

Aus der Abteilung für
Allgemeinmedizin, Präventive und Rehabilitative Medizin
des Fachbereichs Medizin der Philipps-Universität Marburg

Leitung: Prof. Dr. med. Annette Becker, MPH

**Subakuter und chronischer Rückenschmerz in der
hausärztlichen Versorgung:
Eine systematische Übersichtsarbeit symptomevaluierender
Studien**

Inaugural-Dissertation zur Erlangung des Doktorgrades der
gesamten Humanmedizin
im Fachbereich Medizin der Philipps-Universität Marburg

vorgelegt von
Laura Annabelle Sophie Schmidt (Wuppertal)
Marburg, 2017

Angenommen vom Fachbereich Medizin der Philipps-Universität Marburg
am: 10. Juli 2017

Gedruckt mit Genehmigung des Fachbereichs.

Dekan: Prof. Dr. H.Schäfer
Referent: Prof. Dr. A. Becker
Korreferent: Prof. Dr. J. Kruse

Inhaltsverzeichnis

1. Abstract.....	1
2. Einleitung	3
2.1. Rückenschmerz	3
2.2. Symptomevaluierende Studien	5
2.3. Systematische Übersichtsarbeiten.....	6
2.4. Ziel der Arbeit	7
3. Methodik:.....	8
3.1. Literatursuche:.....	8
3.2. Festlegung der Ein- und Ausschlusskriterien	10
3.3. Bewertung der Abstracts.....	12
3.4. Suche der Studienvolltexte.....	13
3.5. Bewertung der Studienvolltexte.....	13
3.6. Datenextraktion.....	13
3.7. Mathematische Auswertung der Prognoseergebnisse.....	15
3.7.1. Prozentualer Baseline-Wert (PBW).....	15
3.7.2. Prozentualer Ergebniswert (PEW).....	16
3.7.3. Prozentuale Veränderung/Verbesserung (PV).....	16
3.8. Beurteilung der Studienqualität.....	17
4. Ergebnisse.....	20
4.1. Allgemeine Ergebnisse	20
4.1.1. Ergebnisse der Literatursuche.....	20
4.1.2. Allgemeine Studienmerkmale.....	24
4.1.3. Ergebnisse der Qualitätsbeurteilung.....	30
4.1.3.1. Allgemeine Studiendetails (Domäne A und B).....	30
4.1.3.2. Ätiologieergebnisse (Domäne C).....	33
4.1.3.3. Prognoseergebnisse (Domäne D).....	33
4.2. Studienergebnisse.....	34
4.2.1. Ergebnisse im Bereich Ätiologie.....	34
4.2.2. Ergebnisse im Bereich Prävalenz.....	34
4.2.3. Ergebnisse im Bereich Prognose	35
4.2.3.1. Allgemeines	35
4.2.3.2. Subakuter Rückenschmerz	36
4.2.3.2.1. Lebensqualität.....	37
4.2.3.2.2. Schmerz.....	39
4.2.3.2.3. Funktionelle Beeinträchtigung.....	39
4.2.3.2.4. Arbeitsbezogene Ergebnisse.....	40
4.2.3.2.5. Abhängigkeit des Verlaufs vom Ausgangswert.....	40
4.2.3.2.6. Fazit.....	41
4.2.3.3. Chronischer Rückenschmerz	42
4.2.3.3.1. Funktionelle Beeinträchtigung.....	43
4.2.3.3.2. Lebensqualität	44
4.2.3.3.3. Stimmung.....	48
4.2.3.3.4. Schmerz.....	48
4.2.3.3.5. Arbeitsbezogene Ergebnisse.....	50
4.2.3.3.6. Verlauf	50
4.2.3.3.7. Abhängigkeit des Verlaufs vom Ausgangswert.....	51
4.2.3.3.8. Fazit.....	52
4.2.3.4. Subakuter und chronischer Rückenschmerz (gemischte Gruppe)	52
4.2.3.4.1. Funktionelle Beeinträchtigung	53
4.2.3.4.2. Schmerz	55
4.2.3.4.3. Lebensqualität.....	57
4.2.3.4.4. Arbeitsbezogene Ergebnisse	57
4.2.3.4.5. Stimmung.....	57
4.2.3.4.6. Verlauf.....	58

4.2.3.4.7. Abhängigkeit des Verlaufs vom Ausgangswert.....	59
4.2.3.4.8. Fazit.....	59
4.2.3.5. Rückenschmerz mit Ausstrahlung in die Beine.....	60
4.2.3.5.1. Schmerz.....	60
4.2.3.5.2. Funktionelle Beeinträchtigung	61
4.2.3.5.3. Arbeitsbezogene Ergebnisse.....	62
5. Diskussion.....	63
5.1. Zusammenfassung der Ergebnisse.....	63
5.2. Methoden	63
5.2.1. Schlussfolgerung	67
5.3. Diskussion der Ergebnisse.....	68
5.3.1. Grundlegendes.....	68
5.3.2. Diskussion und Interpretation der Prävalenzdaten.....	71
5.3.3. Diskussion und Interpretation der Ätiologiedaten.....	73
5.3.4. Diskussion und Interpretation der Prognosedaten	74
5.3.4.1. Subakuter Rückenschmerz	74
5.3.4.2. Chronischer Rückenschmerz.....	76
5.3.4.3. Subakuter und chronischer Rückenschmerz (gemischte Gruppe).....	78
5.3.4.4. Rückenschmerz mit Ausstrahlung in die Beine.....	81
5.4. Fazit	83
6. Literaturverzeichnis:.....	84
7. Anlagen.....	91
8. Akademische Lehrer:.....	94
9. Danksagung:.....	95

Abbildungsverzeichnis

Abbildung 1: Ausschlussprozess der Abstracts und Publikationen.....	21
Abbildung 2: Lebensqualität - Durchschnittswert der Subskalen (subakuter Rückenschmerz).....	38
Abbildung 3: Abhängigkeit des Verlaufs vom Ausgangswert (Subakuter Rückenschmerz).	41
Abbildung 4: Funktionelle Beeinträchtigung (chronischer Rückenschmerz).....	43
Abbildung 5: Lebensqualität (chron. Rückenschmerz) gemittelte Werte der Subskalen	46
Abbildung 6: Schmerz (chronischer Rückenschmerz).....	49
Abbildung 7: Abhängigkeit des Verlaufs vom Ausgangswert (chron. Rückenschmerz)	51
Abbildung 8: Funktionelle Beeinträchtigung (subakuter/chronischer Rückenschmerz)	54
Abbildung 9: Vergleich der funktionellen Beeinträchtigung aller Subgruppen.....	55
Abbildung 10: Schmerz - gemischte Rückenschmerzpatienten (subakut/chronisch).....	56
Abbildung 11: Schmerz - Analyse aller Subgruppen	56
Abbildung 12: Lebensqualität – gemischte Rückenschmerzpatienten (subakut/chronisch)	57
Abbildung 13: Abhängigkeit des Verlaufs vom Ausgangswert gemischte Rückenschmerzpatienten (subakut/chronisch).....	59
Abbildung 14: Schmerz (Rückenschmerz + Ausstrahlung).....	61
Abbildung 15: Funktionelle Beeinträchtigung (Rückenschmerz + Ausstrahlung).....	61

Tabellenverzeichnis

Tabelle 1: Suchsyntax 29.10.2012.....	9
Tabelle 2: Ein- und Ausschlusskriterien der Abstracts.....	10
Tabelle 3: Ein- und Ausschlusskriterien der Volltexte.....	13
Tabelle 4: Allgemeine Studieninformationen.....	24
Tabelle 5: Definition von "usual care" in den eingeschlossenen Studien.....	26
Tabelle 6: Ein- und Ausschlusskriterien in den eingeschlossenen Studien.....	30
Tabelle 7: Qualitätsbeurteilung der Studien.....	32
Tabelle 8: Biasrisiko-Beurteilung.....	33
Tabelle 9: Risiko für klinische Heterogenität/Varianz.....	33
Tabelle 10: Qualitätsbeurteilung der Ätiologieergebnisse.....	34
Tabelle 11: Qualitätsbeurteilung der Prognoseergebnisse	34
Tabelle 12: Übersicht der Messinstrumente	37
Tabelle 13: Lebensqualität (subakuter Rückenschmerz).....	40
Tabelle 14: Schmerz und funktionelle Beeinträchtigung (subakuter Rückenschmerz).....	40
Tabelle 15: Arbeitsbezogene Ergebnisse (subakuter Rückenschmerz).....	41
Tabelle 16: Gemittelte Teilergebnisse, Lebensqualität Subskalen (chronischer Rückenschmerz).....	45
Tabelle 17: Gemitteltes Gesamtergebnis, Subskalen Lebensqualität (chronischer Rückenschmerz).....	45
Tabelle 18: Lebensqualität (chronischer Rückenschmerz).....	49
Tabelle 19: Stimmung (chronischer Rückenschmerz).....	49
Tabelle 20: Arbeitsbezogene Ergebnisse und Verlaufsergebnisse (chronischer Rückenschmerz).....	52
Tabelle 21: Ergebnisse gemischter Rückenschmerzpatienten (subakut und chronisch).....	60

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ICPC	International Code of Primary Care
IQR	Interquartilsabstand (Inter quartile range)
k.A.	Keine Angabe
KI	Konfidenzintervall
M	Mittelwert
Md	Median
N	Anzahl
NHP	Nottingham Health Profile (Lebensqualität Messinstrument)
(r)ODQ	(revised) Oswestry Disability Questionnaire (Beeinträchtigungs Messinstrument)
RCT	Randomisiert kontrollierte Studie
RF	Red flags
RMQ	Roland-Morris Disability Questionnaire
s	Standardabweichung
PEW	Prozentualer Ergebniswert
SF-36	SF-36 (Lebensqualität Messinstrument)
PBW	Prozentualer Baseline-Wert
PV	Prozentuale Veränderung/Verbesserung
VAS	Visuelle Analogskala (Schmerz Messinstrument)

1. Abstract

Background: Subacute and chronic back pain is a frequently seen symptom in the general practice. Reasons for encounter in general practice are usually symptoms. Therefore symptom based studies, which are subject of this review, evaluate important characteristics of the primary care setting: Prognosis, prevalence and etiology are the most interesting outcomes.

