wurde.

Studiendesign: Systematische Literaturübersicht.

Ziele: Wir wollen herausfinden, welchen Effekt das Therapievolumen auf Patienten mit chronischen unspezifisch lumbalen Rückenschmerzen hat. Wir untersuchen die Outcomes Schmerzintensität, Disability und arbeitsunfähige Tage.

Methoden: Die systematische Suche nach RCTs erfolgte in den Datenbanken Pubmed, Cochrane central, CINAHL und PEDro. Die Suche wurde Ende 2008 zum letzten Mal aktualisiert. Mit Hilfe der definierten Einschlusskriterien, des Prüfens der Referenzen und der Kontaktaufnahme mit Experten konnten wir acht Studien einschliessen, die niedriges mit hohem Interventionsvolumen vergleichen. Die Outcomes Schmerzintensität und Disability haben wir statistisch ausgewertet, die Anzahl arbeitsunfähige Tage qualitativ beschrieben.

Resultate: Für die Schmerzintensität und Disability zeigt sich eine Tendenz zugunsten der grossvolumigen Gruppen, wobei nur zwei Studien ein signifikantes Ergebnis aufweisen. Die Anzahl arbeitsunfähiger Tage haben bei den klein- und grossvolumigen Gruppen abgenommen, wobei keine signifikanten Unterschiede zwischen niedrigem und hohem Volumen beschrieben sind.

Schlussfolgerung: Wir können kein signifikantes Ergebnis zugunsten der grossvolumigen Gruppe vorweisen, die Tendenz geht jedoch in diese Richtung.

Schlüsselwörter: chronische unspezifisch lumbale Rückenschmerzen, Interventionsvolumen, Rehabilitation, Schmerzintensität, Disability, Arbeitsfähigkeit, Krankheitstage, systematische Literaturübersicht.

RÉSUMÉ

Cadre théorique: Les lombalgies sont très répandues et ont d'importantes conséquences socioéconomiques. Le traitement des lombalgies chroniques non-spécifiques est néanmoins discuté. Nous partons de l'hypothèse suivante : un volume de traitement important a une meilleure efficacité qu'un petit volume. Ce sujet n'a encore jamais été étudié par une revue systématique.

Design de l'étude: Revue systématique de la littérature.

Buts: Analyser l'effet du volume d'intervention sur des patients présentant des lombalgies chroniques non-spécifiques. Nous observons cet effet sur la douleur, l'incapacité au quotidien (disability) et les jours d'incapacité au travail.

Méthodes: Nous avons effectué la recherche systématique de RCTs dans les banques de données Pubmed, Cochrane central, CINAHL et PEDro et l'avons actualisée pour la dernière fois fin 2008. Par des critères d'inclusion prédéfinis, la vérification des références et la prise de contact avec les auteurs, nous avons pu inclure huit études dans notre revue. La douleur et l'incapacité au quotidien (disability) ont été analysées statistiquement et les jours d'incapacité au travail ont été décrits qualitativement.

Résultats : En ce qui concerne la douleur et l'incapacité au quotidien (disability), seul deux études apportent un résultat significatif. Malgré cela, nous pouvons montrer une tendance allant vers les groupes à haut volume d'intervention. Pour ce qui est des jours d'incapacité au travail, leur nombre a diminué dans les deux groupes. Il n'y a néanmoins pas de différence significative entre les groupes à haut et à bas volume.

Conclusion: Nous ne pouvons pas présenter de résultat significatif en faveur du haut volume d'intervention, la tendance va néanmoins dans ce sens.

Mots-clés: lombalgies chroniques non-spécifiques, volume d'intervention, réadaptation, douleur, incapacité au quotidien (disability), capacité de travail, jours de maladie, revue systématique.

1 EINFÜHRUNG

1.1 Verbreitung von Rückenschmerzen

Das sehr häufige Symptom lumbaler Rückenschmerzen haben 60-90% der Erwachsenen mindestens einmal im Laufe ihres Lebens. Bei 80-90% der Betroffenen klingen die Beschwerden meist innerhalb von sechs Wochen wieder ab. Dennoch haben lumbale Rückenschmerzen eine enorme sozioökonomische Bedeutung. In den USA sind sie die häufigste Ursache einer Arbeitsunfähigkeit und eine der häufigsten Ursachen für einen Arztbesuch. Dabei verursachen die 10% der Patienten, bei denen sich die Beschwerden chronifizieren, 90% der Kosten. Die direkten Kosten machen etwa ein Drittel und die indirekten Kosten zwei Drittel der Gesamtkosten aus. Die direkten Kosten setzen sich aus Diagnostik und Therapie zusammen. Die indirekten Kosten entsprechen der Wertschöpfungsminderung durch Arbeitsausfall.[54]

In Deutschland betragen die Totalkosten für Rückenschmerzen durchschnittlich 1322 € pro Person und pro Jahr. Wenn man dieses auf die Gesamtpopulation Deutschlands zwischen 18 und 75 Jahren (ca. 61,8 Mio. Menschen) ausrechnet, ergibt es Gesamtkosten für Rückenschmerzen von 48,96 Milliarden Euro. Dies entspricht 2,2% des Bruttoinlandproduktes. Die 1322 € teilen sich ziemlich gleichmäßig in direkte (46%) und indirekte Kosten (54%) auf.[61]

Die durch lumbale Rückenschmerzen verursachten Kosten stellen für alle westlichen Gesellschaften ein wirtschaftliches Problem dar. Die Wertschöpfungsminderungskosten in den westlichen Ländern betragen schätzungsweise 84%-96% der Gesamtkosten für lumbale Rückenschmerzen.[26, 57]

Den vorher genannten Ausführungen kann man entnehmen, dass die Prävalenz und der Kostenanteil von chronischen unspezifisch lumbalen Rückenschmerzen nicht in allen Ländern gleich sind. Verschiedene Autoren haben deren Vorkommen untersucht. Volinn 1997[59], Andersson 1999[5] und Loney et Stratford 1999[32] erwähnen verschiedene Gründe, warum dieses von Land zu Land variiert. Eine Auswahl dieser Begründungen ist: Geschlecht, Alter, Hautfarbe, Kultur, soziale Schicht und Verdienst der untersuchten Population, sozioökonomische Lage und Sozialsystem des untersuchten Landes, nicht einheitliche Diagnostik und Definition von lumbalen Rückenschmerzen, Industrialisierung (daraus folgt langes Sitzen oder wenig Bewegung)

und unterschiedliche Arbeitsbedingungen (z.B. Unzufriedenheit, Leistungsdruck oder monotone Arbeit).

1.2 Persönliche Motivation

Wir möchten mit unserer Arbeit herausfinden, welchen Einfluss das Behandlungsvolumen auf Patienten mit chronischen unspezifisch lumbalen Rückenschmerzen hat. Der Begriff Volumen beinhaltet die Dauer (Zeit) und die Häufigkeit (Frequenz) der einzelnen Behandlungen. Die Frequenz setzt sich aus der Anzahl Behandlungen in einem bestimmten Zeitrahmen zusammen. Dieser entspricht in den meisten Rehabilitationsprogrammen bis zu ungefähr acht Wochen.

Uns interessiert, ob ein grosses Behandlungsvolumen ein anderes Ergebnis aufweist als ein kleineres Volumen. Dabei geht es uns nicht darum, stationäre versus ambulante Therapie zu untersuchen, da diese nicht mit einem bestimmten Volumen gleichzusetzen sind. In internationalen Studien werden auch grossvolumige Massnahmen ambulant durchgeführt.[7, 30, 51] Unser Ziel ist es also herauszufinden, was das Therapievolumen im Allgemeinen bei chronischen unspezifisch lumbalen Rückenschmerzen für eine Rolle spielt.

Für uns ist es nebensächlich, welche Behandlungsarten angewandt werden. Es geht uns nicht darum, verschiedene Interventionen miteinander zu vergleichen. Wir möchten uns auch nicht auf eine Intervention beschränken, sondern herausfinden, wie lange und wie häufig man in einem bestimmten Zeitrahmen behandeln sollte, bis in Bezug auf die von uns ausgewählten Outcomes eine Veränderung eintritt. Laut den „European Guidelines for the management of chronic nonspecific low back pain“ [1] gibt es keinen Goldstandard für die Behandlung von chronischen unspezifisch lumbalen Rückenschmerzen. Man kann nicht eindeutig sagen, welche Therapie für welche Beschwerden bei chronischen unspezifisch lumbalen Rückenschmerzen nützlich ist. Aus diesem Grund erlauben wir uns, nicht näher auf die Interventionen einzugehen und sie in grossvolumige und kleinvolumige Gruppen einzuteilen.

Es gibt diverse RCTs, welche die Outcomes Schmerz, Disability und Arbeitsunfähigkeit behandeln, indem sie verschiedene Interventionen miteinander vergleichen.[6, 57] Wir haben bis jetzt jedoch keine Studie gefunden, welche diese Outcomes in Bezug auf das Therapievolumen behandelt. Aus diesem Grund finden wir es sehr interessant, dies zu untersuchen.

2 PROBLEMATIK

2.1 Unspezifisch lumbale Rückenschmerzen

Wissenschaftler, Therapeuten, Ärzte und Patienten haben nicht die gleiche Auffassung von unspezifisch lumbalen Rückenschmerzen. Es liegt eine grosse Diskrepanz vor, da dieser Begriff mehrdeutig ist. Dies stellt ein Problem bei der wissenschaftlichen Arbeit dar, weil die Gefahr besteht, dass man dadurch „Äpfel“ mit „Birnen“ vergleicht.[14, 16] Im Allgemeinen definiert man unspezifisch lumbale Rückenschmerzen als Schmerz und Unbehagen zwischen der untersten Rippe und der inferioren Glutealfalte, mit oder ohne Ausstrahlung ins Bein. Rückenschmerzen werden in drei Kategorien aufgeteilt, die so genannte „diagnostic triage“: spezifische spinale Pathologien, radikuläre Schmerzen und nicht spezifische lumbale Rückenschmerzen.[1] Heute werden Rückenschmerzen in der Literatur vermehrt in mechanische und nicht-mechanische Rückenschmerzen eingeteilt, wobei die mechanischen nochmals in spezifisch und unspezifisch unterteilt werden.[54] Wir lassen uns jedoch nicht auf diese Definition ein, weil sie momentan noch nicht weit verbreitet ist.

Unspezifisch lumbale Rückenschmerzen entstehen durch ein komplexes Zusammenspiel vieler verschiedener Strukturen. Daher ist es oft nicht möglich, eine einzelne anatomische Struktur für die Beschwerden verantwortlich zu machen. Die lumbosakrale Region spielt eine Schlüsselrolle bei der Übertragung des Gewichts vom oberen Rumpf auf die Beine, sowohl in statischer Position wie auch in Bewegung. Bindegewebe, Knochen, Muskeln, Bänder und Nerven bilden ein komplexes, ineinander übergreifendes System. Muskeln setzen zum Teil an Bändern an, Ligamente verbinden knöcherne Strukturen miteinander und Nerven ziehen zwischen diesen verschiedenen Strukturen hindurch. Dieses System muss gleichzeitig zwei grundsätzlich gegensätzliche Eigenschaften aufweisen: Stabilität und Mobilität. Somit wird klar, dass sich ein Problem einer Struktur auf die umliegenden Strukturen auswirken kann. Ein solches strukturelles Ungleichgewicht kann eine Funktionseinschränkung zur Folge haben. Traumatische, entzündliche oder degenerative Prozesse können die verschiedenen Strukturen beeinflussen.[58]

Es ist wichtig zu wissen, dass selten nur ein Syndrom mit seinen Symptomen auftritt. Hier liegt die Schwierigkeit der Differentialdiagnose. Es ist wichtig, alle Symptome wie

auch die verschiedenen Krankheitsbilder zu kennen, um der eventuellen Ursache der Rückenschmerzen auf die Spur zu kommen.[58]

Neben den oben aufgeführten biologischen Faktoren spielen auch psychosoziale Aspekte eine wichtige Rolle. Je nach Lebenssituation ist die Wahrscheinlichkeit, unter Rückenschmerzen zu leiden, anders. Zudem beeinflussen kulturelle Einflüsse, persönliche Eigenschaften, soziales Umfeld und körperliche Aktivitäten unsere Gesundheit.[52, 60]

2.2 Chronizität

Die Klassifikation der Chronizität der lumbalen Rückenschmerzen ist ebenfalls nicht einheitlich. In zeitlicher Hinsicht werden für wissenschaftliche Zwecke oft akute (maximal 4-6 Wochen), subakute (6-12 Wochen), chronische (über 12 Wochen) und rezidivierende Lumbalgien unterschieden.[54] In den meisten von uns gelesenen Studien werden Patienten als chronisch lumbale Rückenschmerzpatienten aufgenommen, wenn sie seit mindestens drei Monaten daran leiden.[14, 22, 57] Darum beschränken wir uns auf diese Definition. Dabei sind wir uns im Klaren, dass die Chronizität nicht nur einen zeitlichen Aspekt hat. Chronizität ist ein individuelles Zusammenspiel von physischen und psychischen Gegebenheiten, einer Krankheitsgeschichte mit vielen Therapien und häufig schwierigen Arbeitsbedingungen (körperliche Leistung und/oder Unzufriedenheit am Arbeitsplatz).[14]

Die Psyche und das soziale Umfeld („Yellow Flags“) spielen bei der Chronifizierung eine grosse Rolle. Die „Yellow Flags“ [Anhang A] sind psychosoziale Faktoren, welche das Risiko, an einer langzeitigen Disability oder Arbeitsunfähigkeit zu erkranken, erhöhen. Dies ist im Zusammenhang mit chronischen unspezifisch lumbalen Rückenschmerzen zu verstehen. Die „Red Flags“ beschreiben die physischen Faktoren. Diese sind als medizinische Warnzeichen zu betrachten.[44]

2.3 Analyisierte Outcomes in unserer Review

2.3.1 Schmerz und Disability

Schmerz ist ein Symptom und nicht ein klinisches Zeichen, eine Diagnose oder eine Krankheit. Schmerz ist eine unangenehme sensorische und emotionale Erfahrung, die mit einer aktuellen oder potenziellen Gewebeschädigung verknüpft ist, oder mit Begriffen einer solchen Schädigung beschrieben wird. Schmerz ist immer subjektiv und dadurch nicht direkt messbar.[1]

Die Neurophysiologie des Schmerzes ist sehr komplex und für uns Physiotherapeuten wichtig zu kennen. Die Weiterleitung des Schmerzreizes setzt sich in Transduktion, Transmission, Modulation und Perzeption zusammen. Die Transduktion ist die Umwandlung der auf die Nozizeptoren wirkenden Reize (mechanische, chemische oder thermische) in elektrische Impulse. Unter Transmission versteht man die elektrische Weiterleitung der afferenten nozizeptiven Informationen im peripheren und zentralen Nervensystem. Die Modulation beinhaltet die Kontrolle und Bearbeitung dieser nozizeptiven Information durch spezifisch neuronale Hemmsysteme. Schliesslich ist die Perzeption die subjektive Schmerzwahrnehmung (bewusst und unbewusst) und die Schmerzäusserung (willkürlich und autonom).[19]

Laut Schumacher und Brähler (2002) haben Melzack und Wall im Jahr 1965 die „Gate-Control-Theorie“ formuliert. Diese integriert zusätzlich zu den physiologischen auch psychologische Faktoren der Schmerzentstehung. Die Theorie besagt, dass es auf mehreren Ebenen des schmerzverarbeitenden Systems zur Modulation nozizeptiver Reize kommt. Diese Modulation wird von zentralen Kontrollprozessen gesteuert, welche wiederum von einem motivational-affektiven System beeinflusst werden. Nach Melzack und Wall übernehmen Teile des Hinterhorns im Rückenmark die Funktion eines Tores, durch welches nicht alle Informationen weitergeleitet werden. Eine Schmerzempfindung entsteht erst dann, wenn die weitergeleitete nozizeptive Information eine bestimmte Schwelle überschreitet.[52]

Jedes Individuum lernt durch persönliche Erlebnisse im Laufe des Lebens den Begriff Schmerz anzuwenden. Darum interpretiert Schmerz nicht jeder gleich. Nicht alle Menschen gehen mit Schmerzen gleich um: Schmerz kann unterdrückt oder im Gegenteil die ganze Aufmerksamkeit darauf gerichtet werden. Dieses Verhalten wird von der Kultur wie auch von persönlichen Eigenschaften beeinflusst. Hinzu kommen Stresssituationen, bei denen Schmerzen stärker wahrgenommen werden.[1, 4]

Garrad und Bennett [20] definieren Disability als Leistungseinschränkung in einer oder mehreren Aktivitäten des täglichen Lebens, so dass diese Einschränkung eine Abhängigkeit zu einer anderen Person erfordert. Die WHO definiert Disability im Jahre 1980 folgendermassen: Disability ist eine Einschränkung (als Folge einer Beeinträchtigung) eine Alltagstätigkeit in der Art und Weise auszuführen, wie es für die Mehrheit der Bevölkerung als normal gilt. [62] Die alltäglichen Aktivitäten kann man in

verschiedene Bereiche einteilen: Mobilität (Gehen, Treppen steigen, Transfer in und aus dem Bett oder Stuhl, Reisen), Körperpflege (Ernährung, An- und Ausziehen, Intimpflege), Haushalt (Einkauf, Kochen, Putzen, Kleider waschen), Beruf (Arbeitsfähigkeit).[20, 62]

Disability wird auf Deutsch mit Wörtern wie *Behinderung*, *Funktionseinschränkung* oder *Unfähigkeit* übersetzt. Es existiert aber kein Wort, das die englische Bedeutung genau wiedergibt. In den von uns eingeschlossenen Studien wird Disability auch als *Function* bezeichnet. Wir haben uns entschieden, in unserer Arbeit den englischen Begriff Disability zu übernehmen.

Wie Schmerzen ist auch Disability bei Rückenschmerzpatienten häufig subjektiv. Sie kann aber einfacher objektiviert werden als Schmerzen, da konkrete Alltagssituationen beschrieben werden können.[60]

Es ist wichtig, Schmerz und Disability zu unterscheiden, da sie nicht zwingend in einem proportionalen Zusammenhang in Erscheinung treten. Ein grosses Ausmass an Disability kann von einer niedrigen Schmerzintensität verursacht werden. Und starke Schmerzen haben nicht unbedingt einen hohen Grad an Disability zur Folge. In der Praxis bedeutet das, dass Ärzte und Therapeuten mit einer rein schmerzorientierten Therapie möglicherweise die Situation des Patienten im Alltag nicht verändern. Es gibt einen Zusammenhang zwischen diesen beiden Begriffen, die Wissenschaft zeigt aber, dass dieser kleiner ist als bisher angenommen.[14, 60]

2.3.2 Anzahl arbeitsunfähige Tage

Schmerz und Disability können zu Arbeitsunfähigkeit führen. Die Tatsache, dass der Betroffene nicht arbeiten kann, kann zu sozialen und psychischen Problemen desjenigen führen. Zudem kommen durch den Arbeitsausfall Kosten für Arbeitgeber, Staat und Gesundheitsinstitutionen hinzu. Es handelt sich also auch um ein gesamtwirtschaftliches Problem.[60]

Disability wird oft mit Arbeitsunfähigkeit gleichgestellt. Diese zwei Begriffe haben aber eine unterschiedliche Bedeutung. Arbeitsunfähigkeit ist ein Teil der Disability, ist aber nicht genau das Gleiche. Die Beurteilung der Disability bezieht sich auf alltägliche Tätigkeiten, was nicht zwingend heisst, dass der Betroffene seiner Arbeit nicht nachgehen kann.[60] Aufgrund der unterschiedlichen Bedeutung von Disability und

Arbeitsunfähigkeit ist es uns wichtig, diese zwei Begriffe als Outcome getrennt zu analysieren.