Objective: This systematic review analyses the literature from the past 50 years assessing the results of symptom based studies about prognosis, prevalence and etiology of subacute and chronic back pain in a primary care setting in order to show the current state of knowledge.

Methods: Studies were identified through systematic search of abstracts in Pubmed and through identification of relevant related publications in the literature list of the respective studies, until 2015. The identified abstracts were independently rated by two reviewers according to the following predefined criteria: qualitative design of included studies, primary care setting, symptom based study design, and subacute or chronic back pain with a minimum duration of 6 weeks and prognostic/prevalence/etiologic outcome. Bibliographic data, methodological items and outcome data from full text manuscripts of the eligible studies were extracted. A standardized outcome tool was applied to enable quality assessment. To ensure comparability of different outcomes, questionnaires percentage outcome measurements were used: percentage baseline value, percentage outcome value and a percentage change.

Results: 3794 abstracts were identified that lead to a total of 885 publications. Finally 26 publications from 15 studies were included, matching the topic of this review. Due to the heterogeneity of the included studies a meta-analysis could not be performed and thus a descriptive analysis was used instead. The vast majority of the studies obtained good to average quality results and no study had to be excluded for quality reasons.

Only one study provided aetiological data in which the authors found a prevalence of 23.6 % (CI: 19.4-28.3) for axial spondyloarthritis, 22 % of those showed results of a spondylitis ankylosans.

Two studies reported results on the prevalence of chronic back pain. The first one found 1.3 % of the practice visits to be caused by chronic back pain patients. The second one declared that 5.95 % of chronically ill patients received an ICPC code for chronic back pain (L84, L86).

Most of the results referred to prognostic outcomes. The key outcomes were quality of life, pain, disability, work-related aspects and mental health. Average percentage change (PC) over 6 months for quality of life ranged from -2.9 to 23.6 % with great variance for each subscale. With regard to pain the results ranged from -2.1 to 11 % (PC), and the vast majority of studies showed an improvement with time. The disability outcome ranged from +5.8 to +6.8 % (PC). 48 % to 58 % of those back pain patients, who were initially on sick-leave, returned to work after 1 year. The category of mental outcomes included many different questionnaires about diverse aspects of mental health and again the data showed improvement over time.

Conclusion: This systematic review of symptom evaluating studies overall showed an improvement of prognostic outcomes for subacute and chronic back pain. There is scarce literature on prevalence and etiology in primary care.

2. Einleitung

2.1. Rückenschmerz

Rückenschmerz als Krankheitsbild ist zunächst für jedermann ein Begriff und scheint in seiner Bedeutung eindeutig. Doch Rückenschmerz zeichnet sich bei eingehender Betrachtung gerade durch seine Komplexität aus. Dies erschwert die einheitliche Erforschung und führt teilweise zu einer inhomogenen und widersprüchlichen Studienlage. [58] Vor allem folgende Aspekte können variieren: Die genaue Lokalisation und Ausdehnung der Schmerzen, die Ausstrahlung der Schmerzen, ihre Dauer und deren Verlauf sowie das Behandlungsumfeld und die Ätiologie der Schmerzen.

Subakuter und chronischer Rückenschmerz ist allgegenwärtig in allen Bevölkerungsschichten und jeder medizinischen Versorgungsebene. Dementsprechend groß ist seine Bedeutung. So zeigten Studien, dass Rückenschmerz der häufigste Grund chronischer Schmerzen in der Bevölkerung ist. [11, 23, 27, 29, 35, 58, 67, 80] Die Prävalenz von chronischem Rückenschmerz in der Allgemeinbevölkerung liegt zwischen 1-25 %. [5, 27, 40, 60, 67] Davon befindet sich der überwiegende Teil der Betroffenen in ambulanter Behandlung. [83] Die bisherige Datenlage legt nahe, dass im hausärztlichen Umfeld etwa 2.2 % der Besuche wegen chronischem Rückenschmerz und 0.8 % wegen subakutem Rückenschmerz stattfinden. [37] Es zeigte sich zudem, dass Frauen im Schnitt häufiger betroffen sind als Männer. [60, 81]

Die Länge der Beschwerden und die ihr zugeordneten Kategorien „akut, subakut und chronisch“ sind in der Wissenschaft nicht vollkommen einheitlich. [2, 14, 94] Im Rahmen dieser Arbeit orientieren wir uns an der Definition der „Nationalen Versorgungsleitlinie“. [2] Danach sind subakute Rückenschmerzen zwischen 6 und 12 Wochen, chronische Rückenschmerzen mehr als 12 Wochen lang. Ein wichtiges Ziel im Bereich der subakuten Rückenschmerzen ist es, die Entstehung chronischer Schmerzen zu vermeiden. Dies geht unter anderem mit der Erfassung von Risikofaktoren („yellow flags“) einher und einer intensivierten multimodalen Therapie. Im Bereich der chronischen Beschwerden ist insbesondere die Patientenedukation, in Hinblick auf die biopsychosozialen Komponenten der Beschwerden und der Umgang mit diesen im Sinne einer Vermeidung von funktionellen Beeinträchtigungen und der Entwicklung einer Behinderung oder Pflegebedürftigkeit, wichtig. [2]

Rückenschmerz hat durch direkte und indirekte Kosten große volkswirtschaftliche Auswirkungen. [17, 18, 22] Der Kostenfaktor ist ein nicht zu vernachlässigender Aspekt, da auch im Gesundheitssystem ökonomische Betrachtungen seit vielen Jahren an Bedeutung gewinnen. So konnten Studien aus den USA zeigen, dass Rückenschmerz in seiner Gesamtheit die sechststeerste Erkrankung und sogar die viertteuerste Erkrankung aus Sicht der dortigen Arbeitgeber ist. [17] Dabei generiert chronischer Rückenschmerz den größten Anteil der Kosten. [83] Im ambulanten Bereich belaufen sich die direkten Ausgaben auf etwa \$35,7 Milliarden oder \$3152 pro Person über zwei Jahre (2006/07, USA). [83]

Europäische Studien dokumentierten jährliche direkte Kosten in Höhe von 3090€ pro Person. Die indirekten Kosten pro Person lagen etwa 5-mal so hoch, bei 17.600€, wobei die Arbeitsunfähigkeit anteilig die größten Kosten verursacht. [22] Neuere Studien zeigten zudem, dass die Häufigkeit von chronischem Rückenschmerz zunimmt. Hochrechnungen zufolge könnten 2020 bis zu 20 Millionen Menschen in den USA chronischen Rückenschmerz haben. [83] Ein Grund für diese Entwicklung ist auch der demografische Wandel, denn die Prävalenz von Rückenschmerzpatienten steigt mit dem Alter an und Betroffene begeben sich mit zunehmendem Alter häufiger in Behandlung. [83]

Rückenschmerz ist ein intensiv beforschtes Feld. Seine Bedeutung begründet sich zum einen in seiner großen Verbreitung und zum anderen in seinen ökonomischen Folgen.

Es ist ein komplexes Krankheitsbild. Bei vielen Betroffenen, lässt sich keine diagnostizierbare körperliche Ursache für die Beschwerden finden. Dies resultiert in der Forschung in der Unterscheidung zwischen spezifischen und nichtspezifischen Rückenschmerz. Beispiele für spezifischen Rückenschmerz sind Frakturen, Osteoporose, rheumatische Erkrankungen, Bandscheibenvorfälle oder auch Karzinome und deren Metastasen. Bei etwa 85% der Betroffenen lässt sich jedoch kein körperliches Korrelat finden. [3]

Dies macht die Unterscheidung zwischen Beratungsanlass (nicht-klassifizierter Rückenschmerz) und Beratungsergebnis (spezifischer/nichtspezifischer Rückenschmerz; Diagnose), insbesondere in Datenbankstudien, schwierig.

2.2. Symptomevaluierende Studien

Symptomevaluierende Studien befassen sich mit dem Beratungsanlass der Patienten. Dadurch grenzen sie sich von Studien ab, die Patienten aufgrund einer bereits gestellten Diagnose einschließen.

Ihr Aufbau spiegelt den tatsächlichen Alltag der Primärversorgung besser wider, deren Aufgabe im Gesundheitswesen die Filterung der Patienten auf Basis von Beschwerden darstellt. So entsprechen die Studienergebnisse symptomevaluierender Studien mit höherer Wahrscheinlichkeit den wahren Werten für das hausärztliche Patientenkollektiv. Die Ergebnisse symptomevaluierender Studien sollen dazu führen, dass primärversorgende Ärzte empirisch basierte Daten zur Häufigkeit der präsentierten Beschwerden, der zugrundeliegenden Ätiologien und zur Prognose erlangen.

Von besonderem Interesse ist diesem Zusammenhang die Frage nach Ursachen für subakuten und chronischen Rückenschmerz und deren Häufigkeitsverteilung. Eng mit diesem Aspekt verbunden ist die Methode des „probabilistic reasoning“, bei der klinische Entscheidungen auf Basis der zugrundeliegenden Prätest-Wahrscheinlichkeit bestimmter Ätiologien gefällt werden. [21, 30] Auf der Basis der Prätest-Wahrscheinlichkeit kann nach diagnostischen Tests (in Anamnese, Untersuchung oder weiterführender Diagnostik) sodann die Posttest-Wahrscheinlichkeit berechnet werden. Der zu erwartende diagnostische Gewinn beeinflusst die Entscheidung, ob ein bestimmter Test bzw. eine bestimmte Diagnostik angefordert wird oder nicht. Dies unterstützt den rationalen Einsatz von Diagnostik und hilft Über- und Unterdiagnostik zu vermeiden.

Eine weitere wichtige Fragestellung kommt aus dem Bereich der Prävalenz und Inzidenz. Wie viele Patienten kommen mit einem bestimmten Symptom zu ihrem Arzt? Dies sind Kennzahlen, die eine Einschätzung der Häufigkeitsverteilung der Beschwerden und des damit verbundenen Versorgungsbedarfs möglich machen.

Und schließlich ermöglichen diese Studien eine Einschätzung der Prognose. Ein entscheidender Aspekt, nicht nur für den Patienten, sondern auch für den Arzt. Symptomevaluierende Studien versetzen ihn in die Lage, dem Patienten eine Beratung auf Grundlage evidenzbasierter Daten zu geben.

Damit diese auf einem soliden Fundament gründen, gibt es Qualitätsmerkmale für symptomevaluierende Studien. Wesentliche Aspekte sind eine prospektive und konsekutive Rekrutierung sowie ein multizentrisches Design, um Verzerrungen des Patientenkollektivs im

Prozess der Rekrutierung zu vermeiden bzw. zu erkennen. Die Formulierung klarer Ein- und Ausschlusskriterien sorgt für Transparenz, um deutlich zu machen aus welchen Patienten das Studienkollektiv besteht und welches Setting es repräsentiert. Eine Kontrollgruppe ohne Symptom kann im Bereich Ätiologie und Prognose, die Erforschung abhängiger oder unabhängiger Variablen im Bezug auf das zu erforschende Symptom ermöglichen. Die Transparenz der durchgeführten Untersuchungen/Messungen sowie gleiche Untersuchungen/Messungen für alle Patienten sind sowohl im Bereich ätiologischer als auch prognostischer Fragestellungen von Bedeutung. Eine ausreichend lange Nachbeobachtungsperiode und ausreichend hohe Patientenzahlen erhöhen die Wahrscheinlichkeit, dass die Ergebnisse nahe am wahren Wert liegen. [20]

Sinn dieser Qualitätskriterien ist es somit, unterschiedliche Ausprägungen von Ergebnisverzerrungen zu vermeiden, um Ergebnisse zu erreichen, die in ihrer Aussage repräsentativ sind für die Gesamtheit des untersuchten Patientenkollektivs. Zudem sollen sie die nötige Transparenz schaffen, um die Vergleichbarkeit von Studien und deren qualitative Analyse zu ermöglichen.

2.3. Systematische Übersichtsarbeiten

Rückenschmerzen sind ein intensiv beforschtes Arbeitsfeld, mit einer entsprechend hohen Zahl an Publikationen. Dies geht mit der großen Herausforderung einher, den Überblick über die stetig wachsende Zahl an Publikationen zu behalten. [77] Ein mögliches Werkzeug sind systematische Übersichtsarbeiten, die vor allem dem Zweck dienen, Aussagen vieler einzelner Studien zu einem Gesamtbild zusammenzufügen. Dabei fassen sie deren Aussagen jedoch nicht nur beschreibend zusammen, sondern bewerten ihre Methodik und analysieren deren Ergebnisse. Die Bedeutung systematischer Übersichtsarbeiten in der heutigen Forschung ist deshalb nicht zu unterschätzen.

Um eine hohe Qualität systematischer Übersichtsarbeiten zu gewährleisten, müssen auch diese bestimmte Anforderungen erfüllen. Ziel ist es die Nachteile klassischer Übersichtsarbeiten zu vermeiden, indem methodische Standards sowohl auf die Identifikation als auch die Selektion der einzuschließenden Originalstudien angewandt werden.

Zur Einhaltung dieser Standards sollten die folgenden Kriterien beachtet werden: Bestimmung der Forschungsfrage a priori, Durchführung einer möglichst umfangreichen Literatursuche und Festlegung der Ein- und Ausschlusskriterien im Vorfeld. Der Ausschluss von Studien sollte begründet dargestellt werden, die Extraktion der eingeschlossenen Studien

standardisiert erfolgen ebenso wie die Beurteilung der Studienqualität, um schließlich zu einer Zusammenfassung und Interpretation der Ergebnisse zu kommen. [6, 28, 39, 55, 96] Werden diese Kriterien eingehalten, wird das Risiko für Ergebnisverzerrungen minimiert und es entsteht eine Arbeit von höchster wissenschaftlicher Evidenz.

Insgesamt sind systematische Übersichtsarbeiten symptomevaluierender Studien bisher selten und diese Arbeit damit Teil einer neuen, höchst spannenden und alltagsrelevanten Forschungsrichtung.

2.4. Ziel der Arbeit

Ziel dieser systematischen Übersichtsarbeit symptomevaluierender Studien ist es, den aktuellen wissenschaftlichen Stand zur Prävalenz, Ätiologie und Prognose des Beratungsanlasses subakuter und chronischer Rückenschmerz in der Primärversorgung auszuwerten.

Dies soll erreicht werden, indem die Literatur nach symptomevaluierenden Studien zu diesem Beratungsanlass durchsucht wird. Hierdurch sollen alltagsrelevante Fragen der hausärztlichen Versorgung im Bezug auf subakuten und chronischen Rückenschmerz geklärt werden. Trotz der Häufigkeit dieses Symptoms und der Schlüsselrolle der Primärversorgung bei deren Behandlung, gibt es aktuell noch keine systematische Übersichtsarbeit symptomevaluierender Studien, die sich auf den Sektor der Primärversorgung fokussiert.

Um diese Wissenslücke zu schließen, werden Studien gesucht, die Patienten in der Primärversorgung rekrutieren und deren Teilnehmer sich mit dem Symptom bzw. dem Beratungsanlass „Rückenschmerz“ vorstellen. Aus dieser Auswahl an Studien werden Ergebnisse für Patienten, die zum Zeitpunkt ihrer Rekrutierung mindestens 6 Wochen Beschwerden hatten, in dieser Übersichtsarbeit analysiert. Dabei sollen folgende Fragen beantwortet werden:

Wie ist die Prognose von Patienten mit subakutem oder chronischem Rückenschmerz?

Wie hoch ist die Prävalenz/Inzidenz dieser Beschwerden in der Hausarztpraxis?

Welche Ursachen verbergen sich hinter dem Symptom und wie ist deren Häufigkeitsverteilung?

3. Methodik:

Um eine hohe methodische Qualität der Arbeit zu gewährleisten, dienten sowohl das Handbuch der „Cochrane Collaboration“ [39] als auch die „PRISMA Richtlinien“ [55, 63] als Orientierung. Dabei ist zu beachten, dass diese Checkliste für systematische Übersichtsarbeiten von Diagnose- und Therapiestudien ausgelegt ist. Dementsprechend war eine Anpassung der Kriterien für diese systematische Übersichtsarbeit symptomevaluierender Studien nötig.

Die methodischen Grundlagen für diese systematische Übersichtsarbeit lassen sich in 7 Arbeitsschritte unterteilen: Literatursuche, Festlegung der Ein- und Ausschlusskriterien, Bewertung von Studienzusammenfassungen (im Folgenden auch Abstract(s) genannt), Suche der Publikationen der eingeschlossenen Abstracts, Bewertung der Publikationen, Extraktion der Daten und Bewertung der Studienqualität. Als Zweitbeurteiler standen Studierende der Humanmedizin zur Verfügung, die ebenfalls eine Dissertation im Rahmen der Arbeitsgruppe „systematische Literaturübersichten symptomevaluierender Studien“ anfertigen.

3.1. Literatursuche:

Die Literatur wurde in PubMed und durch eine Schneeballsuche identifiziert. Vor Beginn der Literatursuche wurde eine komplexe Suchsyntax für Pubmed erstellt. Kernvorgaben der Suchsyntax waren die Sprache der Abstracts/Publikationen, die Auswahl von Zeitschriften, die durchsucht werden sollten, die Festlegung von Suchbegriffen (z.B. Synonyme für den Hausarzt und Synonyme für Rückenschmerz und Rückenschmerz mit Ausstrahlung in die Beine) und die Abschnitte der Publikationen in denen gesucht werden sollte.

Zulässige Sprachen waren zunächst Englisch, Deutsch, Niederländisch, Schwedisch, Russisch, Spanisch, Französisch und Italienisch. Letztlich wurden jedoch nur die deutschen und englischen Studien eingeschlossen, da eine verlässliche Übersetzung der anderen Publikationen nicht möglich war.