Die Arbeitsunfähigkeit wird meistens mit der Anzahl arbeitsunfähiger Tage gemessen. Die Zeitspanne der Messung variiert von Studie zu Studie.

2.4 Begründung der Wahl unserer Outcomes

Zum Thema chronische unspezifisch lumbale Rückenschmerzen werden sehr viele verschiedene Outcomes beschrieben. Unsere Arbeit würde ein zu grosses Ausmass annehmen, wenn wir diese alle behandeln würden. Darum haben wir uns auf die drei oben beschriebenen Outcomes beschränkt.

Unserer Ansicht nach sind diese drei Begriffe sehr wichtig, da sie in den meisten Studien über chronische unspezifisch lumbale Rückenschmerzen vorkommen. Dies ermöglicht uns, Vergleiche zwischen den jeweiligen Studien zu erstellen. Die Anzahl arbeitsunfähiger Tage haben zudem eine enorme sozioökonomische Auswirkung.[39, 60] Die anderen in den Studien untersuchten Outcomes nehmen wir zur Kenntnis, analysieren sie aber nicht. (siehe Tabelle 6)

2.5 Biopsychosoziales Modell

Vor den 1980er Jahren standen für die Behandlung von chronischen Rückenschmerzen Bettruhe und andere passive Therapieformen im Vordergrund. Seit ungefähr 30 Jahren tendiert man immer mehr zu aktiven Behandlungsmassnahmen. Das Krankheitsbild lumbaler Rückenschmerzen wurde mit der fortschreitenden Chronifizierung mehr und mehr unter dem Einfluss psychosozialer Faktoren gesehen. Ungefähr im Jahre 1985 sprach unter anderem der amerikanische Orthopäde Tom Mayer zum ersten Mal von einer neuen Therapieform, die ganzheitlich auf den Patienten eingeht. Bei dieser verlagert sich der Behandlungsschwerpunkt von der symptomatischen Schmerzbehandlung zur Behandlung körperlicher, psychischer und sozialer Funktionen. Dieses biopsychosoziale Modell führte zu den „Functional Restoration“ Rehabilitationsprogrammen, deren Hauptziel die Reintegration in die Alltags- und Arbeitswelt war. Ein Aspekt dieser Programme ist das körperliche Training, in dem das Ziel die Steigerung der allgemeinen Fitness und die Verbesserung der kardiovaskulären und pulmonalen Kapazität ist. Hinzu kommen psychotherapeutische Interventionen, bei denen auf emotionale Bedürfnisse eingegangen wird. Es werden auch spezielle Therapien zur Erlernung eines adäquaten Krankheitsverhaltens in Bezug auf Aktivität

und Arbeitsfähigkeit angeboten.[39, 40, 46] Es existieren wissenschaftliche Studien über die Effizienz einer solchen interdisziplinären Rehabilitation. Diese zeigen, dass eine biopsychosoziale Rehabilitation eine Verbesserung der Schmerzsymptomatik wie auch der Disability erbringt.[6, 22, 43, 51]

Oft bedingt dieser biopsychosoziale Therapieansatz einen teuren stationären Aufenthalt. Die Studien von Bendix et al. 2000[11] in Skandinavien und Turk et al. 1996[55] in den USA zeigen jedoch auf, dass gesamtwirtschaftlich gesehen durch diese biopsychosozialen Therapieprogramme Kosten eingespart werden können. Dieses positive Ergebnis ist vor allem auf die Reduktion der Folgekosten (Wertschöpfungsminderung) von chronischen unspezifisch lumbalen Rückenschmerzen zurückzuführen.[11, 46, 55]

Trotz der Existenz der biopsychosozialen Rehabilitationsprogramme werden sie nicht konsequent und entsprechend der Guidelines eingesetzt. Das Angebot an passiven und aktiven Behandlungsmethoden ist sehr gross. Die Wissenschaft zeigt aber, dass die meisten angewandten Therapien bei Rückenschmerzen ineffizient sind.[22, 57] Die Unmenge von verschiedenen meist wenig effektiven Monotherapien und die effektiven multimodalen Therapien bezeugen, dass bei diesem Krankheitsbild ein multifaktorielles Geschehen zugrunde liegt. Es ist also schwierig herauszufinden, welche Therapie eine Verbesserung der Situation bei chronischen unspezifisch lumbalen Rückenschmerzen bewirkt.[1, 60]

Rückenschmerzen sind ein multifaktorielles und nicht nur ein individuelles, sondern vielmehr auch ein gesamtwirtschaftliches Problem. Dafür ist es wichtig, eine möglichst effiziente Behandlung beziehungsweise ein möglichst effizientes Behandlungsvolumen zu finden.

2.6 Hypothese

Wie schon beschrieben, untersuchen wir nicht den Effekt einer bestimmten physiotherapeutischen Intervention. Uns interessiert das gesamte Behandlungsvolumen bei der Rehabilitation chronischer unspezifisch lumbaler Rückenschmerzen. Die Behandlungen können sich aus verschiedenen Therapiebereichen zusammensetzen. Meist basieren sie heutzutage auf dem biopsychosozialen Modell.

Demzufolge lautet unsere Fragestellung: Effekt des Interventionsvolumens bei Patienten mit chronischen unspezifisch lumbalen Rückenschmerzen. Diesen Effekt untersuchen wir auf die Outcomes Schmerzintensität, Disability und Anzahl arbeitsunfähiger Tage. Wie es schon der Name sagt, haben chronische unspezifisch lumbale Rückenschmerzen keine spezifisch organische Ursache. Bei der Entstehung dieses Krankheitsbildes sind biologische, psychische und soziale Faktoren entscheidend. Daher erscheint es sinnvoll, diese Faktoren in die Behandlung mit einzubeziehen. Durch den Einschluss all dieser Faktoren in die Therapie nimmt das Behandlungsvolumen unserer Meinung nach zu. Aus diesem Grund nehmen wir an, dass ein hohes gegenüber einem tiefen Interventionsvolumen ein signifikant besseres Ergebnis der untersuchten Outcomes bewirkt.

3 METHODEN

3.1 Suchstrategie

Um unsere systematische Literaturübersicht zu machen, haben wir in vier verschiedenen elektronischen Datenbanken nach passenden Artikeln gesucht. Wir haben keinen zeitlichen Rahmen definiert, um keine Studie wegen ihres Alters auszuschliessen. Zusätzlich haben wir bei den gefundenen Artikeln die Referenzen geprüft und Experten kontaktiert, um so auf weitere Studien zu unserem Thema aufmerksam zu werden.

Um die Suche vorzubereiten, haben wir die Suchwörter in „Population (P)“, „Intervention (I)“, „Control (C)“ und „Outcome (O)“ unterteilt. Die nachfolgenden Begriffe dienten uns als Grundlage für die Suche. Die definitiven Suchwörter der Tabellen 1-3 haben sich aus dieser Liste herauskristallisiert.

P: Patients (-65) with unspecific chronic low back pain

I: Manual Therapy

Physical activity

Mobilisation

Manipulation

Muscle training

Rehabilitation

Aquatherapy

Physical therapy

Back school

Education

C: Intervention time

Intervention volume

Intervention dose

O: Back to work

Work status

Work readiness

Return to work

Sick leave days

Off work status

Pain

Function
Quality of life
Well-being
ADL
Disability

Das Interventionsvolumen definieren wir in den Einschlusskriterien. Die Gruppe Control haben wir bei der Suche weggelassen.

Am 28.04.2008 haben wir eine erste Suche im Pubmed gestartet. Anhand einer zufälligen Auswahl an Artikeln dieser Suche, haben wir uns ins Thema eingelese. Die definitive Suche führten wir am 29.09.2008 durch und kamen auf 832 Artikel. Am 26.12.2008 aktualisierten wir diese Suche ein letztes Mal. Wir haben die gleiche Suchstrategie angewandt wie am 29.09.2008. Dabei kamen wir auf 854 Treffer.

Nachstehend ist unsere Suchstrategie in den verschiedenen Datenbanken aufgelistet. Im Protokoll haben wir geschrieben, dass wir unsere Suche im Pubmed, Cochrane central und PEDro durchführen. Entgegen dieser Ausführungen haben wir auch noch im CINAHL gesucht, da in dieser Datenbank verschiedene physiotherapeutisch relevante Zeitschriften gelistet sind, die in Pubmed nicht enthalten sind.

3.1.1 Pubmedsuche

In Tabelle 1 ist die Suche im Pubmed dargestellt. Zuerst haben wir die Suchbegriffe für die Outcomes miteinander kombiniert (#1). Danach haben wir die Schlüsselwörter für die Population miteinander kombiniert (#2 bis #7). Bei der #8 haben wir die Begriffe für die Intervention kombiniert. Dann wurden alle Randomised Controlled Trials im Pubmed gesucht (#9). Wir haben Population, Outcomes und Intervention miteinander kombiniert (#10) und schlussendlich das Ganze mit den Limiten „RCT“ und „keine Kinder unter 18 Jahren“ kombiniert und so die definitive Suche gemacht (#12).

3.1.2 Cochrane central Suche

Da die Suche gleich aufgebaut ist wie beim Pubmed, beschreiben wir die einzelnen Schritte hier nicht. Wir haben das Vorgehen in Tabelle 2 dargestellt.

Suche Cochrane	Schlüsselwörter	Anzahl Treffer
#6	Search #5 nur Clinical trials	719
#5	Search #1 AND #2 AND #3 AND #4	999
#4	Search ((random*)) OR ((RCT)) OR ((randomised controlled trials))	440003
#3	Search ((((((((("Sick Leave"[Mesh])) OR ((Sick Leave))) OR ((back to work))) OR ((work status))) OR ((work readiness))) OR ((return to work))) OR ((work hardening))) OR ((work condition*)) OR (("Pain"[Mesh])) OR (("Pain Measurement"[Mesh])) OR ((Pain)) OR ((Pain Measurement)) OR ((off work status)) OR ((Work Capacity Evaluation)) OR ((Work Capacity Evaluation"[Mesh])) OR ((Disability Evaluation)) OR ((Disability)) OR ((functional status)) OR ((Quality of Life"[Mesh])) OR ((Quality of Life)) OR ((Personal Satisfaction"[Mesh])) OR ((Personal Satisfaction)) OR ((well being)) OR ((Rehabilitation, Vocational"[Mesh])) OR ((Rehabilitation, Vocational)) OR ((Vocational)) OR ((Activities of Daily Living"[Mesh])) OR ((Activities of Daily Living))	145141
#2	Search ((rehabilitation)) OR ((rehabilitation centers)) OR ((Rehabilitation program*))	18186
#1	Search ((sciat*)) OR ((lumbag*)) OR ((LBP)) OR ((back problem)) OR ((back pain)) OR ((low back pain))	6009

Tabelle 2: Suchstrategie vom 29. September 2008 in Cochrane central.

3.1.3 CINAHL Suche

Da es bei CINAHL nicht möglich war, mit Suchnummern zu arbeiten und diese zu kombinieren, haben wir hier die definitive Suche aus Cochrane central direkt angewandt. (siehe Tabelle 3)

Suche CINAHL	Suchwörter	Anzahl Treffer
#1	((randomised controlled trials) or (RCT) or (random*)) and ((rehabilitation*) or (rehabilitation program*) or (rehabilitation centers) or (rehabilitation)) and ((back pain) or (back problem) or (LBP) or (lumbag*) or (sciat*) or (low back pain)) and ((off work status) or (pain measurement) or (pain) or (work condition*) or (work hardening) or (return to work) or (work readiness) or (work status) or (back to work) or (sick leave) or (activities of daily living) or (vocational) or (well being) or (personal satisfaction) or (quality of life) or (functional status) or (disability) or (disability evaluation)) and (DT:NU = CLINICAL-TRIAL)	385

Tabelle 3: Suchstrategie vom 29. September 2008 in CINAHL.

3.1.4 PEDro Suche

Zum Schluss haben wir noch eine Suche in PEDro durchgeführt. Man kann in dieser Datenbank nur nach vorgegebenen Schlüsselwörtern suchen und die Ergebnisse müssen immer einzeln angeklickt werden, um zum Abstract zu kommen. Es ist relativ schwierig einen Überblick über die jeweiligen Artikel zu erhalten. Deshalb haben wir diese Suche ganz zum Schluss gemacht und konnten somit schnell feststellen, ob wir die Artikel bereits in den anderen Datenbanken gefunden hatten oder ob sie uns noch fehlten. Es sind dabei keine neuen Artikel dazu gekommen.

Zusammengefasst haben wir in allen vier Datenbanken 1936 Artikel gefunden (ohne Duplikate). Wir haben alle Titel und Abstracts ins Endnote Programm (Thomson, USA) exportiert, um damit weiterarbeiten zu können.

Beim Prüfen der Referenzen der gefundenen Studien konnten wir zusätzlich zwei Studien einschliessen.

3.2 Einschluss der Studien

Im Protokoll haben wir Ein- und Ausschlusskriterien definiert, mittels derer wir die gefundenen Artikel aussortieren konnten. Wir haben folgende sieben Einschlusskriterien definiert:

- a) Es müssen unspezifische lumbale Rückenbeschwerden vorhanden sein.
- b) Die Beschwerden müssen chronisch, also mehr als 3 Monate andauern.
- c) Es muss mindestens eine physiotherapeutische Intervention vorkommen.
- d) Es müssen Angaben in Stunden über mindestens zwei verschiedene Interventionsvolumen, die miteinander verglichen werden, vorhanden sein.
- e) Die Studie muss in Deutsch, Französisch oder Englisch geschrieben sein.
- f) Das Alter der Population muss zwischen 18 und 65 Jahren sein.
- g) Es muss sich um eine RCT handeln.

Ein Ausschlusskriterium sind Tierversuche.

Wir haben alle gefundenen Artikel ins Programm Endnote übertragen, wo wir direkt die Duplikate löschen konnten und eine übersichtliche Darstellung hatten. Anhand der Endnotedatei teilten wir uns die Studien auf. Eine Studienverfasserin las die Abstracts von 55% der Studien, die zweite Studienverfasserin die Abstracts von den anderen 55%. (siehe Abbildung 1) Die 10%, welche Beide gelesen haben, wurden miteinander verglichen. Falls es Unstimmigkeiten gab, diskutierten wir und versuchten einen Konsens zu finden. Wenn wir uns nicht einig waren, entschied unser Betreuer darüber, ob wir die Studie einschliessen oder nicht. Wenn es bei den 10%, die beide gelesen haben, zu vielen Unstimmigkeiten gekommen wäre, hätten beide alle Abstracts überprüft. Da dies nicht der Fall war, konnten wir die Aufteilung beibehalten.

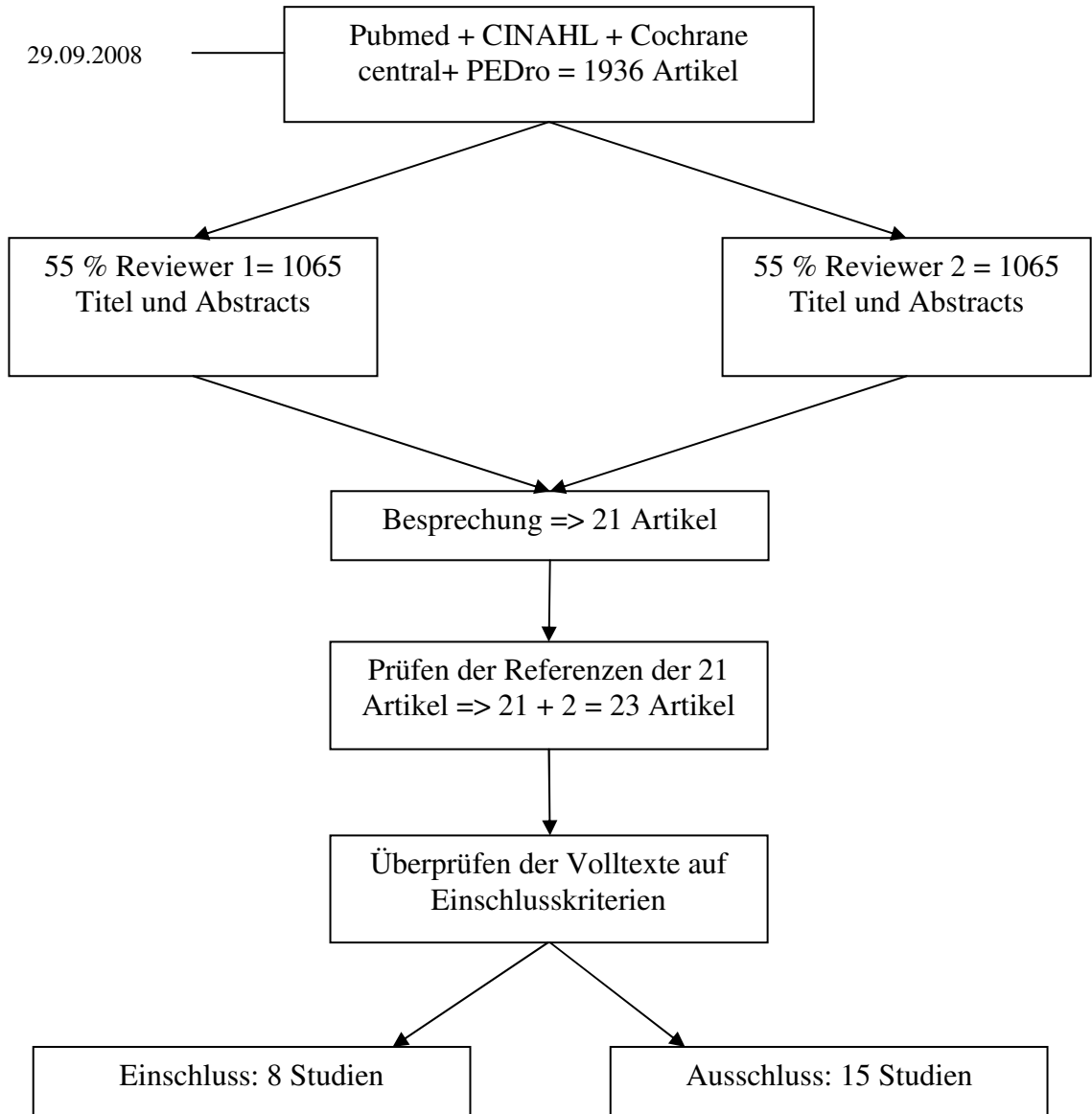


Abbildung 1: Flussdiagramm der Studien

Viele der 1936 Studien konnten bereits anhand der Ein- und Ausschlusskriterien beim ersten Screening ausgeschlossen werden. Es blieben 21 Studien übrig, bei denen wir die Referenzen überprüften. Dabei kamen zwei Studien hinzu. Die nun 23 Studien haben wir im Volltext gelesen. Dabei konnten wir erneut 15 Studien ausschliessen; die Gründe dafür sind in Tabelle 4 aufgeführt.

Studie	Ausschlussgrund
Bendix et al. 1998[8]	Die Beschwerden sind nicht unspezifisch.
Bendix et al. 1998[7]	Die Beschwerden sind nicht unspezifisch.
Callaghan 1994[13]	Es ist nicht klar ersichtlich, ob die Beschwerden unspezifisch sind.
Goldby et al. 2006[21]	Keine genauen Angaben über Interventionsvolumen.
Hansen et al. 1992[23]	Es wird nicht unterschieden zwischen chronisch und subchronisch.
Härkäpää et al. 1989[24]	Keine genauen Angaben über Interventionsvolumen; beide Gruppen erhalten gleiches Volumen an Therapie.
Härkäpää et al. 1990[25]	Keine genauen Angaben über Interventionsvolumen; beide Gruppen erhalten gleiches Volumen an Therapie.
Ludwig et al. 2006[33]	Keine genauen Angaben über Interventionsvolumen. Es ist nicht klar ersichtlich, ob die Beschwerden unspezifisch sind. Andere Definition von chronisch (weniger als 3 Monate).
Maleitzke et al. 1997[35]	Keine genauen Angaben über Interventionsvolumen.
Manniche et al. 1988[37]	Nur eine physiotherapeutische Intervention.
Mellin et al. 1989[42]	Keine genauen Angaben über Interventionsvolumen; beide Gruppen erhalten gleiches Volumen an Therapie.
Mellin et al. 1990[41]	Keine genauen Angaben über Interventionsvolumen; beide Gruppen erhalten gleiches Volumen an Therapie.
Rainville et al. 2002[47]	Keine RCT.
Skouen et al. 2002[53]	Die Beschwerden sind nicht unspezifisch.
Van der Roer et al. 2008[56]	Keine genauen Angaben über Interventionsvolumen.

Tabelle 4: Ausgeschlossene Studien mit Begründung.

Anhand der bereits beschriebenen Einschlusskriterien haben wir folgende acht Studien eingeschlossen: Alaranta et al. 1994[2], Bendix et al. 1997[9], Bendix et al. 1995[10], Bendix et al. 2000[11], Jousset et al. 2004[28], Kääpä et al. 2006[30], Roche et al. 2007[48] und Rose et al. 1997[50]. Davon stammen drei aus Dänemark, zwei aus Frankreich, zwei aus Finnland und eine aus England. Die untersuchte Population beträgt in den verschiedenen Studien zwischen 60 und 293 Patienten. Das Publikationsjahr der Studien erstreckt sich von 1994 bis 2007. (siehe Tabelle 6)

3.3 Bewertung der Studien

Im Protokoll haben wir geschrieben, dass wir die Qualität der von uns eingeschlossenen Studien mittels der PEDro-Skala und der Cochrane-risk-of-bias-Tabelle überprüfen. Wir haben diese zwei Bewertungsbögen gewählt, weil wir diese am besten kennen und weil sie in der Physiotherapie sehr häufig eingesetzt werden. Im Laufe der Arbeit haben wir gemerkt, dass die zwei Bewertungsbögen ähnliche Kriterien untersuchen. Wir finden die PEDro-Skala genauer, da zum Beispiel die Blindierung in die drei Bereiche Personal, Patient und Auswerter unterteilt wird. Bei der Cochrane-risk-of-bias-Tabelle werden diese Punkte unter einem Item zusammengefasst.

Wir konzentrieren uns daher vor allem auf die PEDro-Skala, welche in elf Items eingeteilt und für die methodologische Qualitätsüberprüfung von RCTs entworfen worden ist. Jedes positiv beantwortete Item, ausser dem ersten, wird zu einem Endresultat addiert. Das erste Item wird ausgeschlossen, weil es im Gegensatz zu den zehn anderen Items die externe Validität prüft. Die Skala reicht also von 0-10 Punkten. Die Skala wird in der PEDro-Datenbank regelmässig gebraucht, um die Qualität der eingehenden RCTs und systematischen Reviews zu überprüfen.[34]

Die zwei Studienverfasserinnen haben die Qualität der eingeschlossenen Studien separat beurteilt und später einen Vergleich durchgeführt. Falls keine Übereinstimmung gefunden wurde entschied eine dritte Person. In Tabelle 5 ist ersichtlich, dass kein Artikel weniger als fünf Punkte hat. Für uns sind also alle eingeschlossenen Studien von mittlerer bis guter Qualität.

Studie	Anzahl Punkte auf der PEDro-Skala
Bendix et al. 1997[9]	9 Punkte
Bendix et al. 1995[10]	8 Punkte
Kääpä et al. 2006[30]	8 Punkte
Roche et al. 2007[48]	8 Punkte
Bendix et al. 2000[11]	7 Punkte
Jousset et al. 2004[28]	7 Punkte
Alaranta et al. 1994[2]	7 Punkte
Rose et al. 1997[50]	5 Punkte

Tabelle 5: Bewertung auf der PEDro-Skala

In Abbildung 2 ist die Zusammenfassung der Qualitätsbewertung aller Studien nach der PEDro-Skala zu finden. Im Anhang B sind die entsprechenden Bewertungskriterien beschrieben.

	Eligibility criteria	Random allocation	Allocation concealment	Baseline comparability	Blinding of subjects	Blinding of therapists	Blinding of assessors	1 key outcome with >85%	Intention to treat analysis	Between-group comparisons	Point measures and measures of variability
Alaranta 1994	+	+	?	+	-	-	-	+	-	+	+
Bendix 1995	+	+	?	+	-	-	+	+	+	+	+
Bendix 1997	+	+	+	+	?	-	+	+	+	+	+
Bendix 2000	+	+	?	+	-	-	+	+	-	+	+
Jousset 2004	+	+	+	+	-	-	-	+	-	+	+
Kääpä 2006	+	+	+	+	-	-	?	+	+	+	+
Roche 2007	+	+	+	+	-	-	-	+	+	+	+
Rose 1997	+	+	?	-	-	-	?	+	-	+	+

Abbildung 2: Übersicht der Bewertung nach PEDro-Skala. Grün: Kriterium erfüllt, gelb: Unklar, rot: Kriterium nicht erfüllt.

In Abbildung 2 sind folgende Punkte ersichtlich. Die Einschlusskriterien sind bei allen Studien vorhanden. Die Zuteilung war überall geheim. Zufällig ist sie bei vier Studien; bei den vier anderen Studien ist dieser Punkt unklar. Bei sieben von acht Studien sind die Anfangswerte vergleichbar. Die Blindierung von Patienten und Therapeuten bei der Therapie von chronischen unspezifisch lumbalen Rückenschmerzen ist nicht möglich. In drei Studien waren die Auswerter blind. In allen Studien wurden die angekündigten Outcomes ausreichend untersucht. Bei vier Studien wurde die „Intention to treat-Analyse“ durchgeführt. Bei den anderen vier Studien ist die Behandlung der Ausfälle nicht beschrieben. Die statistischen Voraussetzungen sind bei allen Studien gegeben.

3.4 Datenextraktion

In Tabelle 6 sind die Daten der jeweiligen Studien aufgelistet.

Studie	Population	Volumenangaben	Follow-up	Definition Chronizität	1) Anzahl arbeitsunfähige Tage	2) Schmerzintensität	3) Disability	4) Andere Outcomes
Alaranta et al. 1994[2]	Randomisierung: 293 von 378 HV: N: 152 , drop outs:5/3 LV: N: 141 , drop outs:3	HV: (37h/Wo+5h/Wo)x3Wo = 123h +Heim 3Wo LV: ca. 40-50% von Akseli, 15-20h/Wox3Wo = 45-60h +passive Therapie	3 Wo vor Beginn, 3 Mte, 12 Mte	Pain duration of at least 6 months	0-30T, 31-90T, >90T / letzte 12 Mte bei Beginn (N) 0-7T, 8-30T, 31-200T, >200T / 12 Mte follow-up (in Prozent zu N)	0-100 pain and disability Million index,	0-100 pain and disability Million index,	Subjective physical performance and leisure time physical activities, use of medical care services, occupational handicap, psychological measures, Physical measures
Bendix et al. 1995[10]	Randomisierung: 132 HV: N: 46 , drop outs:6 LV1: N: 43 , drop outs:12 LV2: N: 43 , drop outs:8	HV: 39h/Wox3Wo+follow-up Programm 6h/Tx1T/Wox3Wo = 135h LV1: 2h/Tx2T/Wox6Wo = 24h LV2: 2h/Tx2T/Wox6Wo = 24h	4 Mte	A minimum of 6 months disabling low back trouble	Anzahl Tage (3 Jahre vor Studienbeginn und nach 4 Mte follow-up)	0-10 VAS	0-30 ADL assessment Manniche	Work readiness, contact with the health-care system, leg pain, physical activity.
Bendix et al. 1997[9]	Randomisierung: 132 HV: N: 46 , drop outs:8 LV1: N: 43 , drop outs:12 LV2: N: 43 , drop outs:9	HV: 135h LV1: 24h LV2: 24h	1 Jahr	A minimum of 6 months disabling low back trouble	Anzahl Tage in letzten 3 Jahren und in 1 Jahr follow-up (Median, IQR, Range)	0-10 VAS	0-30 ADL assessment Manniche	Work readiness, contact with the health-care system, leg pain, physical activity, medication.
Bendix et al. 2000[11]	Randomisierung: 138 HV: N: 64 , drop outs:11 LV: N: 74 , drop outs:12	HV: 39h/Wox3Wo = 117h LV: 1,5h/Tx3T/Wox8Wo = 36h	1 Jahr	???	Anzahl Tage (zu Beginn letzte 3 Jahre+nach 1 Jahr follow-up)	0-10 VAS	0-30 ADL assessment Manniche	Health care contact, leg pain, work capability, overall assessment.
Jousset et al. 2004[28]	Randomisierung: 86 HV: N: 44 , drop outs:2 LV: N: 42 , drop outs:1	HV: 6h/Tx5T/Wox5Wo = 150h LV: 1h/Tx3T/Wox5Wo +Heim 0,83h/Tx2T/Wox5Wo = 23,33h	6 Mte	???	Anzahl Tage (2 Jahre vor Beginn, 6 Mte follow-up)	0-10 VAS	0-100 Quebec Back Pain Disability scale.	Physical measures, return to work within 1 week, psychosocial parameters, treatment appreciation
Kääpä et al. 2006[30]	Randomisierung (nur Frauen): 132 von 143 =>2 drop outs HV: N: 64 , drop outs:5 LV: N: 66 , drop outs:5	HV: 6hx5T+8h/Wox5Wo = 70h LV: 10x1h/6-8Wo = 10h	direkt, 6, 12, 24 Mte	They had had daily or nearly daily LBP with or without sciatica during the preceding year	0 T, 1-30T, >30T / letzte 12 Mte	0-10 VAS	0-100 Oswestry Index	Sciatic pain intensity, subjective working capacity, health care consumption, depression, general well being, beliefs of working ability after 2 years
Roche et al. 2007[48]	Randomisierung: 132 HV: N: 68 , drop outs:1 LV: N: 64 , drop outs:0	HV: 6h/Tx5T/Wox5Wo = 150h LV: 1h/Tx3T/Wox5Wo +Heim 0,83h/Tx2T/Wox5Wo = 23,33h	nur direkt nach TTT	LBP for at least 3 months	Wurden zu Beginn der Studie für letzte 2 Jahre erfasst, danach nicht mehr.	0-10 VAS	0-100% Dallas pain Questionnaire	Tabelle mit qualitativen Messungen, ohne aber genaue Erklärungen im Text. Physical measures
Rose et al. 1997[50]	Randomisierung: 60 von 161 HV: N: 22 LV1: N: 22 LV2: N: 16 total drop outs:18	HV: 5x3h/T = 15h LV1: 5x6h/T = 30h LV2: 10x6h/T = 60h	10 Wo vor Beginn, bei Beginn, direkt danach, 6 Mte	Patients had experienced constant benign low back pain for 6 months or longer	-	0-100 VAS	0- 24 Roland Morris Disability Questionnaire	Psychologica l distress, depression, pain management

Tabelle 6: Übersichtstabelle der eingeschlossenen Studien. HV: high volume, LV: low volume, N: Population, h: Stunden, T: Tage, Wo: Wochen, Mte: Monate, VAS: visuelle Analogskala

3.5 Datenanalyse

Anhand der Tabelle 6 können wir herauslesen, ob und was wir miteinander vergleichen können. Das Outcome Anzahl arbeitsunfähige Tage beschreiben wir qualitativ, da wir die Resultate der einzelnen Studien nicht auf einen Nenner bringen konnten. Die Outcomes Schmerz und Disability werten wir statistisch aus.

Die Autoren der eingeschlossenen Studien verwenden verschiedene Follow-ups. Diese haben wir in vier Gruppen zusammengefasst, um die Arbeit übersichtlicher zu machen: direkt nach der Intervention, 1-8 Monate, 9-16 Monate und über 16 Monate danach. Auf Grund dieser verschiedenen Follow-ups können wir die Veränderung innerhalb einer Studie nicht mit der Veränderung innerhalb der anderen Studien vergleichen. Darum berechnen wir den Unterschied zwischen der grossvolumigen und der kleinvolumigen Gruppe zum Follow-up Zeitpunkt. Somit können wir auch einen Gesamteffekt zwischen mehreren Studien ermitteln.

Für die statistische Analyse haben wir mit dem Programm REVMAN (cochrane collaboration, Version 5.0.0) gearbeitet. Wir haben Mittelwert, Standardabweichung und Teilnehmer der beiden Gruppen zu den Follow-up Zeitpunkten eingefügt. Wenn wir die benötigten Werte nicht direkt aus den Studien herauslesen konnten, haben wir diese soweit möglich mit Formeln aus dem Revman-Handbuch [27] selbst ausgerechnet. Das Resultat der Analyse beinhaltet die Gewichtung der einzelnen Studien und den standardisierten Unterschied des Mittelwertes mit einem 95%-Konfidenzintervall zwischen den zwei Interventionsgruppen der einzelnen Studien und für den gepoolten Effekt.

3.6 Messinstrumente

3.6.1 Schmerz

Schmerz ist subjektiv, doch im Allgemeinen wird versucht, die Intensität mit Hilfe der visuellen Analogskala (VAS) zu objektivieren. Dies ist auch bei den von uns eingeschlossenen Studien der Fall. Es handelt sich um eine 10cm lange Linie, auf der der Patient zwischen den Extremen „überhaupt kein Schmerz“ und „maximal vorstellbarer Schmerz“ seine individuelle Schmerzintensität pro Schmerzgebiet einschätzen kann. Manchmal ist die Linie zusätzlich numerisch (0-10 oder 0-100) untergliedert. Es kommt auch vor, dass die Antworten verbal gestuft sind. Zu einer genauen Schmerzanamnese gehört die Erfassung der Häufigkeit, der Qualität, der Lokalisation, des Verhaltens über 24 Stunden und der Medikamenteneinnahme. Diese

Angaben werden bei der VAS nicht berücksichtigt und müssen zusätzlich erfragt werden.[29, 45] Die VAS wird von Oesch et al. 2007 als reliabel und valide beschrieben.[45]

3.6.2 Disability

Es existieren verschiedene Bewertungsbögen, welche den subjektiven Parameter Disability zu objektivieren versuchen. Die Autoren der von uns eingeschlossenen Studien haben mit unterschiedlichen Fragebögen gearbeitet, welche jeweils Alltagssituationen erfragen. Nachfolgend beschreiben wir diese genauer. Die Fragebögen sind in Anhang C-G aufgeführt.

Oswestry Disability Index

In zehn verschiedenen Bereichen kann der Patient aus sechs Möglichkeiten den für ihn zutreffenden Satz anstreichen und gibt somit Auskunft darüber, wie er mit seinen alltäglichen Tätigkeiten zurechtkommt. Jeder dieser sechs Sätze wird mit einer bestimmten Anzahl Punkte gewichtet. Je weniger Punkte der Patient erhält, desto selbstständiger kann er sein Leben führen. Folgende zehn Bereiche werden behandelt: Schmerzintensität, Körperpflege, Anheben, Gehen, Sitzen, Stehen, Schlafen, Sexualleben, Sozialleben, Reisen. Die Beantwortung des Fragebogens bezieht sich auf den Tag des Ausfüllens.