Durchsucht wurden alle zugänglichen Zeitschriftenjahrgänge bis zum 29.10.2012 sowie im Anschluss noch einmal in einer identischen Suche der Zeitraum 29.10.2012 bis zum 14.02.2015. Dieses Vorgehen sollte eine hohe Sensitivität der Literatursuche gewährleisten. Die Suchsyntax vom 29.10.2012 zeigt Tabelle 1. Die Literatursuche war initial im Bezug auf die Dauer der Symptomatik nicht eingeschränkt, sodass die gesamte Bandbreite von akut, subakut, chronisch sowie Gruppen mit gemischter Symptombdauer identifiziert und bewertet wurden.

Suchsyntax vom 29.10.2012:		
Symptom in allen Schreibweisen im Titel oder Abstract	Back [TIAB] AND pain [TIAB], Back [TIAB] AND pain* [TIAB], Backpain [TIAB], Back [TIAB] AND ache [TIAB], Back [TIAB] AND ache* [TIAB], Backache [TIAB], Lumbar [TIAB] AND pain [TIAB], Lumbar [TIAB] AND pain* [TIAB], Lumbar [TIAB] AND ache [TIAB], Lumbar [TIAB] AND ache* [TIAB], Lumbo* AND pain [TIAB], Lumbo* AND pain* [TIAB], Lumbo* AND ache [TIAB], Lumbo* AND ache* [TIAB], Dorsalgia [TIAB], Sciatica [TIAB]	OR
OR		
Symptom [Schlagwort]	Back [MeSH] AND pain* [MeSH], Backpain [MeSH], Back [MeSH] AND ache [MeSH], Backache [MeSH], Lumbar [MeSH] AND pain* [MeSH], Dorsalgia [MeSH], Sciatica [MeSH],	OR
AND		
Begriff Allgemeinmedizin (verschiedene Schreibweisen) in Titel oder Abstract	“general practitioner” [TIAB], “general practitioners” [TIAB], “general practice” [TIAB], “family practice” [TIAB], “family practitioners” [TIAB], “family practitioner” [TIAB], “family medicine” [TIAB], “family physician” [TIAB], “family physicians” [TIAB], “family doctor” [TIAB], “family doctors” [TIAB], “primary care” [TIAB], “family practices” [TIAB], “GP” [TIAB], “GPs” [TIAB], “GP’s” [TIAB]	OR
OR		
In einem der für unser Fachgebiet relevanten Journale	“BMC Fam Pract” [TA], “Fam Pract” [TA], “J Fam Pract” [TA], “Fam Pract Res J” [TA], “J Am Board Fam Pract” [TA], “Br J Gen Pract” [TA], “J R Coll Gen Pract” [TA], “J Coll Gen Pract” [TA], “J Coll Gen Pract Res News” [TA], “Can Fam Physician” [TA], “Ann Fam Med” [TA], “Aust Fam Physician” [TA], “Scand J Prim Health Care” [TA], “Eur J Gen Pract” [TA], “Arch Fam Med” [TA], “J Gen Intern Med” [TA], “Aten primaria” [TA]	OR
OR		
Begriff Allgemeinmedizin (verschiedene Schreibweisen) in Mail-Adresse oder Institutsname des Autors	“general practice” [AD], “family practice” [AD], “family medicine” [AD], “primary care” [AD], community [AD]	OR
OR		
Allgemeinmedizin [Schlagwort]	“General Practitioners” [MeSH], “Family Practice” [MeSH], “Physicians, Family” [MeSH], “Primary Health Care” [MeSH], “Physicians, Primary Care” [MeSH], “General Practice” [MeSH], “Community Health Services” [MeSH], “Rural Health Services” [MeSH], “Rural Health” [MeSH]	OR
OR		
Sprachen	AND (Dutch[lang] OR English[lang] OR French[lang] OR German[lang] OR Italian[lang] OR Russian[lang] OR Spanish[lang] OR Swedish[lang])	
Limits	Limits: NOT (Editorial[pt] OR Addresses[pt] OR Bibliography[pt] OR Biography[pt] OR Case Reports[pt] OR Comment[pt] OR Dictionary[pt] OR Directory[pt] OR Festschrift[pt] OR Government Publications[pt] OR Historical Article[pt] OR In Vitro[pt] OR Interactive Tutorial[pt] OR Interview[pt] OR Introductory Journal Article[pt] OR Lectures[pt] OR Legal Cases[pt] OR Legislation[pt] OR News[pt] OR Patient Education Handout[pt] OR Portraits[pt] OR Webcasts[pt])	

Tabelle 1: Suchsyntax 29.10.2012

3.2. Festlegung der Ein- und Ausschlusskriterien

Im Vorfeld wurde eine Liste von Ein- und Ausschlusskriterien verfasst. Diese wurde genutzt, um zunächst alle Abstracts zu beurteilen und im nächsten Schritt die Publikationen der eingeschlossenen Abstracts. Die Ein- und Ausschlusskriterien waren auf Ebene der Abstracts weniger eng gefasst, um den Einschluss aller relevanten Arbeiten zu gewährleisten. (Tab. 2) Bei der Beurteilung der vollständigen Publikation (im Folgenden auch Volltext genannt) kamen strengere Kriterien zum Einsatz, um schließlich eine hohe Einschliessicherheit relevanter Studien zu gewährleisten. (Tab. 3)

Ein- und Ausschlusskriterien der Abstracts	
Kategorien	Kriterien
Studiendesign	<p>Einschlägig:</p> <ul style="list-style-type: none"> • Originalarbeit mit quantitativen Design (randomisierte kontrollierte Therapiestudie (RCT), Kohortenstudie, Fall-Kontrollstudie) • Rückenschmerzpatienten mussten die Beobachtungseinheit darstellen <p>Nicht-einschlägig:</p> <ul style="list-style-type: none"> • Qualitative Studien, systematische oder narrative Übersichtsarbeiten, Fallberichte
Symptom	<p>Einschlägig:</p> <ul style="list-style-type: none"> • Patienten mit dem primären oder sekundären Beratungsanlass „Rückenschmerz“ • Alle Schmerzlokalisationen im Rücken, außer der Nackenregion, waren zulässig <p>Nicht-einschlägig:</p> <ul style="list-style-type: none"> • Rückenschmerz war nicht der Beratungsanlass
Rekrutierungsort:	<p>Einschlägig:</p> <ul style="list-style-type: none"> • Die Rekrutierung musste in der Primärversorgung statt finden (Hausarzt, Allgemeinmediziner, allgemeiner Internist, Familienarzt, Kinderarzt) <p>Nicht-einschlägig:</p> <ul style="list-style-type: none"> • Sekundärversorgung: Notaufnahmen, spezialisierte Fachärzte, Krankenhäuser, an ein Krankenhaus angegliederte Ambulanzen („outpatient clinics“) • Tertiärversorgung: Rehabilitationszentren • Betriebsärzte, Militärärzte

Tabelle 2: Ein- und Ausschlusskriterien der Abstracts

Ein- und Ausschlusskriterien der Volltexte	
Kategorien	Kriterien
Studiendesign	<p>Einschlägig:</p> <ul style="list-style-type: none"> • Originalarbeit mit quantitativem Design (randomisierte kontrollierte Therapiestudie (RCT), Kohortenstudie, Fall-Kontrollstudie) • mindestens eine Patientengruppe musste die reguläre Hausarztbehandlung erhalten („usual care“), d.h. die Behandlungsfreiheit des Hausarztes ist nicht eingeschränkt oder fremd bestimmt • Eingeschlossen wurden Studien auch dann, wenn sie schriftliche Informationen zum Beratungsanlass verteilten. Je komplexer die Vermittlung der Informationen, desto eher wurden Studien ausgeschlossen. Ausgeschlossen wurden zudem Studien mit schematisierten Medikamentenplan oder standardisierten Rückenübungen. <p>Nicht-einschlägig:</p> <ul style="list-style-type: none"> • Qualitative Studien, systematische oder narrative Übersichtsarbeiten, Fallberichte
Primärversorgung/ Rekrutierungsort:	<p>Einschlägig:</p> <ul style="list-style-type: none"> • Die Rekrutierung musste in der Primärversorgung stattfinden Dazu zählten: Hausarzt, Allgemeinmediziner, allgemeiner Internist, Familienarzt („family doctor“), Kinderarzt <p>Nicht-einschlägig:</p> <ul style="list-style-type: none"> • Sekundärversorgung: Notaufnahmen, spezialisierte Fachärzte, Krankenhäuser, an Krankenhäuser angegliederte Ambulanzen („outpatient clinics“) • Tertiärversorgung, Rehabilitationszentren • Betriebsärzte, Militärärzte • Werbung, Befragung in Wartezimmern die nicht vom Beratungsanlass handelten • Versicherungsdatenbanken mit unklarer Verbindung zum Hausarztsetting • sonstige Settings, die nicht dem obigen Einschlusskriterium entsprachen
Population:	<p>Einschlägig:</p> <ul style="list-style-type: none"> • Rückenschmerzpatienten mussten selbst die Beobachtungseinheit darstellen <p>Nicht einschlägig:</p> <ul style="list-style-type: none"> • Oben genanntes Kriterium nicht erfüllt, • z.B., Hausärzte wurden bezüglich der Thematik Rückenschmerz befragt • Nur Mitarbeiter einer bestimmten Firma wurden bezüglich Rückenschmerz befragt