Die erreichte Gesamtpunktzahl wird in eine Formel eingefügt, um die Prozente der bestehenden Disability auszurechnen. Die Bewertung wird in fünf Gruppen eingeteilt, die von minimaler Disability bis hin zu schwerer Invalidität reichen.[17, 18, 29] [Anhang C]

Quebec Back Pain Disability Scale

Diese Skala beinhaltet zwanzig Aussagen, welche der Patient bewertet. Er kann auf einer Skala von 0 bis 5 das für ihn Zutreffende ankreuzen. 0 entspricht keiner Schwierigkeit und 5 der Unfähigkeit, diese Tätigkeit durchzuführen. Die Inhalte der Fragen liegen in den Bereichen Liegen, Schlafen, Reisen, Stehen, Laufen, Körperpflege, Bewegen, Anziehen, Essen und Heben. Am Schluss werden die einzelnen Punktzahlen addiert, um das Gesamtergebnis auszurechnen. Wer unter 15 Punkten liegt, gilt als selbstständig und nicht im Alltag behindert. Je höher die Summe ausfällt, desto höher ist der Grad der Disability.[18, 31] [Anhang D]

ADL-Assessment nach Manniche

In diesem Fragebogen werden 15 Fragen über die alltäglichen Tätigkeiten (Anziehen, Körperpflege, Laufen, Stehen, Sitzen, Schlafen, Essen) gestellt. Bei jeder Frage kann der Patient 0, 1 oder 2 Punkte erreichen. Je tiefer die zusammengezählte Punktzahl ist, umso weniger ist der Patient in seinem alltäglichen Leben eingeschränkt.[36] [Anhang F]

Roland-Morris Disability Questionnaire

Dieser Fragebogen besteht aus 24 Aussagen zum Funktionsstatus. Die Aussagen werden mit dem Satz: „wegen meines Rückens“ ergänzt. Die Beschwerden beziehen sich auf den Tag der Befragung. Die folgenden Eigenschaften werden berücksichtigt: Aktivitätsebene, Bewegung, Aktivitäten des täglichen Lebens, Essen und Schlafen. Die Anzahl angekreuzter Aussagen entspricht der Anzahl Punkte. Je mehr Punkte erreicht werden, umso höher ist der Grad der Disability.[49] [Anhang E]

Schlussfolgerung

Es wurde wissenschaftlich erwiesen, dass die genannten Fragebögen und Skalen die Veränderungen der Disability aufzeigen können. Laut Beurskens et al. 1995, Deyo et al. 1998 und Fritz et Irrgang 2001 sind diese reliabel, valide sowie änderungssensitiv.[12, 15, 18]

Die Fragebögen unterscheiden sich durch die Art der Fragestellung. Im Oswestry Disability Index werden zu jedem Bereich sechs verschiedene Sätze formuliert, von denen der Patient den für ihn zutreffenden anstreichen kann. Beim Roland-Morris Disability Questionnaire kann der Patient aus 24 möglichen Aussagen die für ihn zutreffenden ankreuzen. Im Quebec Back Pain Disability Scale hingegen werden die Aussagen stichwortartig zu jedem Bereich dargestellt. Der Patient kann jeder Aussage einen Score von 0 bis 5 zuteilen. Das ADL-Assessment nach Manniche ist nach dem selben Prinzip aufgebaut wie die Quebec Back Pain Disability Skala, ausser dass der Score von 0 bis 2 reicht.

Trotz dieser Unterschiede behandeln alle vier Fragebögen die gleichen Themen. Am Schluss werden jeweils die Punkte zusammengezählt und die Gesamtpunktzahl gibt Auskunft über den Grad der Einschränkungen im Alltag. Bei allen vier Skalen nimmt

die Disability mit steigender Punkt- oder Prozentzahl zu. Weil die Skalen alle auf dem gleichen Prinzip beruhen, erlauben wir uns, das Outcome Disability aus den verschiedenen Studien zu vergleichen, auch wenn nicht immer exakt der gleiche Fragebogen angewandt wurde.

3.6.3 Arbeitsunfähigkeitstage

Für die Messung der Arbeitsunfähigkeit ist die Anzahl der arbeitsunfähigen Tage weit verbreitet. In den von uns eingeschlossenen Studien werden die arbeitsunfähigen Tage in einem bestimmten Zeitraum gezählt. Dieser Zeitraum ist von Studie zu Studie unterschiedlich. Aus diesem Grund konnten wir die Resultate für dieses Outcome nicht direkt vergleichen und haben sie qualitativ beschrieben.

4 RESULTATE

Weil in den eingeschlossenen Studien die Werte über unterschiedliche Zeitspannen gemessen wurden, haben wir für die statistische Analyse die Ergebnisse nach Follow-up aufgeteilt. Somit ergeben sich für beide Outcomes vier verschiedene Gruppen. Das heisst, dass in jeder Gruppe nur die Studien mit dem entsprechenden Follow-up direkt verglichen werden.

Bei Bendix et al. 1995[10], Bendix et al. 1997[9] und Bendix et al. 2000[11] mussten wir die Standardabweichung für jedes Outcome selbst ausrechnen. Mit Hilfe des Revman-Handbuches [27] haben wir diese vom Interquartilbereich abgeleitet. Bendix et al. 1995[10], 1997[9] und 2000[11] haben jeweils nur den Median angegeben. Wir haben diesen mit dem Mittelwert gleichgestellt, da die Daten normalverteilt waren. Bei allen anderen eingeschlossenen Studien waren Standardabweichung und Mittelwert gegeben.

Bei Bendix et al. 1995[10] und Bendix et al. 1997[9] ist die Population in drei Gruppen aufgeteilt. Da die zwei Gruppen mit kleinerem Interventionsvolumen die gleiche Anzahl Stunden aufwiesen und die Art der Intervention für uns nicht relevant ist, haben wir diese zwei Gruppen zu einer Gruppe zusammengefasst. Für die Berechnungen haben wir die Formeln des Revman-Handbuches[27] verwendet. So konnten wir die Grösse der Gruppe, den Mittelwert und die Standardabweichung der gepoolten Low-Volume-Gruppe für die Outcomes Schmerzintensität und Disability ausrechnen.(siehe Tabelle 7)

Group	1	2	Combined groups
Sample size	N_1	N_2	$N_1 + N_2$
Mean	M_1	M_2	$\frac{N_1 M_1 + N_2 M_2}{N_1 + N_2}$
SD	SD_1	SD_2	$\sqrt{\frac{(N_1 - 1)SD_1^2 + (N_2 - 1)SD_2^2 + \frac{N_1 N_2}{N_1 + N_2} (M_1^2 + M_2^2 - 2M_1 M_2)}{N_1 + N_2 - 1}}$

Tabelle 7: Berechnungsformeln zum Poolen von Gruppen. SD: Standardabweichung

4.1 Outcome Schmerzintensität

In Tabelle 8 sind die Baseline-Daten der Schmerzintensität von den Studien der quantitativen Analyse aufgeführt.

Studie	Messinstrument	HV Mean	LV Mean
Bendix et al. 1995[10]	VAS	5.3	5.9
Bendix et al. 1997[9]	VAS	5.3	5.6
Bendix et al. 2000[11]	VAS	5.1	6.0
Jousset et al. 2004[28]	VAS	5.0	4.6
Kääpä et al. 2006[30]	VAS	4.6	5

Tabelle 8: Anfangswerte der Messungen der Schmerzintensität. Der Mean entspricht dem standardisierten Mittelwert. HV: high volume, LV: low volume, VAS: Visuelle Analogskala

4.1.1 Analyse direkt nach der Intervention

Kääpä et al. 2006[30] haben die Schmerzintensität direkt nach der Intervention getestet. Der Unterschied zwischen den Gruppen ist nicht signifikant. (siehe Abbildung 3)

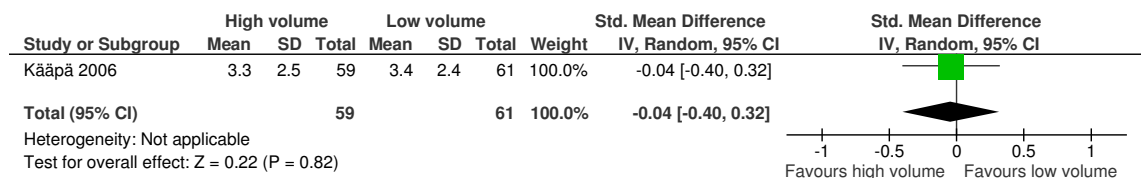


Abbildung 3: Forest Plot der Schmerzintensität direkt nach der Intervention. Grünes Quadrat: Gewichtung der Studie, Diamant: Mittlerer Effekt über alle Studien, SD: Standardabweichung, IV: Inverse Varianz, CI: Konfidenzintervall. High volume 70h, Low volume 10h

4.1.2 Analyse 1-8 Monate nach der Intervention

Drei Studien haben die Schmerzintensität in der Zeitspanne zwischen 1 und 8 Monate gemessen. Jousset et al. 2004[28] und Kääpä et al. 2006[30] zeigen keinen signifikanten Unterschied zwischen den Interventionsgruppen. Bendix et al. 1995[10] zeigen, dass die Patienten der grossvolumigen Gruppe nach vier Monaten signifikant weniger Rückenschmerzen haben als die Patienten der Gruppe mit dem kleineren Volumen. (siehe Abbildung 4)

Die Anfangswerte der drei untersuchten Studien [10, 28, 30] weisen bei beiden Interventionsgruppen keine grossen Unterschiede auf. (siehe Tabelle 8)

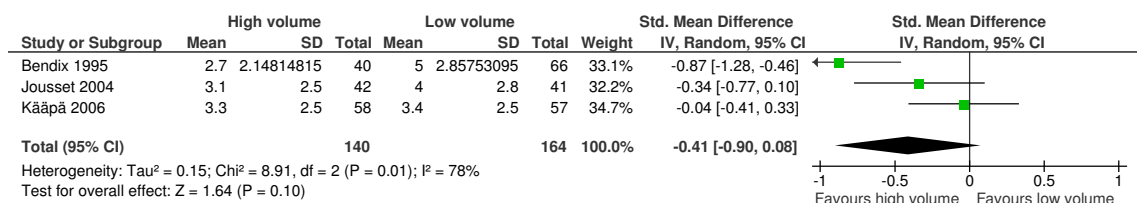


Abbildung 4: Forest Plot mit Vergleich der Schmerzintensität 1-8 Monate nach der Intervention. Grünes Quadrat: Gewichtung der Studie, Diamant: Mittlerer Effekt über alle Studien, SD: Standardabweichung, IV: Inverse Varianz, CI: Konfidenzintervall. Bendix 95: High volume(HV) 135h, Low volume(LV) 24h. Jousset: HV 150h, LV 23h. Kääpä: HV 70h, LV 10h

4.1.3 Analyse 9-16 Monate nach der Intervention

In dieser Zeitspanne haben drei Studien die Schmerzintensität untersucht. Bendix et al. 1997[9] zeigen einen signifikanten Unterschied zugunsten der grossvolumigen Gruppe auf. Bendix et al. 2000[11] und Kääpä et al. 2006[30] weisen keinen signifikanten Unterschied zwischen den Gruppen auf. (siehe Abbildung 5)

Die Anfangswerte der drei untersuchten Studien [9, 11, 30] sind bei beiden Interventionsgruppen genau gleich oder sehr ähnlich. (siehe Tabelle 8)

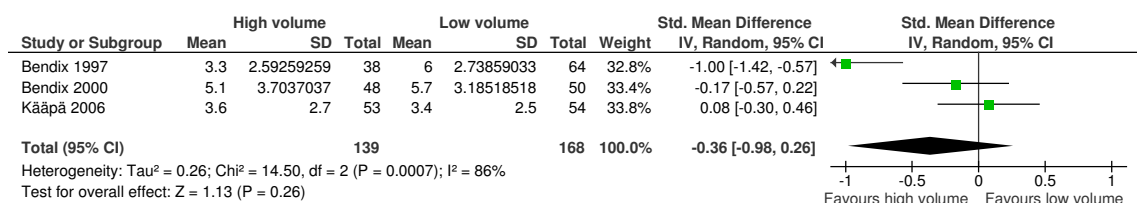


Abbildung 5: Forest Plot mit Vergleich der Schmerzintensität 9-16 Monate nach der Intervention. Grünes Quadrat: Gewichtung der Studie, Diamant: Mittlerer Effekt über alle Studien, SD: Standardabweichung, IV: Inverse Varianz, CI: Konfidenzintervall. Bendix 97: High volume(HV) 135h, Low volume(LV) 24h. Bendix 2000: HV 117h, LV 36h. Kääpä: HV 70h, LV 10h

4.1.4 Analyse mehr als 16 Monate nach der Intervention

Auch zu diesem Messzeitpunkt ist der Unterschied zwischen den zwei Gruppen bei Kääpä et al. 2006[30] nicht signifikant. (siehe Abbildung 6)

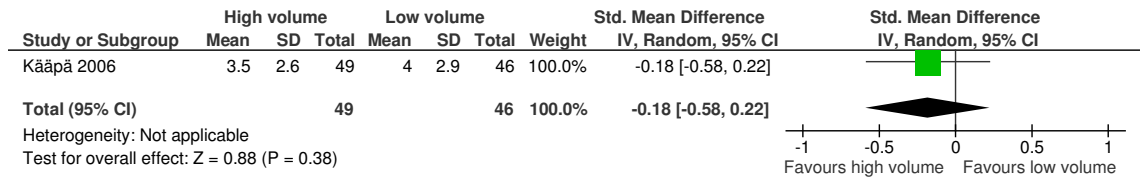


Abbildung 6: Forest Plot der Schmerzintensität mehr als 16 Monate nach der Intervention. Grünes Quadrat: Gewichtung der Studie, Diamant: Mittlerer Effekt über alle Studien, SD: Standardabweichung, IV: Inverse Varianz, CI: Konfidenzintervall. High volume 70h, Low volume 10h

Zusammenfassend sieht man, dass für das Outcome Schmerzintensität nur Bendix et al. 1995[10] und Bendix et al. 1997[9] einen signifikanten Unterschied zwischen den Gruppen zum jeweiligen Follow-up Zeitpunkt aufweisen.

4.2 Outcome Disability

Die Autoren der von uns eingeschlossenen Studien haben insgesamt vier verschiedene Messinstrumente für die Disability verwendet. Dies sind der „Oswestry-Index“ (eine Studie), die „Quebec-Back-Pain-Disability Skala“ (eine Studie), der „Roland-Morris Disability Questionnaire“ (eine Studie) und ein von Manniche et al. 1994[36] entwickeltes ADL-Assessment (drei Studien). Alaranta et al. 1994[2] verwenden den Pain and Disability Million Index. Diese Studie haben wir jedoch nicht in die statistische Analyse einbezogen, da Schmerz und Disability als ein Outcome zusammengefasst werden. Die Messinstrumente für die Disability sind im Methodenteil beschrieben und im Anhang ersichtlich. [Anhang C-G]

In Tabelle 9 sind die Baseline-Daten der Disability aufgeführt. Da die Skalen in den verschiedenen Studien entweder von 0 bis 100 oder von 0 bis 30 reichen, haben wir die Werte in Prozente umgerechnet, um sie vergleichbar zu machen.

Studie	Messinstrument	High volume		Low volume	
		Mean	Prozent	Mean	Prozent
Kääpä et al. 2006[30]	Oswestry Disability Index	25.4	25.4	23.8	23.8
Jousset et al. 2004[28]	Quebec Backpain Disability Scale	34.6	34.6	31.6	31.6
Bendix et al. 1995[10]	ADL-Assesment nach Manniche	15.5	51.6	15.3	50.9
Bendix et al. 1997[9]	ADL-Assesment nach Manniche	15	50	14.6	48.6
Bendix et al. 2000[11]	ADL-Assesment nach Manniche	16	53.3	16	53.3

Tabelle 9: Anfangswerte der Messungen der Disability. Der Mean entspricht dem standardisierten Mittelwert.

4.2.1 Analyse direkt nach der Intervention

Kääpä et al. 2006[30] haben die Disability direkt nach der Intervention untersucht und konnten zu diesem Zeitpunkt keinen signifikanten Unterschied zwischen den beiden Gruppen feststellen. (siehe Abbildung 7)

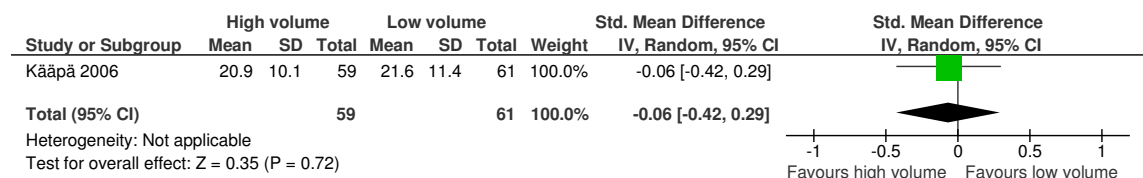


Abbildung 7: Forest Plot der Disability direkt nach der Intervention. Grünes Quadrat: Gewichtung der Studie, Diamant: Mittlerer Effekt über alle Studien, SD: Standardabweichung, IV: Inverse Varianz, CI: Konfidenzintervall. High volume 70h, Low volume 10h

4.2.2 Analyse 1-8 Monate nach der Intervention

Drei Studien haben die Disability in dieser Zeitspanne gemessen. Bei Kääpä et al. 2006[30] und Jousset et al. 2004[28] wird kein signifikanter Unterschied zwischen den Gruppen festgestellt. Bendix et al. 1995[10] hingegen zeigen einen signifikanten Unterschied der Gruppen zugunsten der grossvolumigen Gruppe auf. (siehe Abbildung 8)

Bei Bendix et al. 1995[10] entsprechen die Anfangswerte der Disability ca. der Hälfte des möglichen Scores. Verglichen mit Kääpä et al. 2006[30] entspricht dies fast dem Doppelten mehr und 15.4% bzw. 7.8% mehr als bei Jousset et al. 2004[28]. Bei Jousset et al. 2004[28] sind die Anfangswerte höher als bei Kääpä et al. 2006[30] (9.2% bzw.

7.8%). (siehe Tabelle 9) Das weite Spektrum der Anfangswerte beeinflusst den mittleren Effekt über alle Studien, nicht aber das Verhältnis zwischen den zwei Gruppen innerhalb einer Studie.

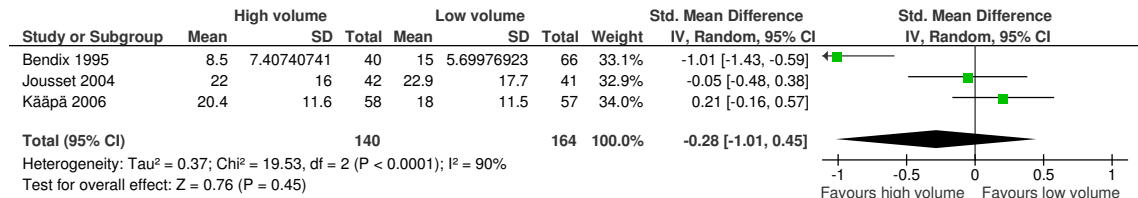


Abbildung 8: Forest Plot mit Vergleich der Disability 1-8 Monate nach der Intervention. Grünes Quadrat: Gewichtung der Studie, Diamant: Mittlerer Effekt über alle Studien, SD: Standardabweichung, IV: Inverse Varianz, CI: Konfidenzintervall. Bendix 95: High volume(HV) 135h, Low volume(LV) 24h. Jousset: HV 150h, LV 23h. Kääpä: HV 70h, LV 10h

4.2.3 Analyse 9-16 Monate nach der Intervention

In diesem Zeitrahmen haben drei von uns eingeschlossene Studien das Outcome Disability behandelt. Kääpä et al. 2006[30] und Bendix et al. 2000[11] zeigen keinen signifikanten Unterschied zwischen den Gruppen. Bendix et al. 1997[9] hingegen stellen einen signifikanten Unterschied zwischen den Gruppen zugunsten der Gruppe mit höherem Volumen fest. (siehe Abbildung 9)

In Bendix et al. 1997[9] und Bendix et al. 2000[11] sind die Anfangswerte ähnlich (Unterschied 3.3% bzw. 4.7%). Die Werte von Bendix et al. 1997[9] entsprechen fast dem Doppelten vom Anfangswert von Kääpä et al. 2006[30] (24.6% bzw. 24.8%). Auch beim Vergleich zwischen Bendix et al. 2000[11] und Kääpä et al. 2006[30] sieht man grosse Unterschiede (27.9% bzw. 29.5%). (siehe Tabelle 9) Das weite Spektrum der Anfangswerte beeinflusst den mittleren Effekt über alle Studien, nicht aber das Verhältnis zwischen den zwei Gruppen innerhalb einer Studie.

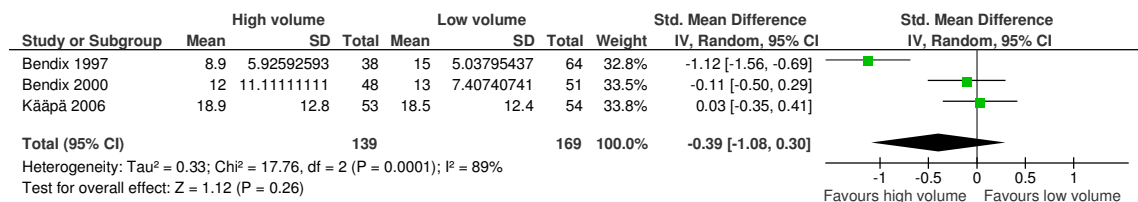


Abbildung 9: Forest Plot mit Vergleich der Disability 9-16 Monate nach der Intervention. Grünes Quadrat: Gewichtung der Studie, Diamant: Mittlerer Effekt über alle Studien, SD: Standardabweichung, IV: Inverse Varianz, CI: Konfidenzintervall. Bendix 97: High volume(HV) 135h, Low volume(LV) 24h. Bendix 2000: HV 117h, LV 36h. Kääpä: HV 70h, LV 10h

4.2.4 Analyse mehr als 16 Monate nach der Intervention

Kääpä et al. 2006[30] konnten nach zwei Jahren keinen signifikanten Unterschied zwischen den beiden Gruppen beschreiben. (siehe Abbildung 10)

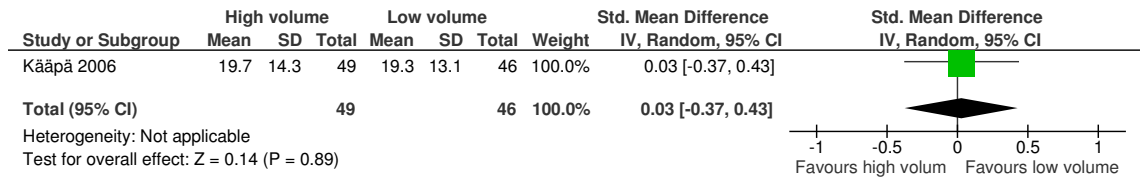


Abbildung 10: Forest Plot der Disability mehr als 16 Monate nach der Intervention. Grünes Quadrat: Gewichtung der Studie, Diamant: Mittlerer Effekt über alle Studien, SD: Standardabweichung, IV: Inverse Varianz, CI: Konfidenzintervall. High volume 70h, Low volume 10h

Zusammenfassend sieht man, dass für das Outcome Disability nur Bendix et al. 1995[10] und Bendix et al. 1997[9] einen signifikanten Unterschied zwischen den Gruppen zum jeweiligen Follow-up Zeitpunkt aufweisen.

4.3 Andere Studien

Die Studie von Rose et al. 1997[50] entspricht allen Einschlusskriterien. Sie behandelt unter anderem die Outcomes Schmerzintensität und Disability. Wir haben sie jedoch nicht in die quantitativen Analysen miteinbezogen, da keine genauen Werte vorliegen. Obwohl die Messinstrumente für statistische Berechnungen geeignet wären (VAS, Roland-Morris Disability Questionnaire), sind die Angaben der Outcomes ungenau. Sie sind als Graphiken dargestellt, aus deren Skalierung keine genauen Werte herausgelesen werden können. Es ist also nicht möglich, diese Studie direkt mit den anderen zu vergleichen. Aus diesem Grund beschreiben wir die Resultate dieser Studie qualitativ für die Outcomes Schmerzintensität und Disability. Dabei berücksichtigen wir die Gruppen mit 15-stündiger und 60-stündiger Intervention.

Die Schmerzintensität nimmt innerhalb von sechs Monaten bei beiden Gruppen ab. Bei der grossvolumigen Gruppe nimmt die Schmerzintensität auf der VAS von 6.0 auf 5.1, bei der kleinvolumigen Gruppe von 4.8 auf 4.3 ab. Beide Abnahmen sind nicht signifikant. Die Disability nimmt innerhalb von sechs Monaten bei beiden Gruppen signifikant ab. Sie verändert sich in der grossvolumigen Gruppe von 12.4 auf 9.4 und in der kleinvolumigen Gruppe von 12.4 auf 10.0 Punkte auf dem Roland-Morris Disability Questionnaire.

Die Anfangswerte der beiden Gruppen für die beiden Outcomes unterscheiden sich nicht signifikant. Über die Signifikanz des Unterschiedes zwischen beiden Gruppen am Schluss des Follow-ups können wir keine Aussagen machen, da genaue Angaben fehlen.

Alaranta et al. 1994[2] haben die Outcomes Schmerz und Disability mit dem Pain and Disability Million Index [3] untersucht. Dieser Fragebogen beurteilt die beiden Outcomes zusammen; es steht also keine separate Analyse von Schmerz und Disability zur Verfügung. Aus diesem Grund können wir die Werte nicht in unsere Berechnungen miteinbeziehen. Laut Alaranta et al. 1994[2] nimmt der Million Index nach drei und zwölf Monaten bei beiden Gruppen ab. Die Abnahme ist bei der Gruppe mit dem höheren Volumen signifikant grösser.

Roche et al. 2007[48] haben mit der visuellen Analogskala die Schmerzintensität untersucht. Da in der Studie nur die Differenzen zwischen zwei Zeitpunkten angegeben sind, können wir diese trotz des gleichen Messinstrumentes nicht in unsere statistische Analyse einschliessen. Laut den Autoren ist die Schmerzintensität direkt nach der Intervention (fünf Wochen) in beiden Gruppen signifikant gesunken. Der Unterschied zwischen den beiden Gruppen ist nicht signifikant.

4.4 Zusammenfassung der Outcomes Schmerzintensität und Disability

In Tabelle 10 sind die statistischen Daten nochmals zusammenfassend aufgeführt. Die Autoren der Studien haben im Gegensatz zu uns den Verlauf untersucht. Deren Resultate sind auch in nachfolgender Tabelle zu finden.

Studie	HV	LV	Effektgrösse SMD [95% KI]	Veränderung innerhalb der Gruppe (Verlauf)		Unterschied des Verlaufs zwischen den beiden Gruppen
				HV	LV	
Alaranta et al. 1994[2]	123h	45-60h	Keine Angaben	44.8 auf 26.8 s ¹ 44.8 auf 28.4 s ²	44.8 auf 34.3 s ¹ 44.8 auf 34.3 s ²	s zg HV ¹ s zg HV ²
Bendix et al. 1995[10]	135h	24h	-1.01 [-1.43 bis -0.59] s ¹ -0.87 [-1.28 bis 0.46] s ²	15.5 auf 8.5 ka ¹ 5.3 auf 2.7 ka ²	15.3 auf 15.0 ka ¹ 5.9 auf 2.8 ka ²	s zg HV ¹ s zg HV ²
Bendix et al. 1997[9]	135h	24h	-1.12 [-1.56 bis -0.69] s ¹ -1.00 [-1.42 bis -0.57] s ²	15.0 auf 8.9 s ¹ 5.3 auf 3.3 s ²	14.6 auf 15.0 ns ¹ 5.6 auf 2.7 ns ²	s zg HV ¹ s zg HV ²
Bendix et al. 2000[11]	117h	36h	-0.11 [-0.50 bis 0.29] ns ¹ -0.17 [-0.57 bis 0.22] ns ²	16.0 auf 12.0 s ¹ 5.1 auf 5.1 ns ²	16.0 auf 13.0 s ¹ 6.0 auf 5.7 ns ²	ns ¹ ns ²
Joussset et al. 2004[28]	150h	23.33h	-0.05 [-0.48 bis 0.38] ns ¹ -0.34 [-0.77 bis 0.10] ns ²	34.6 auf 22.0 s ¹ 5.0 auf 3.1 s ²	31.6 auf 22.9 ns ¹ 4.6 auf 4.0 ns ²	ns ¹ ns ²
Kääpä et al. 2006[30]	70h	10h	0.03 [-0.37 bis 0.43] ns ¹	25.4 auf 19.7 s ¹	23.8 auf 19.3 s ¹	ns ¹
			0.03 [-0.35 bis 0.41] ns ¹	25.4 auf 18.9 s ¹	23.8 auf 18.5 s ¹	ns ¹
			0.21 [-0.16 bis 0.57] ns ¹	25.4 auf 20.4 s ¹	23.8 auf 18.0 s ¹	ns ¹
			-0.06 [-0.42 bis 0.29] ns ¹	25.4 auf 20.9 s ¹	23.8 auf 21.6 s ¹	ns ¹
			-0.18 [-0.58 bis 0.22] ns ²	4.6 auf 3.5 ns ²	5.0 auf 4.0 s ²	ns ²
			0.08 [-0.3 bis 0.46] ns ²	4.6 auf 3.6 ns ²	5.0 auf 3.4 s ²	ns ²
			-0.04 [-0.41 bis 0.33] ns ²	4.6 auf 3.3 ns ²	5.0 auf 3.4 s ²	ns ²
			-0.04 [-0.40 bis 0.32] ns ²	4.6 auf 3.3 ns ²	5.0 auf 3.4 s ²	ns ²
Roche et al. 2007[48]	150h	23.33h	Keine Angaben	4.7 auf 2.8 s ²	4.5 auf 3.0 s ²	ns ²
Rose et al. 1997[50]	60h	15h	Keine Angaben	12.4 auf 8.4 s ¹ 12.4 auf 9.4 s ¹ 6.0 auf 5.0 ns ² 6.0 auf 5.1 ns ² *	12.3 auf 9.0 s ¹ 12.4 auf 10.0 s ¹ 4.8 auf 4.9 ns ² 4.8 auf 4.3 ns ² *	Interaktion Zeit- Volumen : keine Angaben. Haupteffekt Volumen: ns

Tabelle 10: Volumenübersichtstabelle mit Effektgrössen (Standardised Mean Difference), entsprechenden 95% Konfidenzintervallen (95% KI) und Verlaufsdaten. HV: high volume, LV: low volume, zg: zugunsten, s: signifikant, ns: nicht signifikant, ¹ Disability, ² Schmerzintensität, ka: keine Angaben über die Signifikanz, * Werte wurden anhand der Grafik geschätzt

In Tabelle 10 sieht man, dass für das Outcome Schmerzintensität im Verlauf vier Studien (Alaranta et al. 1994[2], Bendix et al. 1997[9], Jousset et al. 2004[28], Roche et al. 2007[48]) eine signifikante Verbesserung der Gruppe mit grossem Volumen zeigen. Die jeweiligen kleinvolumigen Gruppen verzeichnen auch eine Abnahme; diese ist in zwei Fällen signifikant (Alaranta et al. 1994[2], Roche et al. 2007[48]), in den anderen zwei Studien nicht (Bendix et al. 1997[9], Jousset et al. 2004[28]). Die Follow-ups der genannten Studien sind verschieden lang.

Betreffend Disability verzeichnen wir sechs Studien (Alaranta et al. 1994[2], Bendix et al. 1997[9], Bendix et al. 2000[11], Jousset et al. 2004[28], Kääpä et al. 2006[30], Rose et al. 1997[50]) mit einer signifikanten Verbesserung im Verlauf für die Gruppe mit dem grossen Volumen. Davon nimmt die Disability auch bei vier kleinvolumigen Gruppen signifikant ab (Alaranta et al. 1994[2], Bendix et al. 2000[11], Kääpä et al. 2006[30], Rose et al. 1997[50]). Bei der kleinvolumigen Gruppe in Bendix et al. 1997[9] nimmt die Disability minim zu (nicht signifikant) und bei Jousset et al. 2004[28] können wir eine nicht signifikante Abnahme verzeichnen. Die Follow-ups der genannten Studien sind unterschiedlich.

Für das Outcome Schmerzintensität wie auch für die Disability zeigen im Verlauf drei Studien einen signifikanten Unterschied zwischen den zwei Gruppen zugunsten der grossvolumigen Gruppe auf. Dies sind Alaranta et al. 1994[2], Bendix et al. 1995[10] und Bendix et al. 1997[9]. In den restlichen Studien ist der Unterschied entweder nicht signifikant (Bendix et al. 2000[11], Jousset et al. 2004[28], Kääpä et al. 2006[30], Roche et al. 2007[48]) oder nicht untersucht worden (Rose et al. 1997[50]).

4.5 Outcome Anzahl arbeitsunfähige Tage

Wie schon erwähnt erfassen wir das Outcome Anzahl arbeitsunfähige Tage in Bezug auf das Interventionsvolumen qualitativ. Im Folgenden gehen wir auf jede Studie einzeln ein.

Alaranta et al. 1994[2]:

Die Anzahl arbeitsunfähiger Tage pro Jahr wurde drei Wochen vor Beginn der Intervention und zwölf Monate später notiert. In beiden Gruppen ist während der zwölf Monate nach Intervention die durchschnittliche Anzahl arbeitsunfähiger Tage gesunken. In der Akseli Gruppe (high volume) hat sich der Durchschnitt von 57.8 auf 33.9 Tage, in der kleinvolumigen Gruppe von 58.5 auf 36.9 Tage vermindert. Zwischen den Gruppen konnte kein statistisch signifikanter Unterschied festgestellt werden.

Bendix et al. 1995[10]:

Die Arbeitsunfähigkeit wurde in den drei Jahren vor Interventionsbeginn und in den vier Monaten nach der Intervention gemessen. Bei der Gruppe mit hohem Interventionsvolumen ist der Medianwert von 296 auf 25 gesunken. Bei der zweiten Gruppe (low volume) nahm der Wert von 300 auf 13 ab und bei der dritten Gruppe (low volume) von 440 auf 122. Zwischen der Gruppe 3 und den zwei anderen Gruppen besteht ein signifikanter Unterschied, jedoch nicht zwischen der Gruppe 1 und 2. Die Werte der zwei Gruppen mit niedrigem Volumen wollten wir statistisch zusammenführen, da beide das gleiche Interventionsvolumen aufweisen. Wegen fehlender Werte in der Studie ist dies aber nicht möglich. Darum führen wir alle drei Gruppen einzeln auf.

Bendix et al. 1997[9]:

Die Erfassung der arbeitsunfähigen Tage erfolgte über die drei Jahre vor Interventionsbeginn und in einem Jahr nach Interventionsbeginn. Bei der grossvolumigen Rehabilitationsgruppe sank der Medianwert in dieser Zeitspanne von 273 auf 52. Die Anzahl Tage sank in der zweiten Gruppe (low volume) von 300 auf 100 und in der dritten Gruppe (low volume) von 415 auf 295. Der Vergleich zwischen den drei Gruppen ergibt einen signifikanten Unterschied zugunsten der grossvolumigen Gruppe.

Wenn man die zwei Gruppen mit niedrigerem Volumen statistisch zu einer Gruppe zusammenfasst, sinkt die Anzahl arbeitsunfähiger Tage von 360 auf 202 Tage. Wenn man dies mit der Gruppe mit dem höheren Interventionsvolumen vergleicht, ist zu sehen, dass in der Gruppe mit dem höheren Volumen die Arbeitsunfähigkeitstage um 63 Tage mehr abnehmen. Zu diesem Unterschied haben wir keine Angabe über die Signifikanz.

Bendix et al. 2000[11]:

Der Median der arbeitsunfähigen Tage ist bei der funktionellen Rehabilitationsgruppe (high volume) von 200 auf 5.5 und bei der ambulanten Therapiegruppe (low volume) von 220 auf 2.5 gesunken. Der erste Wert entspricht den arbeitsunfähigen Tagen der letzten drei Jahre vor Therapiebeginn und der zweite Wert zeigt die arbeitsunfähigen Tage ein Jahr nach der Intervention. Der Unterschied zwischen den Gruppen ist nicht signifikant, das heisst beide Therapieformen haben fast gleichermassen zu einer Verminderung der arbeitsunfähigen Tage geführt. Die Zeitspanne, in der die Arbeitsunfähigkeit gemessen wurde, ist nicht einheitlich für die erwähnten Messzeitpunkte. Daher sind die Werte nicht direkt vergleichbar.

Jousset et al. 2004[28]:

Die Anzahl arbeitsunfähiger Tage wurde in den sechs Monaten vor und in den sechs Monaten nach der Intervention von den Teilnehmern der Studie selber dokumentiert. Beim funktionellen Rehabilitationsprogramm (high volume) ist die Anzahl Tage in der oben erwähnten Zeitspanne von 101.3 auf 28.7 gesunken. Bei der aktiven individuellen Therapie (low volume) nahm die Zahl arbeitsunfähiger Tage im gleichen Zeitraum von 109.8 auf 48.3 ab. Es konnten also beide Gruppen eine Abnahme der arbeitsunfähigen Tage verbuchen, welche aber beim funktionellen Rehabilitationsprogramm grösser war. In dieser Gruppe nahm die Anzahl der Kranktage um 11,1 Tage mehr ab als in der anderen Gruppe. Dieser Unterschied ist allerdings statistisch nicht signifikant

Kääpä et al. 2006[30]:

Die Anzahl arbeitsunfähiger Tage wurde ein und zwei Jahre nach der Intervention gemessen. Die Autoren verwendeten dafür eine Skala mit folgender Einteilung: Prozent der Studienteilnehmer mit 0 Tagen, mit 1-30 Tagen oder mit mehr als 30 Tagen. Dabei kamen sie zum Schluss, dass die Anzahl arbeitsunfähiger Tage bei beiden

Interventionen stetig abnahm, dass die Senkung aber in der multidisziplinären Rehabilitationsgruppe (high volume) grösser war. In dieser Gruppe sind nach einem Jahr 5,1%, nach zwei Jahren 4,2% mehr Patienten enthalten die keinen Arbeitsunfähigkeitstag aufweisen, als in der Gruppe mit weniger Volumen. Der Unterschied zwischen den beiden Gruppen ist nicht signifikant. Mögliche feine Differenzen zwischen den Gruppen können durch die oben erwähnte Skala nicht erkannt werden.

Roche et al. 2007[48]:

Zu Beginn der Therapie gaben die Patienten der aktiven individuellen Therapie (low volume) im Durchschnitt 135.1 arbeitsunfähige Tage in den letzten zwei Jahren an. Beim funktionellen Rehabilitationsprogramm (high volume) waren es 149.8 Tage. Die Werte unterschieden sich nicht signifikant von einer Gruppe zur anderen. Zum Follow-up Zeitpunkt wurden in dieser Studie die Anzahl arbeitsunfähiger Tage nicht mehr gemessen. Darum können wir keine Aussage über den Einfluss der Therapie in Bezug auf dieses Outcome machen.