Ein- und Ausschlusskriterien der Volltexte	
Kategorien	Kriterien
Symptom:	<p>Einschlägig:</p> <ul style="list-style-type: none"> • Mindestens eine Subgruppe stellte Patienten, die wegen Rückenschmerz zum Hausarzt kam und diesen als primären oder sekundären Grund des Besuchs angab • Wurde das Symptom in einer Datenbank gesucht, musste dieses als Beratungsanlass hinterlegt sein • Einschlägiges Codierungssystem war der ICPC (International Classification of Primary Care) • Ischialgien („Sciatica“) und chronischer Rückenschmerz galten nicht als Diagnose, sondern als Symptom bzw. Beratungsanlass • Der Begriff Rückenschmerz wurde wie folgt interpretiert: Er umfasst Schmerzen im gesamten Rücken (ausgenommen die Nackenregion), sowie die potenzielle Ausstrahlung von Schmerzen in die Beine • Subakuter Rückenschmerz wurde definiert als Rückenschmerz mit einer Symptombdauer von 6-12 Wochen (zum Zeitpunkt der Rekrutierung) • Chronischer Rückenschmerz wurde definiert als Rückenschmerz mit einer Symptombdauer von min. 12 Wochen (zum Zeitpunkt der Rekrutierung) • Sowohl Patienten, die zum ersten Mal Rückenschmerzen beklagten als auch Patienten mit rekurrendem Verlauf wurden eingeschlossen <p>Nicht einschlägig:</p> <ul style="list-style-type: none"> • Oben genannte Bedingungen nicht erfüllt. • Insbesondere die Rekrutierung über Diagnosen und Diagnosekodierungen wie: ICD (International Classification of Diseases), Read Codes. • Schwangere mit Rückenschmerzen
Krankheitswahrscheinlichkeit, Selektion:	<p>Einschlägig:</p> <ul style="list-style-type: none"> • Die relevante Rückenschmerzgruppe wurde nicht im Hinblick auf eine bestimmte Symptomursache selektiert • Beispiel: Patienten wurden vom Hausarzt aufgrund der Ein- und Ausschlusskriterien der jeweiligen Studie oder aufgrund des Studien-Behandlungs- bzw. Diagnostikpfades an eine andere Stelle überwiesen <p>Nicht-einschlägig:</p> <ul style="list-style-type: none"> • Oben genannte Bedingungen nicht erfüllt • Beispiel: Patienten wurden vom Hausarzt aufgrund einer Verdachtsdiagnose (selektiert) an einen Facharzt überwiesen und dort für eine Studie rekrutiert.

Table 3: Ein- und Ausschlusskriterien der Volltexte

3.3. Bewertung der Abstracts

Alle Abstracts wurden in Microsoft Access 2010 eingepflegt und von zwei Beurteilern unabhängig voneinander gelesen und nach oben stehenden Kriterien bewertet. Abstracts konnten entweder als „vielleicht einschlägig“ oder „nicht-einschlägig“ bewertet werden. Im Anschluss wurden die Bewertungen von beiden Gutachterinnen verglichen. Einstimmig als „einschlägig“ bewertete Abstracts wurden zur weiteren Bearbeitung eingeschlossen, die einstimmig als „nicht-einschlägig“ bewerteten Abstracts ausgeschlossen. Bei voneinander ab-

weichender Beurteilung wurde der Abstract noch einmal gemeinsam besprochen und eine endgültige Bewertung festgelegt. Im Zweifel wurde die Studie mit einem weiteren Mitarbeiter der Abteilung für Allgemeinmedizin besprochen. War aus einem Abstract nicht eindeutig ersichtlich, ob er „einschlägig“ war oder nicht, wurde er eingeschlossen und erst auf Ebene der Volltextbewertung über die Einschlägigkeit entschieden.

3.4. Suche der Studienvolltexte

Für die eingeschlossenen Abstracts wurden im Folgenden die Studienvolltexte aufgesucht.

3.5. Bewertung der Studienvolltexte

Alle erhältlichen Volltexte wurden von zwei Beurteilern unabhängig von einander gelesen und auf ihre Einschlägigkeit hin überprüft. In Access 2010 wurden sie hinsichtlich der vorformulierten Ein-/Ausschlusskriterien beurteilt. Dafür gab es für jedes Ein- und Ausschlusskriterium (d.h. Studiendesign, Population, Primärversorgung, Symptom, Selektion, Ergebnisse) ein Bewertungsfeld mit den Optionen „einschlägig“/“nicht-einschlägig“. Sobald bei der Durchsicht der Publikation eines der Kriterien als nicht erfüllt angesehen wurde, wurde diese Bewertung in Access übertragen und der Bewertungsprozess beendet. Für jeden Volltext wurde somit nur ein Ausschlussgrund festgehalten, auch wenn es darüber hinaus noch weitere gegeben haben könnte. Schließlich wurde die Bewertung der Beurteiler miteinander verglichen und so geklärt, wie viele Volltexte letztlich „einschlägig“ waren.

3.6. Datenextraktion

Es wurden im Folgendem Studiencharakteristika (siehe unten) sowie Studienergebnisse aus den Bereichen Prognose, Ätiologie und Prävalenz aus den Volltexten extrahiert. Die Volltexte wurden von zwei Personen extrahiert, wobei die Gesamtzahl der zu extrahierenden Volltexte halbiert wurde und je eine Hälfte durch eine Person extrahiert wurde. Die entsprechenden Daten wurden in eine standardisierte SPSS (Version 22) Datenbank eingepflegt.

Im allgemeinen Teil der Studiencharakteristika wurden folgende Aspekte erhoben: Land, Publikationsjahr, Ein- und Ausschlusskriterien, Ausstrahlung in die Beine (eingeschlossen, ausgeschlossen, unklar), Symptomdefinition, Symptombdauer, Arztbeschreibung, „red flags“ (eingeschlossen, ausgeschlossen), Definition von „usual-care“, Rekrutierungszeitraum, Anzahl der Praxen, Anzahl der Ärzte, Anzahl der Patienten, Durchschnittsalter der Patienten, Frauenanteil der Patienten, Altersspannweite (laut Ein- Ausschlusskriterien oder ggf. als Be-

schreibung des Patientenkollektivs), Art der Rekrutierung (In der Praxis bei Konsultation oder durch eine Suche in der Datenbank), Art des berichteten Ergebnisses (Ätiologie/ Prävalenz/ Prognose).

Zunächst wurden alle Prognoseergebnisse aus Fragebögen oder Messinstrumente extrahiert, die folgende Aspekte zum Gegenstand hatten: Lebensqualität, Schmerzintensität, arbeitsbezogene Ergebnisse (Dauer der Krankschreibung, Anzahl der krankgeschriebenen Teilnehmer, etc.), Einschränkungen durch die Beschwerden, Einnahme von Medikamenten, Anzahl der Wiedervorstellungen beim Hausarzt, Verlauf der Rückenschmerzen (Dauer bis zum Ende einer Episode, Wiederauftreten von Beschwerden, Verlauf der Beschwerden (kontinuierlich, rezidivierend, einmalig, etc.), Messung emotionaler Aspekte (Stimmung/Sorgen/Depression/Ängste, etc.). Im Anschluss wurden die gefundenen Ergebnisse gesichtet und darüber entschieden, welche Prognoseparameter eingehender analysiert werden sollten. Dabei fielen folgende indirekte Prognosemesswerte weg: Einnahme von Medikamenten, Anzahl der Wiedervorstellungen beim Hausarzt.

Im Bereich Ätiologie wurde extrahiert, welche Rückenschmerzätiologien in welchem Häufigkeitsverhältnis vorkamen und um welche Beobachtungseinheit es sich handelte (Beratung (Konsultation), Episode, Patienten). Im Bereich Prävalenz wurden entsprechend die Kennzahlen für die Prävalenz oder Inzidenz extrahiert, zusammen mit der entsprechenden Beobachtungseinheit als Freitext (z.B. Konsultationen, Patienten).

3.7. Mathematische Auswertung der Prognoseergebnisse

3.7.1. Prozentualer Baseline-Wert (PBW)

Da nicht jede Studie das gleiche Messinstrument für die gleiche Ergebniskategorie gewählt hat, wurde die prozentuale Veränderung des jeweiligen Ergebniswertes berechnet. Hierzu wurde das Ergebnis in Bezug zum maximalen Skalenwert gesetzt.

$$\frac{\text{Ergebniswert zum Baseline - Zeitpunkt}}{\text{Maximaler Skalenwert}}$$

Wurde beispielsweise der „Roland-Morris Disability Questionnaire“ (Skala 0-24) mit dem „revised Oswestry Disability Questionnaire“ (Skala 0-100) verglichen, wurde folgende Berechnung durchgeführt.

Beispiele:

- Roland-Morris Disability Questionnaire:

Ergebnis: Baseline: 8.5

$$PBW = \frac{8.5}{24} = 0.354$$

- revised Oswestry Disability Questionnaire:

Ergebnis: Baseline: 49.7

$$PBW = \frac{49.7}{100} = 0.497$$

Bei der Interpretation des prozentualen Baseline-Wertes galt es vor allen Dingen zu beachten, wie die jeweilige Skala ausgerichtet ist. Einige der eingeschlossenen Studien verwendeten Skalen, deren bestes Ergebnis der höchste Wert ist, während ein hohes Ergebnis in anderen Messinstrumenten der schlechteste Wert ist. (Tab. 12)

3.7.2. Prozentualer Ergebniswert (PEW)

Auf die gleiche Art und Weise wie beim prozentualen Baseline-Wert wurden auch die Ergebnisse für die jeweiligen Endpunkte prozentual ausgedrückt.

$$\frac{\text{Ergebniswert zum Endpunkt } X}{\text{Maximaler Skalenwert}}$$

- Roland-Morris-Disability-Questionnaire:

Ergebnis: 6 Monate: 7.4

$$PEW = \frac{7.4}{24} = 0.308$$

3.7.3. Prozentuale Veränderung/Verbesserung (PV)

Ebenso wurde die prozentuale Veränderung berechnet, die im Bezug zum maximalen Skalenwert angegeben wird. Damit ist gemeint, dass das Ergebnis zu einem bestimmten Endpunkt mit dem Baseline-Wert in Bezug gesetzt wird und in Prozent angegeben wird.

Prozentuale Veränderung:

$$\frac{\text{Ergebniswert zum Zeitpunkt } X \text{ minus Ergebnis zum Baseline - Zeitpunkt}}{\text{Maximaler Skalenwert}}$$

Beispiele:

- Roland-Morris Disability Questionnaire:

Ergebnisse: Baseline: 8.5; 6 Monate: 7.4

$$PV = \frac{8.5 - 7.4}{24} = 0.0458 * 100 \approx 4.6 \%$$

- revised Oswestry Disability Questionnaire:

Ergebnisse: Baseline: 49.7; 12 Monate: 35.3

$$PV = \frac{49.7 - 35.3}{100} = 0.144 * 100 = 14.4 \%$$

Die unterschiedliche Polarisierung der Messinstrumente (1=bestes Ergebnis oder 1=schlechtestes Ergebnis) wurde gehandhabt, indem die prozentuale Veränderung immer so angegeben wurde, dass eine Verbesserung der Ergebnisse über die Zeit einen positiven Wert besitzt und eine Verschlechterung des Ergebnis einen negativen Wert.

Beispiele:

- NHP (100=schlechtestes Ergebnis):

Ergebnisse: Baseline: 60; 6 Monate: 30

$$PV = \frac{60-30}{100} = 0.3 \times 100 = 30\% \rightarrow \text{wird angegeben als } 30\% \text{ Verbesserung}$$

- SF-36 (100=bestes Ergebnis):

Ergebnisse: Baseline:50; 12 Monate: 40

$$PV = \frac{50-40}{100} = (0.1) \times 100 = 10\% \rightarrow \text{wird angegeben als } (-10)\% \rightarrow \text{Verschlechterung}$$

- SF-36 (100=bestes Ergebnis):

Ergebnisse: Baseline:50; 12 Monate: 70

$$PV = \frac{50-70}{100} = (-0.2) \times 100 = (-20)\% \rightarrow \text{wird angegeben als } 20\% \rightarrow \text{Verbesserung}$$

3.8. Beurteilung der Studienqualität

Zur Beurteilung der Studienqualität der eingeschlossenen Publikationen fand ein „Bias-Fragebogen“ Anwendung, der im Rahmen einer weiteren Promotionsarbeit in Anlehnung an die Publikation von Donner-Banzhoff et al. [20] entwickelt wird.

Es besteht aus vier Domänen (A-D), mit insgesamt 16 Items („I“). Domäne A (I1-I6) und B (I7-I9) bewerten die Auswahl der Patienten und Ärzte der jeweiligen Studie, Domäne C (I10-I12) und D (I13-I16) Aspekte der Studienergebnisse. Jede Domäne wird abschließend im Bezug auf das Risiko für Verzerrungen mit „niedrig, unklar oder hoch“ bewertet.

Domäne A bewertet die Auswahl der Patienten und Ärzte in den folgenden Punkten:

- Wurde das Symptom klar beschrieben (I1)?
- Wurden die Auswahlkriterien der Patienten klar beschrieben (I2)?
- Wurden die Patienten konsekutiv oder zufällig ausgewählt (I3)?
- War das Design multizentrisch (I4)?
- War die Studienpopulation repräsentativ für das volle Spektrum der Rückenschmerzpopulation (I5)?
- Waren die Ärzte repräsentativ für die Primärversorgung (I6)?

Die Wahrscheinlichkeit für eine Verzerrung (Bias) in Domäne A stützt sich vor allem auf die Art der Patientenrekrutierung (konsekutiv vs. zufällig) und das Vorliegen eines multi-zentrischen Designs.

Als Grundlage zur Einschätzung der klinischen Heterogenität/Variation wird das Patientenkollektiv sowie die teilnehmenden Arztpraxen auf ihre Repräsentativität überprüft.

Domäne B bewertet die Datenerfassung:

- Wie wurden die Daten über das Symptom und die Einschlusskriterien erhoben, direkt oder indirekt? (I7)
- Wurde für alle Patienten die gleiche Art der Datenerfassung verwendet? (I8)
- War die Anzahl der Dropouts unter 20 %? (I9)

Diese Signalfragen sind Grundlage zur Beurteilung des Verzerrungsrisikos im Bereich der Datenerhebung.

Domäne C (I10-12) befasst sich mit der Bewertung von Ätiologieergebnissen:

- Wurde die ätiologische Kategorie klar definiert? (I10)
- Waren die diagnostischen Methoden ausreichend um die Ätiologie zu bestimmen? (I11)
- Wurde bei jedem Patienten die gleiche Diagnostik betrieben? (I12)

Im Bereich der Ätiologie stützt sich die Beurteilung des Verzerrungsrisikos vornehmlich auf die Art des diagnostischen Vorgehens und darauf, ob dieses für jeden Teilnehmer gleich ist.

Domäne D (I13-15) beschäftigt sich mit der Bewertung von Prognoseergebnissen:

- War das Prognoseergebnis klar definiert? (I13)
- Gab es eine Vergleichsgruppe ohne Symptom? (I14)
- War die Art der Messung geeignet, um das jeweilige Prognoseergebnis korrekt zu bestimmen? (I15)
- Wurde bei jedem Patienten die gleiche Messung vorgenommen? (I16).

Die Gesamtbeurteilung des Biasrisikos der Domäne D basiert dann, ähnlich der Domäne C, darauf, ob die Art der Messung geeignet ist das jeweilige Prognoseergebnis zu bestimmen und darauf, ob das Vorgehen für alle Teilnehmer gleich ist.

Enthielt eine Publikation keine Information über den Aspekt „Ausstrahlung in die Beine“, wurde dieser Aspekt als unklar bewertet. Wurde die Beurteilung unklar vergeben, erhielt

Item 1 (I1) automatisch die Beurteilung „Nein“. Item 5 (I5) erhielt dementsprechend maximal die Bewertung „unklar“ oder „Nein“, je nachdem, ob noch weitere Gründe für eine schlechte Bewertung sprachen. Im Bezug auf Item 5 (I5) wurde festgelegt, dass die Studie nur dann für alle Rückenschmerzpatienten repräsentativ sein kann, wenn Patienten mit „Ausstrahlung in die Beine“ eingeschlossen wurden.

Wurden Patienten mit „red flags“ nicht explizit ausgeschlossen, wurde davon ausgegangen, dass sie Teil der Studienpopulation waren. Wurden sie ausgeschlossen, wurde Item 5 (I5) dennoch mit „ja“ bewertet, da Patienten mit „red flags“ nur einen geringen Teil der Patientenpopulation ausmachen und daher Ergebnisverzerrungen im Bezug auf die Gesamtpopulation nicht zu erwarten sind.

Studien, die nur zu einem einzigen Zeitpunkt Daten erhoben haben, wurden im Bezug auf Dropouts mit einem „Ja“ beurteilt (I9), da dann vom Design her keine Dropouts möglich waren.

Im Bezug auf Item 3 (I3) „konsekutive vs. zufällige Rekrutierung“ wurden Studien, die Teilnehmer persönlich, prospektiv und konsekutiv rekrutierten, als Optimum angesehen. Studien, die ihre Patienten über die Datenbank/ICPC rekrutierten, wurden im Bezug auf I3, als qualitativ gut bewertet, da davon ausgegangen wurde, dass Suchvorgänge über Datenbanken alle relevanten Patienten identifizieren.

Nach diesem Schema wurde jede Publikation von zwei Beurteilern unabhängig voneinander bewertet, sowohl der allgemeine Teil jeder Studie als auch jedes Ergebnis. Im Folgenden wurde die Qualitätsbeurteilung miteinander abgeglichen und im Zweifel eine endgültige Bewertung festgelegt. Der vollständige Bias-Fragebogen findet sich im Anhang. Weitere Details finden sich in der Dissertationsschrift von K.Hörner, Philipps-Universität Marburg (in Bearbeitung).

4. Ergebnisse

4.1. Allgemeine Ergebnisse

4.1.1. Ergebnisse der Literatursuche

Durch die erste Suche in Pubmed am 29.10.2012, für den gesamten Zeitraum vor diesem Datum, wurden 3219 Studienzusammenfassungen gefunden. Am 14.02.2015, für den Zeitraum zwischen der ersten und zweiten Suche, weitere 575. Somit konnten insgesamt 3794 Abstracts identifiziert werden. Am Ende des Bewertungsprozesses wurden 885 Abstracts als einschlägig beurteilt. Nach Durchsicht der dazugehörigen Studienvolltexte entsprachen 191 Volltexte den Ein- und Ausschlusskriterien. Zusätzlich zu diesen 191 Volltexten konnten weitere 14 Publikationen über die Schneeballsuche identifiziert werden. Insgesamt lieferten 26 Publikationen zu 15 Studien (teilweise mehrere Publikationen einer Studie) Ergebnisse für subakute und/oder chronische Rückenschmerzpatienten. (Abb. 1)

Die Analyse der Studien erfolgt in vier Subgruppen, die sich in wesentlichen Aspekten unterscheiden: subakuter Rückenschmerz, chronischer Rückenschmerz, gemischt subakuter/chronischer Rückenschmerz und Rückenschmerz mit Ausstrahlung in die Beine.