Zusammenfassung der Studien für dieses Outcome

Insgesamt haben sieben von uns eingeschlossene Studien das Outcome arbeitsunfähige Tage untersucht. Bei allen Studien ist diese Anzahl Tage bei beiden Interventionsgruppen während des Studienverlaufs gesunken. Bei drei Studien (Jousset et al 2004[28], Käätä et al. 2006[30], Bendix et al. 1997[9]) ist die Abnahme bei der grossvolumigen Gruppe grösser als bei der Gruppe mit kleinerem Interventionsvolumen; der Unterschied zwischen den Gruppen ist allerdings nicht signifikant. Zweimal (Bendix et al. 1995[10], Bendix et al. 2000[11]) ist die Abnahme bei der kleinvolumigen Gruppe grösser als bei der Gruppe mit höherem Volumen, auch hier ist der Unterschied zwischen den Gruppen nicht signifikant. Die Abnahme der arbeitsunfähigen Tage beträgt einmal (Roche et al. 2007[48]) in beiden Gruppen gleich viel. In einem Fall (Alaranta et al. 1994[2]) haben wir keine Information über die Anzahl arbeitsunfähiger Tage während des Studienverlaufs.

Da die Follow-up-Zeiten und die Messmethoden nicht übereinstimmen, können die Studien nicht direkt miteinander verglichen werden.

5 DISKUSSION

Der Zweck dieser systematischen Literaturübersicht ist, bei chronischen unspezifisch lumbalen Rückenschmerzen den Effekt des Interventionsvolumens in Bezug auf die Outcomes Schmerzintensität, Disability und Anzahl arbeitsunfähiger Tage zu untersuchen. Dazu haben wir Schmerzintensität und Disability statistisch ausgewertet. Den Verlauf der Anzahl arbeitsunfähiger Tage beschrieben wir qualitativ.

Von den acht eingeschlossenen Studien haben wir die Daten von fünf in quantitative Messungen einbezogen. Für die qualitative Auswertung verwendeten wir die Daten von sieben Studien.

5.1 Bewertung der Studien

Zu Beginn hatten wir vor, die Cochrane-risk-of-bias-Tabelle und die PEDro-Skala für die Qualitätsbewertung der Studien zu benutzen. Wir sind davon ausgegangen, dass sich die zwei verschiedenen Skalen ergänzen und wir somit mehr Informationen über die Qualität der Studien erhalten. Im Verlauf der Arbeit stellten wir fest, dass diese zwei sehr ähnliche Punkte behandeln. Wir haben uns entschieden, die für uns genauere PEDro-Skala zu benutzen.

Maher et al. 2003[34] stufen die Reliabilität der PEDro-Skala als akzeptabel ein. Laut diesen Autoren ist sie ein gutes Instrument, um die Qualität von RCTs für die Zusammenstellung einer systematischen Literaturübersicht zu prüfen. Es wird empfohlen, die Skala für eine Übersicht der Qualität zu brauchen, nicht aber um Studien definitiv ein- oder auszuschliessen. Wir haben in unserer systematischen Literaturübersicht keine Studie anhand ihres Gesamtscores ausgeschlossen. Die PEDro-Skala benutzten wir, um eine Qualitätsübersicht zu erhalten.

Für uns ist nicht jedes Item der PEDro-Skala von gleicher Bedeutung. Wichtig ist zum Beispiel die randomisierte Zuweisung, da dies auch zu unseren Einschlusskriterien zählt. Die Blindierung der Patienten, Therapeuten sowie der Auswerter ist bei physiotherapeutischen Interventionen aus ethischen und praktischen Gründen oft nicht möglich. Daher ist es für den weiteren Verlauf unserer Review nicht relevant, dass diese Items bei den meisten Studien nicht gewährleistet waren. [Anhang B]

5.2 Limitationen unserer Review

Mittels unserem Einschlussverfahren haben wir drei Studien des gleichen Autors eingeschlossen (Bendix et al. 1995[10], Bendix et al. 1997[9] und Bendix et al. 2000[11]). Die zwei Studien mit dem signifikanten Ergebnis stammen beide von dem Autor Bendix (Bendix et al. 1995[10] und Bendix et al. 1997[9]). Sie weisen sehr ähnliche Anfangsdaten und Populationen auf. Mit einer Generalisierung muss man vorsichtig sein, da wir nicht wissen, ob eventuell die Studie von Bendix et al. 1997[9] die Fortsetzung der Bendix et al. 1995[10] ist. Um dies herauszufinden, haben wir den Autor angeschrieben. Seine Antwort bleibt momentan noch aus.

Die Interventionsvolumen von Jousset et al. 2004[28] und Roche et al. 2007[48] sind für beide Gruppen identisch. Die Populationen, das Follow-up und die Messinstrumente sind jedoch unterschiedlich. Wir gehen davon aus, dass die Studien unabhängig voneinander durchgeführt wurden.

Obwohl wir bei der Methodik systematisch vorgegangen sind, können wir nicht ausschliessen, relevante Studien verpasst zu haben. Durch die Spracheingrenzung auf deutsche, französische und englische Studien wissen wir nicht, ob relevante Studien zu unserem Thema in einer anderen Sprache existieren. Bei der Suche haben wir die für uns wichtigsten Suchwörter definiert. Möglicherweise wäre die Suche mit anderen Wörtern nicht gleich ausgefallen. Wir wissen aber nicht, ob wir dadurch mehr oder weniger Studien gefunden hätten. Mit dem Prüfen der Referenzen und dem Autorenkontakt haben wir versucht, das Risiko Studien zu verpassen, zu minimieren.

Zwischen den eingeschlossenen Studien herrscht eine Heterogenität bezüglich Population, Follow-up, Interventionen, untersuchte Outcomes, Messinstrumente und Interventionsvolumen.

In den von uns eingeschlossenen Studien sind die Interventionen nicht genau gleich, die Therapieprogramme jedoch ähnlich aufgebaut. Jeweils ein Teil der Programme bezieht sich auf das körperliche Training, ein Teil auf die psychologischen Aspekte und ein Teil auf die theoretische Aufklärung der Patienten.

Um die Studien vergleichen zu können, haben wir verschiedene Angaben zusammengefasst. Zum Beispiel sprechen wir bei der Analyse von high und low volume. Wir haben jeweils die Einteilung der Originalautoren übernommen, obwohl

sich die Stundenanzahl von einer Studie zur anderen stark unterscheidet (z.B. entspricht high volume bei Kääpä et al. 2006[30] 70h und bei Bendix et al. 1995[10] 135h).

Weiter hatten die einzelnen Studien sehr unterschiedliche Untersuchungszeiträume. Daher konnten wir die Studien nicht nach dem Verlauf der Gruppen miteinander vergleichen, sondern haben für die statistische Analyse die jeweiligen Schlusswerte der Messungen benützt. Dafür haben wir die Studien nach Follow-up-Zeiten in vier Gruppen eingeteilt.

Um das Outcome Disability trotz verschiedener Messinstrumente vergleichen zu können, haben wir die Fragebögen analysiert. Da sie ähnlich aufgebaut sind, entschieden wir, dass wir sie im Rahmen unserer Arbeit miteinander vergleichen können. Bei der Messung des Outcomes Schmerzintensität wurde bei den in die quantitative Analyse eingeschlossenen Studien die VAS gebraucht. Nur bei zwei Studien ist der Messzeitraum der VAS genau angegeben. Bei Jousset et al. 2004[28] bezieht sich der Messwert auf die letzten 48 Stunden, bei Bendix et al. 2000[11] auf die letzten zwei Wochen. Bei den anderen Studien haben wir darüber keine Information. Aus diesem Grund haben wir diesen Aspekt nicht berücksichtigt.

Trotz diesen Gruppierungen konnten wir nicht alle Studien mit gleichen Outcomes miteinander vergleichen. Alaranta et al. 1994[2] haben zwar Schmerz und Disability untersucht, dies aber nur in Kombination. Weil die Werte nicht einzeln für das betreffende Outcome gegeben waren, konnten wir sie nicht in unsere Berechnungen miteinbeziehen. Bei Roche et al. 2007[48] stehen uns für das Outcome Schmerzintensität nur die Angaben über den Verlauf zur Verfügung nicht aber zum Zeitpunkt des Follow-ups, was nötig wäre, um sie in unsere statistische Analyse aufzunehmen. Rose et al. 1997[50] haben Schmerz und Disability nur als Graphik mit ungenauen Angaben dargestellt.

Es ist uns bewusst, dass durch diese verschiedenen Zusammenfassungen und Gruppierungen die Genauigkeit abnimmt. Wir haben aber keine andere Möglichkeit gesehen, wie man die verschiedenen Angaben in den Studien miteinander vergleichen könnte.

5.3 Interpretation der Resultate

5.3.1 Outcome Schmerzintensität

In den Forest Plots der Abbildungen 3, 4, 5 und 6 tendieren die Ergebnisse der untersuchten Studien aller Follow-ups im Allgemeinen in Richtung der Gruppen mit höherem Interventionsvolumen. Dies wird durch den Diamanten in den jeweiligen Abbildungen dargestellt. In Abbildung 4 erkennt man, dass bei Bendix et al. 1995[10] die Gruppe mit höherem Interventionsvolumen eine signifikant tiefere Schmerzintensität gegenüber der Gruppe mit niedrigerem Volumen aufweist. Bei Bendix et al. 1997[9] weist die Gruppe mit dem höheren Interventionsvolumen eine signifikant tiefere Schmerzintensität auf als die Gruppe mit niedrigerem Volumen. (siehe Abbildung 5)

Bei allen anderen untersuchten Studien ist der Unterschied zwischen den zwei Gruppen nicht signifikant.

5.3.2 Outcome Disability

In den Abbildungen 7, 8 und 9 tendieren die Ergebnisse der untersuchten Studien im Allgemeinen in Richtung der Gruppen mit höherem Interventionsvolumen. Dies wird durch den Diamanten in den jeweiligen Abbildungen dargestellt. In Abbildung 10 verschiebt sich die Tendenz eher zu der Gruppe mit niedrigerem Interventionsvolumen. Bei Bendix et al. 1995[10] und Bendix et al. 1997[9] ist der Unterschied zwischen den beiden Gruppen signifikant zugunsten der Gruppe mit dem höheren Interventionsvolumen.

5.3.3 Outcome Anzahl arbeitsunfähige Tage

Für das Outcome Anzahl arbeitsunfähige Tage haben wir den Verlauf der beiden Gruppen innerhalb jeder einzelnen Studie beschrieben. Es konnten keine Vergleiche zwischen den Studien aufgestellt werden, da die Angaben und Messinstrumente sehr unterschiedlich waren.

Die Anzahl arbeitsunfähiger Tage ist in allen Studien, die das Outcome untersuchen, bei beiden Interventionsgruppen gesunken. Der Unterschied zwischen den Gruppen ist in keiner der sieben Studien signifikant. Die Abnahme ist bei drei Studien (Jousset et al 2004[28], Käätä et al. 2006[30], Bendix et al. 1997[9]) bei der grossvolumigen Gruppe grösser als bei der Gruppe mit kleinerem Interventionsvolumen. In zwei Studien (Bendix et al. 1995[10], Bendix et al. 2000[11]) ist das Gegenteil der Fall. In einer

Studie (Roche et al. 2007[48]) ist die Abnahme für beide Gruppen identisch. In der Studie von Alaranta et al. 1994[2] ist keine Information über den Verlauf dieses Outcomes verfügbar.

5.3.4 Interpretation der statistischen Analyse

Zu Beginn der Arbeit haben wir uns entschieden, ein hohes Interventionsvolumen mit einem tiefen zu vergleichen. Wir haben uns nicht genauer auf die Anzahl Stunden eingelassen. Trotzdem ist es wichtig, auf auffallende Differenzen einzugehen, welche das Resultat beeinflussen könnten.

Bei den eingeschlossenen Studien zeigt sich, dass manche Autoren z.B. 135h und andere 70h als hohes Volumen bezeichnen. Es ist schwierig eine generalisierte Aussage anhand unserer Resultate zu machen. Wir stellen aber fest, dass Bendix et al. 1995[10] und Bendix et al. 1997[9] jeweils von einem high volume von 135h reden und nur diese zwei Studien ein signifikantes Ergebnis zugunsten der grossvolumigen Gruppe aufweisen. Daraus könnten wir schliessen, dass ein Interventionsvolumen von z.B. über 100h bessere Effekte erbringt als ein kleineres Volumen. Dies würde die Ergebnisse von Guzmán et al. 2001[22] bestätigen, welcher Programme mit über 100h als intensiv einordnet.

In Tabelle 10 sieht man auf den ersten Blick, dass sich die Volumen in den Studien von Kääpä et al. 2006[30] und Rose et al. 1997[50] stark von den anderen unterscheiden ; sowohl das Volumen der grossvolumigen Gruppen wie auch der kleinvolumigen ist tiefer als in den anderen untersuchten Studien. Bei Kääpä et al. 2006[30] sind in unseren Berechnungen keine signifikanten Ergebnisse zu sehen und bei Rose et al. 1997[50] fehlen die Angaben für die statistische Analyse. Wir sind lediglich über die Signifikanz im Verlauf informiert. Rose et al. 1997[50] haben wir einzeln beschrieben, also nicht direkt mit einer anderen Studie verglichen. Aus diesen Gründen können wir in diesen zwei Fällen den grossen Volumenunterschied zu den anderen Studien vernachlässigen.

Bei den zwei Studien (Bendix et al. 1997[9], Bendix et al. 1995[10]), die einen signifikanten Unterschied zwischen den zwei Gruppen für die Outcomes Schmerzintensität und Disability aufweisen, beträgt das Interventionsvolumen 135 bzw. 24 Stunden. Aus diesem Resultat könnte man schliessen, dass 135 Stunden ein besseres

Interventionsvolumen bezüglich der Outcomes Schmerzintensität und Disability in einem Zeitrahmen bis 16 Monate nach Intervention wären. Wir erläutern im Anschluss warum wir aber diese Schlussfolgerung nicht machen dürfen.

Beim Follow-up von 1-8 Monaten vergleichen wir für die Outcomes Schmerz und Disability die Studien von Bendix et al. 1995[10] mit Jousset et al. 2004[28] und Kääpä et al. 2006[30]. Damit stellen wir in der Gruppe vom high Volume die Volumenangaben 135h, 150h und 70h gleich. Wir fassen also verschiedene Stundenangaben in einer Gruppe zusammen. Wenn man betrachtet, dass 150h mehr als doppelt so viel ist wie 70h, kann dies eine Fehlerquelle darstellen. Da bei Bendix et al. 1995[10] nur das Interventionsvolumen von 135h ein signifikantes Ergebnis zeigt, könnte man daraus schliessen, dass 70h zu wenig und 150h zu viel Volumen sind. Wir können aber nicht wissen, ob nur das Therapievolumen dieses signifikante Resultat ausmacht. Es können auch andere Faktoren wie z.B. die Art der Therapie, die Population, die Erfahrung der Therapeuten, die Compliance oder sozioökonomische Einflüsse eine wesentliche Rolle spielen.

Beim Follow-up von 9-16 Monate vergleichen wir für die Outcomes Schmerz und Disability die Studien von Bendix et al. 1997[9], Bendix 2000[11] und Kääpä et al. 2006[30]. Auch hier stellen wir damit in der Gruppe vom high Volume die Volumenangaben 135h, 117h und 70h gleich. Nur das Volumen von 135h von Bendix et al. 1997[9] zeigt einen signifikanten Unterschied zwischen den zwei Interventionsgruppen. Trotzdem können wir nicht einfach zum Schluss kommen, dass 117h und 70h zu wenig sind, nur weil sie kein signifikantes Ergebnis aufweisen. Auch hier können diverse andere Faktoren wie zuvor geschrieben den Unterschied ausmachen.

Ein weiterer Grund für das signifikante Ergebnis von Bendix et al. 1995[10] und Bendix et al. 1997[9] könnte auch der Stundenunterschied zwischen dem high und dem low Volume in einer Studie sein. Bei Bendix et al. 1995[10] und Bendix et al. 1997[9] beträgt dieser Unterschied 111h, was verglichen mit den anderen Studien eine grosse Differenz darstellt. Daher könnte es auch sein, dass die Verbesserung der Outcomes Schmerz und Disability am Ende der Intervention bei allen andern untersuchten Studien nicht signifikant ist, weil der Stundenunterschied zwischen den zwei Gruppen innerhalb

einer Studie zu klein ist. Dieses Thema könnte in einer eigenen Studie weiter erforscht werden.

Bezüglich Bendix et al. 1995[10] und Bendix et al. 1997[9] bestätigt sich für die zwei Outcomes Schmerzintensität und Disability unsere Arbeitshypothese. Ein höheres Interventionsvolumen wirkt sich signifikant mehr positiv auf diese zwei Outcomes bei chronischen unspezifisch lumbalen Rückenschmerzen aus. Wie schon erwähnt, können wir daraus keine allgemeingültige Aussage aufstellen. Bei der qualitativen Analyse ist keine signifikante Entwicklung zu sehen. Unsere Arbeitshypothese bestätigt sich also auch nicht für das Outcome Anzahl arbeitsunfähige Tage.

5.3.5 Analyse des Verlaufs

Wenn man den Verlauf der Gruppen innerhalb der Studien betrachtet, sieht man viele signifikante Ergebnisse, dies sowohl für die gross- wie auch für die kleinvolumige Gruppe. Der Unterschied des Verlaufs zwischen den zwei Gruppen ist aber nur in drei von den acht eingeschlossenen Studien signifikant. Bei Alaranta et al. 1994[2], Bendix et al. 1995[10] und Bendix et al. 1997[9] ist dies jeweils zugunsten der Gruppe mit dem grösseren Volumen zu beobachten.

Von den drei Studien (Alaranta et al. 1994[2], Bendix et al. 1995[10] und Bendix et al. 1997[9]), die im Verlauf ein signifikantes Ergebnis aufweisen, sind zwei (Bendix et al. 1995[10] und Bendix et al. 1997[9]) auch in unserer statistischen Analyse am Follow-up Zeitpunkt signifikant. Über die Effektgrösse am Ende der Intervention haben wir bei Alaranta et al. 1994[2] keine Informationen. Vielleicht wären diese Werte auch signifikant, wenn eine statistische Analyse möglich gewesen wäre.

5.4 Zusammenfassung der Diskussion

Anhand unserer Analysen ist es uns nicht möglich, eine spezifische Schlussfolgerung bezüglich des Interventionsvolumens zu machen. Das Zusammenspiel vieler anderer Faktoren kann neben dem Volumen die untersuchten Outcomes beeinflussen. Daher ist es sehr schwierig, das Therapievolumen isoliert zu untersuchen. Es ist uns nicht möglich, unsere Arbeitshypothese zu bestätigen. Aufgrund der zwei signifikanten Ergebnisse von Bendix et al. 1995[10] und Bendix et al. 1997[9] können wir sie auch nicht eindeutig verwerfen.