Vier Studien [32, 42, 65, 85, 89] erbrachten Prognoseergebnisse für Patienten mit subakutem Rückenschmerz, definiert als Rückenschmerz mit einer Dauer von 6-12 Wochen.

Acht Studien [26, 48, 76, 78, 86, 89, 91, 93] berichteten Prognoseergebnisse für chronische Rückenschmerzpatienten (>12 Wochen Schmerzen bei Rekrutierung). In dieser Gruppe fanden sich neben Prognoseergebnissen auch zwei Studien, die Daten zur Prävalenz von chronischem Rückenschmerz in der Primärversorgung lieferten [26, 78] sowie eine Studie, die eine ätiologische Fragestellung bearbeitete. [91]

Fünf Studien [31, 34, 45, 51, 52, 62, 69, 70, 90] lieferten Prognoseergebnisse für Rückenschmerzpatienten der Gruppe „gemischt subakut/chronisch“. Diese Subgruppe wurde definiert als Rückenschmerz mit einer Dauer von mindestens 6 Wochen bis zu einer Dauer von über 12 Wochen. Zwei Studien [69, 86] analysierten die Prognose von Rückenschmerzpatienten mit Ausstrahlung in die Beine. Die wichtigsten allgemeinen Informationen zu allen Studien finden sich in Tabelle 4.

Eine Metaanalyse der Studien war nicht möglich, da die verwendeten Messinstrumente gemessen an der Gesamtzahl der identifizierten Studien zu vielseitig waren. Die Analyse verbleibt deshalb deskriptiv.

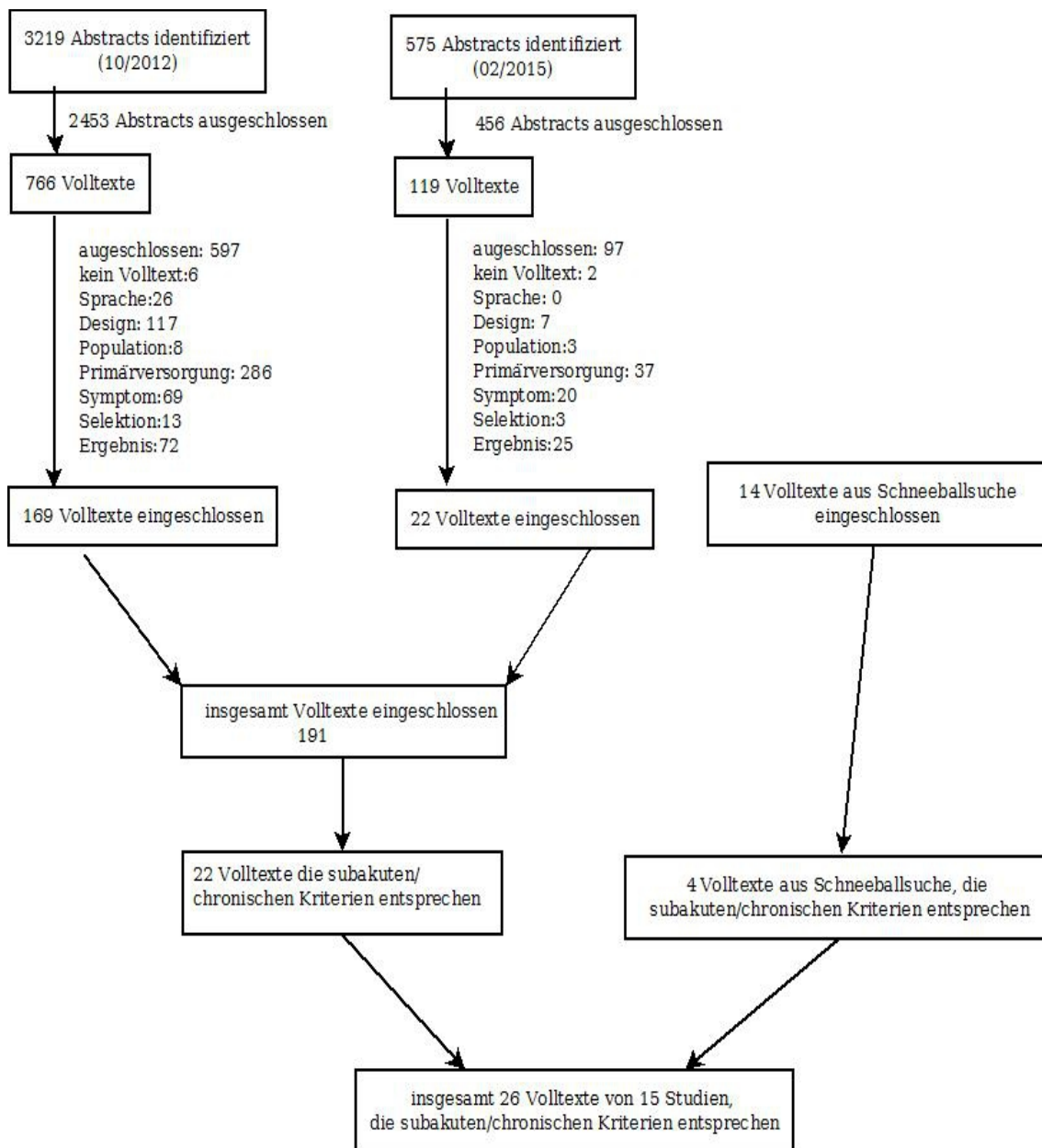


Abbildung 1: Ausschlussprozess der Abstracts und Publikationen

Allgemeine Studieninformationen

Studie	Land	Design	Rekrutierungsart ^h	Pat N	Episodendauer	Praxis N ⁱ	GP N ^j	Alter Ø ^k	Frauen %	Altersgrenzen ^l	Ausstrahlung	Behandlungsumfeld	Rekrutierung
Subakut/chronisch:													
Lamb et al. 2010/2014 [45, 51, 52]	GB	RCT	Praxis/ Datenbank	233	>6 Wochen	56	k.A.	54	61	>18	ja	Hausarzt, nurse practitioner	04/2005-04/2007
Nyiendo et al. 2001/2005 [31, 69, 70]	USA	Kohortenstudie	Praxis	310	>6 Wochen	14	111	39.4	52.6	>18	Ja NR:27.3 ÜK:32.3 UK:40.3	Hausarzt („Family practice, general internal medicine“)	12/1994-06/1996
Van den Hoogen et al. 1998 [90]	NL	Kohortenstudie	Praxis	101	>7 Wochen	11	15	k.A.	k.A.	>16	ja	Hausarzt	05/1990-05/1992
Miller et al. 2002 [62]	GB	RCT	Datenbank	211	>6 Wochen < 6Monate ^f	52	k.A.	Md: 39	60.1	k.A.	k.A.	Hausarzt	11/1995-01/1999
Haldorsen et al. 1998 [34]	NO	RCT	Datenbank	81	>8Wochen <6 Monate ^f	k.A.	k.A.	43	57	19-66 ^c	ja	Hausarzt	k.A.
Subakut:													
Van den Hoogen et al 1997 [89]	NL	Kohortenstudie	Praxis	31	6-12 Wochen	11	15	41	55	>16	36 %	Hausarzt	05/1990-05/1992
Hagen et al. 2000/2003 [32, 65]	NO	RCT	ICPC	220	8-12 Wochen ^f	k.A.	k.A.	41.1	47.7	18-60	39.5%	Hausarzt	k.A.
Indahl et al. 1995 [42]	NO	RCT	ICPC	512	8-12 Wochen ^f	k.A.	k.A.	44.3	39.5	19-66 ^c	ja	Hausarzt	08/1991-03/1993
Storheim et al. 2003 [85]	NO	RCT	ICPC, Praxis	29	8-12 Wochen ^f	k.A.	k.A.	38.9	55.2	20-60	nein	Hausarzt	03/1998- 04/2001
Chronisch:													
Poole et al. 2007 [76]	GB	RCT	Praxis	75	>12 Wochen	12	k.A.	47,45	50,7	18-65	k.A.	Hausarzt	k.A.