Wir können mit unserer Studie eine Tendenz Richtung high volume erkennen, dies aber nicht generalisieren. In unserer Review zeigt keine Studie ein signifikantes Resultat zugunsten der kleinvolumigen Gruppe auf. Wir können also ausschliessen, dass eine kleinvolumige Therapie bessere Effekte auf die Schmerzintensität, die Disability und die arbeitsunfähigen Tage erzielt als eine grossvolumig durchgeführte Rehabilitation. Wir können annehmen, dass diese Aussage generalisierbar ist, da trotz der Heterogenität der Studien bezüglich Population, Follow-up, Interventionen, untersuchter Outcomes, Messinstrumente, Interventionsvolumen und sozioökonomischer Einflüsse alle Studien zu diesem Ergebnis kommen.

Für die Rehabilitationszentren bedeutet dies, dass sie eher grossvolumige Programme aufbauen sollten. Trotzdem bleiben einige Fragen offen. Ab welchem Volumen ist eine Therapie wirksam? Welche Therapieinhalte sind für welchen Patienten wirksam? Um diese Fragen zu beantworten ist weitere Forschung nötig. Zur Begründung einer Kostengutsprache einer grossvolumigen Rehabilitationsmassnahme kann mit Hilfe dieser Review gegebenenfalls besser argumentiert werden.

Für die Versicherungen stellt unsere Arbeit neben anderen Studien [6, 11, 22, 51, 55] eine Grundlage für weitere Forschung dar. Es existieren Studien, welche die verschiedenen Interventionsarten bei chronischen unspezifisch lumbalen Rückenschmerzen analysieren oder miteinander vergleichen [21, 38]. Es gibt auch diverse Studien, welche die Effizienz einer multidisziplinären Rehabilitation aufzeigen.[6, 22, 43, 51] Für die Leistungszahler ist es nötig, mit weiteren Studien herauszufinden, welche Interventionen mit welchem Volumen kosteneffizient sind.

Guzmán et al. 2001[22] haben in einer systematischen Literaturübersicht multidisziplinäre Therapieprogramme mit nicht-multidisziplinären Rehabilitationsprogrammen verglichen. Im Laufe der Untersuchung haben sie die eingeschlossenen Studien anhand ihres Interventionsvolumens eingeteilt. Die Programme mit über 100 Stunden wurden als intensiv eingeordnet. Die Autoren kommen zum Schluss, dass ein intensives multidisziplinäres Rehabilitationsprogramm bessere Ergebnisse in Bezug auf Schmerz und Disability aufweist als z.B. ein weniger intensives oder ein nicht-multidisziplinäres Programm. Guzmán et al. 2001[22] wollten die Effizienz der multidisziplinären Programme an sich untersuchen und sind erst in einem weiteren Schritt auf die Volumeneinteilung gekommen. Es bleibt die Frage, ob

ihre Ergebnisse auf das Volumen oder den multidisziplinären Ansatz zurückzuführen sind.

Scascighini et al. 2008[51] haben in einer systematischen Literaturübersicht den Effekt eines multidisziplinären Therapieansatzes in Bezug auf verschiedene Interventionen und Outcomes untersucht. In diesem Zusammenhang haben sie auch den Einfluss des Behandlungsvolumens der einzelnen Programme angeschaut. Laut ihrer Aussage beeinflusst das Therapievolumen die untersuchten Outcomes nicht. Die Einschlusskriterien für ihre Studien basierten aber nicht auf Volumenangaben und sie haben auch Patienten mit Fibromyalgie untersucht. Somit können sie mit ihrer systematischen Review nicht ein abschliessendes Ergebnis in Bezug auf das Interventionsvolumen und chronische unspezifisch lumbale Rückenschmerzen vorweisen.[51]

Wir bringen mit der Untersuchung des Interventionsvolumens eine andere Sichtweise in das Thema chronische unspezifisch lumbale Rückenschmerzen ein. RCTs, welche effektiv ein Interventionsvolumen gegen ein anderes Volumen vergleichen, haben wir nur wenige gefunden. Die, welche wir gefunden haben, weisen eine grosse Heterogenität bezüglich Outcomes, Messinstrumente, Population und Follow-ups auf. Somit war es schwierig eine generalisierbare systematische Literaturübersicht hierüber zu schreiben. In Zukunft wäre es also interessant und wichtig, Studien bezüglich des Volumens zu starten, damit eventuell zu einem späteren Zeitpunkt weitere Reviews möglich wären. Daneben sollten Studien über die Kosten von Rehabilitationsprogrammen durchgeführt werden, um zu sehen, welche Therapien sich auch von den Kosten her lohnen.[51] Die Voraussetzungen für die Durchführung der genannten Studien wären ein einheitliches Therapieprogramm und Follow-ups, vergleichbare Volumenangaben, Outcomes, Messinstrumente sowie Population.

6 SCHLUSSFOLGERUNG

In unserer statistischen Analyse zum Follow-up Zeitpunkt weisen, in Bezug auf die Outcomes Schmerzintensität und Disability, zwei Studien ein signifikantes Ergebnis zugunsten der grossvolumigen Gruppe auf. Betrachten wir den Verlauf innerhalb einer Studie, so sehen wir, dass drei Studien ein signifikantes Resultat zeigen. Die Tendenz geht also in Richtung grossvolumige Interventionen. Trotzdem können wir mit unserer Arbeit nicht eindeutig sagen, dass grosses Volumen besser ist als kleines.

Es ist schwierig den Einfluss des Interventionsvolumens auf chronische unspezifisch lumbale Rückenschmerzen zu untersuchen. Verschiedene Faktoren wie z.B. die Art der Behandlung, die Population, die Erfahrung der Therapeuten, die Compliance oder sozioökonomische Gegebenheiten beeinflussen die Resultate. Um bezüglich Interventionsvolumen weiterzuforschen, braucht es mehr Studien, welche in einem standardisierten Programm ein Interventionsvolumen mit einem anderen vergleichen.

Zukünftige Forschung auf diesem Gebiet ist sehr wichtig, um einen Anhaltspunkt bezüglich des Therapievolumens in der komplexen Behandlung von chronischen unspezifisch lumbalen Rückenschmerzen zu erhalten, Möglichkeiten für die Kostensenkung zu erarbeiten und eine generalisierbare Festigung der bisherigen Ergebnisse zu erzielen.

Unsere systematische Literaturübersicht soll ein Ansporn für Untersuchungen in diesem Gebiet darstellen.

7 BIBLIOGRAPHIE

1. Airaksinen, O., J.I. Brox, C. Cedraschi, J. Hildebrandt, J. Klaber-Moffett, F. Kovacs, A.F. Mannion, S. Reis, J.B. Staal, H. Ursin, and G. Zanoli, *Chapter 4. European guidelines for the management of chronic nonspecific low back pain*. Eur Spine J, 2006. 15 Suppl 2: p. S192-300.
2. Alaranta, H., U. Rytokoski, A. Rissanen, S. Tallo, T. Ronnema, P. Puukka, S.L. Karppi, T. Videman, V. Kallio, and P. Slati, *Intensive physical and psychosocial training program for patients with chronic low back pain. A controlled clinical trial*. Spine, 1994. 19(12): p. 1339-49.
3. Anagnostis, C., T.G. Mayer, R.J. Gatchel, and T.J. Proctor, *The million visual analog scale: its utility for predicting tertiary rehabilitation outcomes*. Spine, 2003. 28(10): p. 1051-60.
4. Anand, K.J. and K.D. Craig, *New perspectives on the definition of pain*. Pain, 1996. 67(1): p. 3-6; discussion 209-11.
5. Andersson, G.B., *Epidemiological features of chronic low-back pain*. Lancet, 1999. 354(9178): p. 581-5.
6. Becker, A., C. Leonhardt, M.M. Kochen, S. Keller, K. Wegscheider, E. Baum, N. Donner-Banzhoff, M. Pflingsten, J. Hildebrandt, H.D. Basler, and J.F. Chenot, *Effects of two guideline implementation strategies on patient outcomes in primary care: a cluster randomized controlled trial*. Spine, 2008. 33(5): p. 473-80.
7. Bendix, A.E., T. Bendix, C. Hastrup, and E. Busch, *A prospective, randomized 5-year follow-up study of functional restoration in chronic low back pain patients*. Eur Spine J, 1998. 7(2): p. 111-9.
8. Bendix, A.F., T. Bendix, M. Labriola, and P. Boekgaard, *Functional restoration for chronic low back pain. Two-year follow-up of two randomized clinical trials*. Spine, 1998. 23(6): p. 717-25.
9. Bendix, A.F., T. Bendix, C. Lund, S. Kirkbak, and S. Ostfeld, *Comparison of three intensive programs for chronic low back pain patients: a prospective, randomized, observer-blinded study with one-year follow-up*. Scand J Rehabil Med, 1997. 29(2): p. 81-9.
10. Bendix, A.F., T. Bendix, S. Ostfeld, E. Bush, and Andersen, *Active treatment programs for patients with chronic low back pain: a prospective, randomized, observer-blinded study*. Eur Spine J, 1995. 4(3): p. 148-52.
11. Bendix, T., A. Bendix, M. Labriola, C. Hastrup, and N. Ebbelohj, *Functional restoration versus outpatient physical training in chronic low back pain: a randomized comparative study*. Spine, 2000. 25(19): p. 2494-500.
12. Beurskens, A.J., H.C. de Vet, A.J. Koke, G.J. van der Heijden, and P.G. Knipschild, *Measuring the functional status of patients with low back pain. Assessment of the quality of four disease-specific questionnaires*. Spine, 1995. 20(9): p. 1017-28.
13. Callaghan, M.J., *Evaluation of a back rehabilitation group for chronic low back pain in an out-patient setting*. Physiotherapy, 1994. 80(10): p. 677-81.
14. Cedraschi, C., J. Robert, D. Goerg, E. Perrin, W. Fischer, and T.L. Vischer, *Is chronic non-specific low back pain chronic? Definitions of a problem and problems of a definition*. Br J Gen Pract, 1999. 49(442): p. 358-62.
15. Deyo, R.A., M. Battie, A.J. Beurskens, C. Bombardier, P. Croft, B. Koes, A. Malmivaara, M. Roland, M. Von Korff, and G. Waddell, *Outcome measures for*

- low back pain research. A proposal for standardized use. Spine*, 1998. 23(18): p. 2003-13.
16. Dionne, C.E., K.M. Dunn, P.R. Croft, A.L. Nachemson, R. Buchbinder, B.F. Walker, M. Wyatt, J.D. Cassidy, M. Rossignol, C. Leboeuf-Yde, J. Hartvigsen, P. Leino-Arjas, U. Latza, S. Reis, M.T. Gil Del Real, F.M. Kovacs, B. Oberg, C. Cedraschi, L.M. Bouter, B.W. Koes, H.S. Picavet, M.W. van Tulder, K. Burton, N.E. Foster, G.J. Macfarlane, E. Thomas, M. Underwood, G. Waddell, P. Shekelle, E. Volinn, and M. Von Korff, *A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine*, 2008. 33(1): p. 95-103.
 17. Fairbank, J.C., J. Couper, J.B. Davies, and J.P. O'Brien, *The Oswestry low back pain disability questionnaire. Physiotherapy*, 1980. 66(8): p. 271-3.
 18. Fritz, J.M. and J.J. Irrgang, *A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. Phys Ther*, 2001. 81(2): p. 776-88.
 19. Gallacchi, G. and B. Pilger, *Schmerzkompodium: Schmerzen verstehen und behandeln*. 2nd ed. 2005: Georg Thieme Verlag. p. 303. ISBN: 3131283726
 20. Garrad, J. and A.E. Bennett, *A validated interview schedule for use in population surveys of chronic disease and disability. Br J Prev Soc Med*, 1971. 25(2): p. 97-104.
 21. Goldby, L.J., A.P. Moore, J. Doust, and M.E. Trew, *A randomized controlled trial investigating the efficiency of musculoskeletal physiotherapy on chronic low back disorder. Spine*, 2006. 31(10): p. 1083-93.
 22. Guzmán, J., R. Esmail, K. Karjalainen, A. Malmivaara, E. Irvin, and C. Bombardier, *Multidisciplinary rehabilitation for chronic low back pain: systematic review. BMJ*, 2001. 322: p. 1511-1516.
 23. Hansen, F.R., T. Bendix, P. Skov, C.V. Jensen, J.H. Kristensen, L. Krohn, and H. Schioeler, *Intensive, dynamic back-muscle exercises, conventional physiotherapy, or placebo-control treatment of low-back pain. Spine*, 1992. 18(1): p. 98-108.
 24. Harkapaa, K., A. Jarvikoski, G. Mellin, and H. Hurri, *A controlled study on the outcome of inpatient and outpatient treatment of low back pain. Part I. Pain, disability, compliance, and reported treatment benefits three months after treatment. Scand J Rehabil Med*, 1989. 21(2): p. 81-9.
 25. Harkapaa, K., G. Mellin, A. Jarvikoski, and H. Hurri, *A controlled study on the outcome of inpatient and outpatient treatment of low back pain. Part III. Long-term follow-up of pain, disability, and compliance. Scand J Rehabil Med*, 1990. 22(4): p. 181-8.
 26. Heymans, M.W., H.C. de Vet, P.M. Bongers, D.L. Knol, B.W. Koes, and W. van Mechelen, *The effectiveness of high-intensity versus low-intensity back schools in an occupational setting: a pragmatic randomized controlled trial. Spine*, 2006. 31(10): p. 1075-82.
 27. Higgins, J.P.T. and S. Green, *Cochrane Handbook for Systematic Reviews of Interventions*, Version 5.0.0 [updated February 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org
 28. Jousset, N., S. Fanello, L. Bontoux, V. Dubus, C. Billabert, B. Vielle, Y. Roquelaure, D. Penneau-Fontbonne, and I. Richard, *Effects of functional restoration versus 3 hours per week physical therapy: a randomized controlled study. Spine*, 2004. 29(5): p. 487-93; discussion 494.
 29. Junge, A. and A.F. Mannion, *Fragebögen für Patienten mit Rückenschmerzen. Orthopäde*, 2004. 33: p. 545-552.

30. Kaapa, E.H., K. Frantsi, S. Sarna, and A. Malmivaara, *Multidisciplinary group rehabilitation versus individual physiotherapy for chronic nonspecific low back pain: a randomized trial*. Spine, 2006. 31(4): p. 371-6.
31. Kopec, J.A., J.M. Esdaile, M. Abrahamowicz, L. Abenhaim, S. Wood-Dauphinee, D.L. Lamping, and J.I. Williams, *The Quebec Back Pain Disability Scale: conceptualization and development*. J Clin Epidemiol, 1996. 49(2): p. 151-61.
32. Loney, P.L. and P.W. Stratford, *The prevalence of low back pain in adults: a methodological review of the literature*. Phys Ther, 1999. 79(4): p. 384-96.
33. Ludwig, F.J., A.I. Ünlü, C. Melzer, and H.H. Daalman, *Effekte der stationären medizinischen Rehabilitation bei Rückenschmerzpatienten*. Z Orthop, 2006. 144: p. 569-76.
34. Maher, C.G., C. Sherrington, R.D. Herbert, A.M. Moseley, and M. Elkins, *Reliability of the PEDro scale for rating quality of randomized controlled trials*. Phys Ther, 2003. 83(8): p. 713-21.
35. Maleitzke, R., K.H. Göttl, B. Tschirdewahn, and N. Gerdes, *Stationäre Rehabilitation beim chronischen Rückenschmerz - Vergleich zwischen 3- und 4wöchiger Behandlungsdauer*. Akt. Rheumatol., 1997. 22: p. 112-6.
36. Manniche, C., K. Asmussen, B. Lauritsen, H. Vinterberg, S. Kreiner, and A. Jordan, *Low Back Pain Rating scale: validation of a tool for assessment of low back pain*. Pain, 1994. 57(3): p. 317-26.
37. Manniche, C., G. Hesselsoe, L. Bentzen, I. Christensen, and E. Lundberg, *Clinical trial of intensive muscle training for chronic low back pain*. Lancet, 1988. 2: p. 1473-6.
38. Mannion, A.F., M. Muntener, S. Taimela, and J. Dvorak, *Comparison of three active therapies for chronic low back pain: results of a randomized clinical trial with one-year follow-up*. Rheumatology (Oxford), 2001. 40(7): p. 772-8.
39. Mayer, T., P. Polatin, B. Smith, R. Gatchel, D. Fardon, S. Herring, C. Smith, R. Donelson, and D. Wong, *Spine rehabilitation: secondary and tertiary nonoperative care*. Spine J, 2003. 3(3 Suppl): p. 28S-36S.
40. Mayer, T.G., R.J. Gatchel, N. Kishino, J. Keeley, P. Capra, H. Mayer, J. Barnett, and V. Mooney, *Objective assessment of spine function following industrial injury. A prospective study with comparison group and one-year follow-up*. Spine, 1985. 10(6): p. 482-93.
41. Mellin, G., K. Harkapaa, H. Hurri, and A. Jarvikoski, *A controlled study on the outcome of inpatient and outpatient treatment of low back pain. Part IV. Long-term effects on physical measurements*. Scand J Rehabil Med, 1990. 22(4): p. 189-94.
42. Mellin, G., H. Hurri, K. Harkapaa, and A. Jarvikoski, *A controlled study on the outcome of inpatient and outpatient treatment of low back pain. Part II. Effects on physical measurements three months after treatment*. Scand J Rehabil Med, 1989. 21(2): p. 91-5.
43. Negrini, S., C. Fusco, S. Atanasio, M. Romano, and F. Zaina, *Low back pain: state of art*. European journal of pain, 2008. 2: p. 52-56.
44. New Zealand Guidelines Group, *New Zealand acute low back pain guide*, October 2004.
45. Oesch, P., R. Hilfiker, S. Keller, J. Kool, S. Schädler, A. Tal-Akabi, M. Verra, and C. Widmer Leu, *Assessments in der muskuloskelettalen Rehabilitation*. 1. ed. 2007, Bern: Verlag Hans Huber.978-3-456-84452-7
46. Pfungsten, M., *Multimodale Verfahren - auf die Mischung kommt es an!* Schmerz, 2001. 15: p. 492-498.

47. Rainville, J., C.A. Jouve, C. Hartigan, E. Martinez, and M. Hipona, *Comparison of short- and long-term outcomes for aggressive spine rehabilitation delivered two versus three times per week*. Spine J, 2002. 2(6): p. 402-7.
48. Roche, G., A. Ponthieux, E. Parot-Shinkel, N. Jousset, L. Bontoux, V. Dubus, D. Penneau-Fontbonne, Y. Roquelaure, E. Legrand, D. Colin, I. Richard, and S. Fanello, *Comparison of a functional restoration program with active individual physical therapy for patients with chronic low back pain: a randomized controlled trial*. Arch Phys Med Rehabil, 2007. 88(10): p. 1229-35.
49. Roland, M. and J. Fairbank, *The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire*. Spine, 2000. 25(24): p. 3115-24.
50. Rose, M.J., J.P. Reilly, B. Pennie, K. Bowen-Jones, I.M. Stanley, and P.D. Slade, *Chronic low back pain rehabilitation programs: a study of the optimum duration of treatment and a comparison of group and individual therapy*. Spine, 1997. 22(19): p. 2246-51; discussion 2252-3.
51. Scascighini, L., V. Toma, S. Dober-Spielmann, and H. Sprott, *Multidisciplinary treatment for chronic pain: a systematic review of interventions and outcomes*. Rheumatology (Oxford), 2008. 47(5): p. 670-8.
52. Schumacher, J. and E. Brähler, *Psychologische Aspekte akuter und chronischer Schmerzen*, in *Handlungsfelder der psychosozialen Medizin*, E. Brähler and B. Strauss. 2002, Hogrefe: Göttingen. p. 187-207.
53. Skouen, J.S., A.L. Grasdahl, E.M. Haldorsen, and H. Ursin, *Relative cost-effectiveness of extensive and light multidisciplinary treatment programs versus treatment as usual for patients with chronic low back pain on long-term sick leave: randomized controlled study*. Spine, 2002. 27(9): p. 901-10.
54. Stebler, R., R. Putzi, and B.A. Michel, *Lumbale Rückenschmerzen - Diagnostik*. Schweizerisches medizinisches Forum, 2001. 9: p. 205-208.
55. Turk, D., *Efficacy of multidisciplinary pain centers in the treatment of chronic pain*, in *Pain Treatment Centers at a Crossroads*, M. Cohen and J. Campbell. 1996, IASP Press: Seattle. p. 297-306.
56. van der Roer, N., M. van Tulder, J. Barendse, D. Knol, W. van Mechelen, and H. de Vet, *Intensive group training protocol versus guideline physiotherapy for patients with chronic low back pain: a randomised controlled trial*. Eur Spine J, 2008. 17(9): p. 1193-200.
57. Van der Roer, N., M. Van Tulder, W. Van Mechelen, and H. De Vet, *Economic evaluation of an intensive group training protocol compared with usual care physiotherapy in patients with chronic low back pain*. Spine, 2008. 33(4): p. 445-451.
58. Vleeming, A., V. Mooney, and R. Stoeckart, *Movement, stability and lumbopelvic pain; integration of research and therapy*. 2nd ed. 2007, Amsterdam: Elsevier. p. 658. ISBN: 978 0 443 10178 6
59. Volinn, E., *The epidemiology of low back pain in the rest of the world. A review of surveys in low- and middle-income countries*. Spine, 1997. 22(15): p. 1747-54.
60. Waddell, G., *The back pain revolution*. 1st ed. 1998, Edinburgh: Churchill Livingstone. p. 438. ISBN: 0443060398
61. Wenig, C.M., C.O. Schmidt, T. Kohlmann, and B. Schweikert, *Costs of back pain in Germany*. Eur J Pain, 2009. 13(3): p. 280-6.
62. WHO, *International classification of impairments, disabilities and handicaps*. 1980, World Health Organization: Geneva.

Anhang A: Red und Yellow Flags

RED AND YELLOW FLAGS – A QUICK OVERVIEW

RED FLAGS

Red Flags help identify potentially serious conditions. They include:

- »» Features of Cauda Equina Syndrome (see page 8 for symptoms)
- »» Severe worsening pain, especially at night or when lying down
- »» Significant trauma
- »» Weight loss, history of cancer, fever
- »» Use of intravenous drugs or steroids
- »» Patient over 50 years old

YELLOW FLAGS

Yellow Flags indicate psychosocial barriers to recovery. They include:

- »» Belief that pain and activity are harmful
- »» ‘Sickness behaviours’ (like extended rest)
- »» Low or negative moods, social withdrawal
- »» Treatment that does not fit best practice
- »» Problems with claim and compensation
- »» History of back pain, time-off, other claims
- »» Problems at work, poor job satisfaction
- »» Heavy work, unsociable hours
- »» Overprotective family or lack of support

Von New Zealand Guidelines Group, *New Zealand acute low back pain guide*, October 2004.

Anhang B: Definitionen der elf Bewertungskriterien der PEDro-Skala

Criterion	Operational Definition
All criteria	Points are awarded only when a criterion is clearly satisfied. If, on a literal reading of the trial report, it is possible that a criterion was not satisfied, a point should not be awarded for that criterion.
Criterion 1	This criterion is satisfied if the report describes the source of subjects and a list of criteria used to determine who was eligible to participate in the study.
Criterion 2	A study is considered to have used random allocation if the report states that allocation was random. The precise method of randomization need not be specified. Procedures such as coin tossing and dice rolling should be considered random. Quasi-randomization allocation procedures such as allocation by hospital record number or birth date, or alternation, do not satisfy this criterion.
Criterion 3	Concealed allocation means that the person who determined if a subject was eligible for inclusion in the trial was unaware, when this decision was made, of which group the subject would be allocated to. A point is awarded for this criterion, even if it is not stated that allocation was concealed, when the report states that allocation was by sealed opaque envelopes or that allocation involved contacting the holder of the allocation schedule who was "off-site."
Criterion 4	At a minimum, in studies of therapeutic interventions, the report must describe at least one measure of the severity of the condition being treated and at least one (different) key outcome measure at baseline. The rater must be satisfied that the groups' outcomes would not be expected to differ, on the basis of baseline differences in prognostic variables alone, by a clinically significant amount. This criterion is satisfied even if only baseline data of subjects completing the study are presented.
Criteria 4, 7-11	Key outcomes are those outcomes that provide the primary measure of the effectiveness (or lack of effectiveness) of the therapy. In most studies, more than one variable is used as an outcome measure.
Criteria 5-7	Blinding means the person in question (subject, therapist, or assessor) did not know which group the subject had been allocated to. In addition, subjects and therapists are only considered to be "blind" if it could be expected that they would have been unable to distinguish between the treatments applied to different groups. In trials in which key outcomes are self-reported (eg, visual analog scale, pain diary), the assessor is considered to be blind if the subject was blind.
Criterion 8	This criterion is satisfied only if the report explicitly states both the number of subjects initially allocated to groups and the number of subjects from whom key outcome measurements were obtained. In trials in which outcomes are measured at several points in time, a key outcome must have been measured in more than 85% of subjects at one of those points in time.
Criterion 9	An intention-to-treat analysis means that, where subjects did not receive treatment (or the control condition) as allocated and where measures of outcomes were available, the analysis was performed as if subjects received the treatment (or control condition) they were allocated to. This criterion is satisfied, even if there is no mention of analysis by intention to treat, if the report explicitly states that all subjects received treatment or control conditions as allocated.
Criterion 10	A between-group statistical comparison involves statistical comparison of one group with another. Depending on the design of the study, this may involve comparison of 2 or more treatments or comparison of treatment with a control condition. The analysis may be a simple comparison of outcomes measured after the treatment was administered or a comparison of the change in one group with the change in another (when a factorial analysis of variance has been used to analyze the data, the latter is often reported as a group \times time interaction). The comparison may be in the form hypothesis testing (which provides a <i>P</i> value, describing the probability that the groups differed only by chance) or in the form of an estimate (eg, the mean or median difference, a difference in proportions, number needed to treat, a relative risk or hazard ratio) and its confidence interval.
Criterion 11	A point measure is a measure of the size of the treatment effect. The treatment effect may be described as a difference in group outcomes or as the outcome in (each of) all groups. Measures of variability include standard deviations, standard errors, confidence intervals, interquartile ranges (or other quartile ranges), and ranges. Point measures and/or measures of variability may be provided graphically (eg, standard deviations may be given as error bars in a figure) as long as it is clear what is being graphed (eg, as long as it is clear whether error bars represent standard deviations or standard errors). Where outcomes are categorical, this criterion is considered to have been met if the number of subjects in each category is given for each group.

Von Maher et al., *Reliability of the PEDro scale for rating quality of randomized controlled trials*. Phys Ther, 2003. 83(8): p. 713-21.

Anhang C: Oswestry Disability Questionnaire

This questionnaire has been designed to give us information as to how your back or leg pain is affecting your ability to manage in everyday life. Please answer by checking **one box in each section** for the statement which best applies to you. We realise you may consider that two or more statements in any one section apply but please just shade out the spot that indicates the statement **which most clearly describes your problem**.

Section 1: Pain Intensity

- I have no pain at the moment
- The pain is very mild at the moment
- The pain is moderate at the moment
- The pain is fairly severe at the moment
- The pain is very severe at the moment
- The pain is the worst imaginable at the moment

Section 2: Personal Care (eg. washing, dressing)

- I can look after myself normally without causing extra pain
- I can look after myself normally but it causes extra pain
- It is painful to look after myself and I am slow and careful
- I need some help but can manage most of my personal care
- I need help every day in most aspects of self-care
- I do not get dressed, wash with difficulty and stay in bed

Section 3: Lifting

- I can lift heavy weights without extra pain
- I can lift heavy weights but it gives me extra pain
- Pain prevents me lifting heavy weights off the floor but I can manage if they are conveniently placed eg. on a table
- Pain prevents me lifting heavy weights but I can manage light to medium weights if they are conveniently positioned
- I can only lift very light weights
- I cannot lift or carry anything

Section 4: Walking*

- Pain does not prevent me walking any distance
- Pain prevents me from walking more than 2 kilometres
- Pain prevents me from walking more than 1 kilometre
- Pain prevents me from walking more than 500 metres
- I can only walk using a stick or crutches
- I am in bed most of the time

Section 5: Sitting

- I can sit in any chair as long as I like
- I can only sit in my favourite chair as long as I like
- Pain prevents me sitting more than one hour
- Pain prevents me from sitting more than 30 minutes
- Pain prevents me from sitting more than 10 minutes
- Pain prevents me from sitting at all

Section 6: Standing

- I can stand as long as I want without extra pain
- I can stand as long as I want but it gives me extra pain
- Pain prevents me from standing for more than 1 hour
- Pain prevents me from standing for more than 30 minutes
- Pain prevents me from standing for more than 10 minutes
- Pain prevents me from standing at all

Section 7: Sleeping

- My sleep is never disturbed by pain
- My sleep is occasionally disturbed by pain
- Because of pain I have less than 6 hours sleep
- Because of pain I have less than 4 hours sleep
- Because of pain I have less than 2 hours sleep
- Pain prevents me from sleeping at all

Section 8: Sex Life (if applicable)

- My sex life is normal and causes no extra pain
- My sex life is normal but causes some extra pain
- My sex life is nearly normal but is very painful
- My sex life is severely restricted by pain
- My sex life is nearly absent because of pain
- Pain prevents any sex life at all

Section 9: Social Life

- My social life is normal and gives me no extra pain
- My social life is normal but increases the degree of pain
- Pain has no significant effect on my social life apart from limiting my more energetic interests e.g. sport
- Pain has restricted my social life and I do not go out as often
- Pain has restricted my social life to my home
- I have no social life because of pain

Section 10: Travelling

- I can travel anywhere without pain
- I can travel anywhere but it gives me extra pain
- Pain is bad but I manage journeys over two hours
- Pain restricts me to journeys of less than one hour
- Pain restricts me to short necessary journeys under 30 minutes
- Pain prevents me from travelling except to receive treatment

Von Fairbank JCT & Pynsent, PB *The Oswestry Disability Index*. Spine, 2000. 25(22):p. 2940-2953.

Anhang D: Quebec Back Pain Disability Scale

This questionnaire is about the way your back pain is affecting your daily life. People with back problems may find it difficult to perform some of their daily activities. We would like to know if you find it difficult to perform any of the activities listed below, because of your back. For each activity there is a scale of 0 to 5. Please choose one response option for each activity (do not skip any activities) and circle the corresponding number.

Today, do you find it difficult to perform the following activities because of your back?

		0 Not difficult at all	1 Minimally difficult	2 Somewhat difficult	3 Fairly difficult	4 Very difficult	5 Unable to do
1	Get out of bed	0	1	2	3	4	5
2	Sleep through the night	0	1	2	3	4	5
3	Turn over in bed	0	1	2	3	4	5
4	Ride in a car	0	1	2	3	4	5
5	Stand up for 20-30 minutes	0	1	2	3	4	5
6	Sit in a chair for several hours	0	1	2	3	4	5
7	Climb one flight of stairs	0	1	2	3	4	5
8	Walk a few blocks (300-400 m)	0	1	2	3	4	5
9	Walk several kilometres	0	1	2	3	4	5
10	Reach up to high shelves	0	1	2	3	4	5
11	Throw a ball	0	1	2	3	4	5
12	Run one block (about 100m)	0	1	2	3	4	5
13	Take food out of the refrigerator	0	1	2	3	4	5
14	Make your bed	0	1	2	3	4	5
15	Put on socks (pantyhose)	0	1	2	3	4	5
16	Bend over to clean the bathtub	0	1	2	3	4	5
17	Move a chair	0	1	2	3	4	5
18	Pull or push heavy doors	0	1	2	3	4	5
19	Carry two bags of groceries	0	1	2	3	4	5
20	Lift and carry a heavy suitcase	0	1	2	3	4	5

Add the numbers for a total score: _____

Minimum detectable change (90% confidence) 15 points

Von Kopec, JA et al. *The Quebec Back Pain Disability Scale*. Spine, 1995. 20 (3):p. 341-352.

Anhang E: Roland-Morris Disability Questionnaire

When your back hurts, you may find it difficult to do some of the things you normally do.

This list contains sentences that people have used to describe themselves when they have back pain. When you read them, you may find that some stand out because they describe you today.

As you read the list, think of yourself today. When you read a sentence that describes you today, put a tick against it. If the sentence does not describe you, then leave the space blank and go on to the next one. Remember, only tick the sentence if you are sure it describes you today.

1. I stay at home most of the time because of my back.
2. I change position frequently to try and get my back comfortable.
3. I walk more slowly than usual because of my back.
4. Because of my back I am not doing any of the jobs that I usually do around the house.
5. Because of my back, I use a handrail to get upstairs.
6. Because of my back, I lie down to rest more often.
7. Because of my back, I have to hold on to something to get out of an easy chair.
8. Because of my back, I try to get other people to do things for me.
9. I get dressed more slowly than usual because of my back.
10. I only stand for short periods of time because of my back.
11. Because of my back, I try not to bend or kneel down.
12. I find it difficult to get out of a chair because of my back.
13. My back is painful almost all the time.
14. I find it difficult to turn over in bed because of my back.
15. My appetite is not very good because of my back pain.
16. I have trouble putting on my socks (or stockings) because of the pain in my back.
17. I only walk short distances because of my back.
18. I sleep less well because of my back.
19. Because of my back pain, I get dressed with help from someone else.
20. I sit down for most of the day because of my back.
21. I avoid heavy jobs around the house because of my back.
22. Because of my back pain, I am more irritable and bad tempered with people than usual.
23. Because of my back, I go upstairs more slowly than usual.
24. I stay in bed most of the time because of my back.

Von Roland MO, Morris RW. *A study of the natural history of back pain. Part 1: Development of a reliable and sensitive measure of disability in low back pain.* Spine, 1983. 8:p. 141-144

Anhang F: ADL-Assessment nach Manniche

Disability Index

Questions (Table 1 page 319)

- (1) Can you sleep at night without low back pain interfering?
- (2) Can you do your daily work without low back pain reducing your activities?
- (3) Can you do the easy chores at home such as watering flowers or cleaning the table?
- (4) Can you put on shoes and stockings by yourself?
- (5) Can you carry two full shopping bags (10 kilograms total)?
- (6) Can you get up from a low armchair without difficulty?
- (7) Can you bend over the wash basin to brush your teeth?
- (8) Can you climb stairs from one floor to another without resting because of low back pain?
- (9) Can you walk 400 meters without resting because of low back pain?
- (10) Can you run 100 meters without resting because of low back pain?
- (11) Can you ride a bike or drive a car without feeling any low back pain?
- (12) Does low back pain influence your emotional relationship to your nearest family?
- (13) Did you have to give up contact with other people within the last 2 weeks because of low back pain?
- (14) If it was a present interest do you think that there are certain jobs which you would not be able to manage because of your back trouble?
- (15) Do you think that the low back pain will influence your future?

Responses	Points	Forward	Reverse
not a problem	0	yes	no
can be a problem	1	can be	can be
is a problem	2	no	yes

Forward questions: 1 2 3 4 5 6 7 8 9 10 11

Reverse questions: 12 13 14 15

NOTE: In the paper scoring is given as yes = 0; can be problem = 1; no = 2. However these responses for the last 4 questions reverse the general trend of the first 11 questions. It makes more sense to me to reverse the scoring for the last 4 questions.

disability index = SUM(points for all 15 questions)

Von Manniche C Asmussen K et al. *Low back pain rating scale: Validation of a tool for assessment of low back pain*. Pain, 1994. 57(3):p. 317-326.

Anhang G: Pain and Disability Million Index

DATE: _____ NAME: _____

PLEASE MAKE AN "X" ALONG THE LINE TO SHOW HOW FAR FROM NORMAL TOWARD THE WORST POSSIBLE SITUATION YOUR PAIN PROBLEM HAS TAKEN YOU DURING THE PAST WEEK.

1. How bad is your pain? .
no pain _____ worst possible
2. How bad is the pain at night?
no pain _____ worst possible
3. Does the pain interfere with your lifestyle?
no problem _____ total change in lifestyle
4. How good are pain killers for your pain?
complete relief _____ no relief
5. How stiff is your injured area?
no stiffness _____ worst possible stiffness
6. Does your pain interfere with walking?
no problem _____ cannot walk
7. Do you hurt when walking?
no pain _____ worst possible pain
8. Does your pain keep you from standing still?
can stand as long as I want _____ cannot stand at all
9. Does your pain keep you from twisting?
no problem _____ cannot twist
10. Does your pain allow you to sit in an upright hard chair?
sit as long as I like _____ cannot use a hard chair at all
11. Does your pain allow you to sit in a soft arm chair?
sit as long as I like _____ cannot use a soft chair at all
12. Does your injured body part hurt when lying in bed?
no pain _____ no relief at all
13. How much does your pain limit your normal lifestyle?
no limit _____ cannot do anything
14. Does your pain interfere with your activities of daily living or work?
no problem _____ major problem
15. How much have you had to change your home or work place activities because of pain?
no change _____ a great deal of change

Von Anagnostis C. et al. *The Million Visual Analog Scale. Its Utility For Predicting Tertiary Rehabilitation Outcomes.* Spine, 2003. 28(10): p. 1051-1060.