Subakuter und chronischer Rückenschmerz in der Primärversorgung. Eine systematische Übersichtsarbeit symptomevaluierender Studien - Ergebnisse -

Korff et al. 2005 [48]	GB	RCT	k.A.	121	>12 Wochen	k.A.	k.A.	49,8	60,3	25-64	k.A.	GHC ^g	04/2001-05/2005
Storrø et al. 2004 [86]	NO	RCT	ICPC	72	>16 Wochen <52 Wochen ^c	k.A.	k.A.	44.5 ^a 43.2 ^b	56.5 ^a 61.5 ^b	25-66 ^{ac} 25-63 ^{bc}	36.1%	Hausarzt	k.A.
Van Tulder et al. 1998 [93]	NL	Kohortenstudie	ICPC	368	>12 Wochen	k.A.	26	41,1	49	20-60	66%	Hausarzt	k.A.
Van den Hoogen et al. 1997 [89]	NL	Kohortenstudie	Praxis	72	>12 Wochen	11	15	43	62	>16	53 %	Hausarzt	05/1990-05/1992
van Hoeven et al. 2014 [91]	NL	Valdierungsstudie	ICPC	364	>12 Wochen	k.A.	19	36.3	57	20-45	k.A.	Hausarzt	01/2010-07/2010
Rizza et al. 2012 [78]	CH	Querschnittstudie	ICPC	66.212	>6 Monate ^d	k.A.	63	49.6	52.8	>20 20-111 ^c	ja	Hausarzt	01/2009-07/2011
Frolund et al. 1986 [26]	DK	Querschnittstudie	Praxis	2886	>12 Wochen	k.A.	26	k.A.	k.A.	Alle	k.A.	Hausarzt	04/1986
Rückenschmerz mit Ausstrahlung:													
Nyiendo et al. 2001 [69]	USA	Kohortenstudie	Praxis	310	>6 Wochen	14	111	39.4	52.6	>18	Ja NR:27.3 ÜK:32.3 UK:40.3	Hausarzt („Family practice, general internal medicine“ ^e)	12/1994-06/1996
Storrø et al. 2004 [86]	N	RCT	ICPC	72	>16 Wochen <52 Wochen ^c	k.A.	k.A.	44.5 ^a 43.2 ^b	56.5 ^a 61.5 ^b	25-66 ^{ac} 25-63 ^{bc}	36.1%	Hausarzt	k.A.
<p>a – Patienten mit Schmerzen im unteren Rücken, b – Patienten mit Schmerzen mit unteren Rücken und ausstrahlenden Schmerzen ins Bein, c – Altersgrenzen nicht als Ein- und Ausschlusskriterien, sondern als Beschreibung des Patientenkollektivs, d – Information aus O'Halloran et al. 2004[71], e – oder wiederkehrend min. 6 Wochen in den letzten Wochen + Beschwerden am Tag der Randomisierung, f – Dauer der Krankschreibung, g – GHC: Group Health Cooperative, h – Praxis: Patienten direkt rekrutiert, ICPC: Patienten über ICPC-Codes in Datenbank rekrutiert; Datenbank-Patienten über Datenbank rekrutiert, i – Anzahl der Praxen, j – Anzahl der Hausärzte, k – Durchschnittsalter der Kontrollgruppe wie in Studie beschrieben, l – Altersgrenzen wie in Ein- und Ausschlusskriterien beschrieben ; NR-Nur Rückenschmerz, ÜK:Ausstrahlung bis über das Knie, UK:Ausstrahlung bis unter das Knie</p>													

Tabelle 4: Allgemeine Studieninformationen

4.1.2. Allgemeine Studienmerkmale

Neun Studien [32, 34, 45, 48, 51, 52, 62, 65, 76, 86] waren randomisiert kontrollierte Studien. Die Kontrollgruppe erhielt die übliche hausärztliche Behandlung („usual care“). Die in den eingeschlossenen Studien verwendeten Definitionen von „usual care“ wurden in Tabelle 5 dargestellt, um eine hohe Transparenz und Vergleichbarkeit der Studien zu erzielen.

Des Weiteren erbrachte die Suche drei Kohortenstudien [31, 69, 70, 90, 92], die die Patienten im Verlauf ihrer regulären Behandlung beobachteten, eine Validierungsstudie [91], die darauf abzielte, die Häufigkeit von axialer Spondyloarthritis bei chronischen Rückenschmerzpatienten zu untersuchen und zwei Querschnittstudien [26, 78], die Prävalenzdaten erbrachten.

Studien	Beschreibung der Behandlung „usual care“
Subakut:	
Van den Hoogen et al.[89]	„The study did not interfere with the usual management of the low back pain by the general practitioners involved.“
Hagen et al.[32, 65]	„Patients in the control group were not examined at the spine clinic, but treated with primary health care. They had at least one visit to a general practitioner because is required to obtain sick leave. In this study the kind of treatments or the number of visits the patients made for primary health care was not registered.“
Indahl et al.[42]	The control group was not called in for examination, but were treated within the conventional medical system, which seems to be similar to the practice in other Norwegian counties and probably in other industrialized communities.
Storheim et al. 2003[85]	Patients in the control group were treated by their GP and had no restrictions of treatments or referrals.
Chronisch:	
Poole et al. 2007[76]	„non-intervention““usual care“ (tatsächlich erhalten: verschreibungspflichtige und frei verkäufliche Medikamente(32), Physiotherapie(5), Aromatherapiemassage(2), TENS(1),Akupunktur(2), pflanz. Mittel(3), „pain management program“(1))
Korff et al. 2005[48]	„The Control Group received care as usual, whose content is highly variable across patients. Usual care often included use of prescription and non-prescription pain medications, infrequent primary care visits for back pain, and use of ancillary services such as physical therapy by a minority of patients.“
Storrø et al. 2004[86]	„Patients in the C group were offered “treatment as usual” where the general practitioner refers the patient to the physiotherapist, chiropractor etc., and the patient often initiates a range of paramedical “treatments” to cope with everyday pain problems adjunct to the recommendations given by their doctor.“
Van Tulder et al. 1998[93]	k.A.
Van den Hoogen et al.1997[89]	„The study did not interfere with the usual management of the low back pain by the general practitioners involved.“

van Hoeven et al. 2014[91]	k.A.
Rizza et al. 2012[78]	k.A.
Frølund et al. 1986[26]	k.A.
Subakut/chronisch (gemischt)	
Lamb et al. 2010/2014[45, 51, 52]	„10–15-min session of best practice advice to remain active, provided by a trained health professional and supplemented by The Back Book“
Nyiendo et al. 2001/2005[31, 69, 70]	„the study did not seek to alter or interfere with usual management of LBP by the practitioners involved“ (...) „Medical patients received prescription drugs, exercise plan, and self-care advice;; approximately 25% were referred for physical therapy“
Van den Hoogen et al. 1998[90]	„The study did not interfere with the usual management of the low back pain by the general practitioners involved.“
Miller et al. 2002[62]	„usual care“
Haldorsen et al. 1998[34]	Patients in the control group were followed up by their general practitioners, without any feedback or advice on therapy. These patients, therefore, were subjected to ordinary treatments given by the general practitioners, particularly physiotherapy.
Rückenschmerz mit Ausstrahlung	
Nyiendo et al. 2001 [69]	„the study did not seek to alter or interfere with usual management of LBP by the practitioners involved“ (...) „Medical patients received prescription drugs, exercise plan, and self-care advice;; approximately 25% were referred for physical therapy“
Storrø et al. 2004[86]	„Patients in the C group were offered “treatment as usual” where the general practitioner refers the patient to the physiotherapist, chiropractor etc., and the patient often initiates a range of paramedical “treatments” to cope with everyday pain problems adjunct to the recommendations given by their doctor.“

Tabelle 5: Definition von "usual care" in den eingeschlossenen Studien

Die Studien wurden in Europa und den USA durchgeführt: USA (1 Studie) [31, 69, 70], Großbritannien (4 Studien) [45, 48, 51, 52, 62, 76], Niederlande (3 Studien) [89–91, 93], Norwegen (5 Studien) [32, 34, 42, 65, 85, 86], Niederlande (3 Studien) [89–91, 93], Dänemark (1 Studie) [26], Schweiz (1 Studie) [78]. Alle Studienteilnehmer wurden somit in einem westlichen Gesundheitssystem behandelt. Diese weichen zwar strukturell voneinander ab, ähneln sich jedoch insbesondere im Bereich der Primärversorgung in wesentlichen Aspekten der Behandlung. Patienten konnten auf unterschiedliche Weise rekrutiert werden. Entweder wurden Patienten in der Praxis erfasst, wenn sie den Arzt aufgesucht haben oder sie wurden indirekt über ausgewählte ICPC-Codes im Computersystem identifiziert. Folgende Codes fanden Berücksichtigung in den Studien:

- L02 (Rückensympt./-beschwerd. BWS/n.s.)
- L03 (Untere Rückensympt./-beschwerd.)
- L84 (Rückensyndrom ohne Schmerzausstrahlung)
- L86 (Rückenschmerz mit Schmerzausstrahlung) [7]

Zwei Studien [45, 51, 52, 62]) identifizierten ihre Patienten über die Praxisdatenbanken, ohne das genaue Vorgehen anzugeben.

Alle Studien sind im Bereich der Primärversorgung durchgeführt worden. Dies beinhaltete für die meisten Studien die hausärztliche Versorgung, jedoch wurden auch Studien mit etwas anderen Settings eingeschlossen. Manche Studien präzisieren das genaue Umfeld der Primärversorgung nicht. Eine Studie wurde in einer „Group Health Cooperative“ (GHC) durchgeführt. [48] In einer anderen [45, 51, 52] wurden die Patienten nicht nur vom Hausarzt, sondern auch von einem/einer speziell weitergebildeten Arzthelferin („nurse practitioner“) untersucht. Eine Studie [31, 69, 70] rekrutierte Patienten zudem von einer „family practice“ (Familienmediziner) und einer „general internal medicine“-Praxis. (Allgemeiner Internist) (Tab. 4).

Zudem unterschieden sich die Studien in der Rekrutierung von Patienten mit Schmerzausstrahlung in die Beine. Neun Studien schlossen Patienten mit Ausstrahlung in die Beine ein. [31, 32, 34, 42, 45, 51, 52, 65, 69, 70, 78, 86, 89, 90, 93] Vier der neun Studien machten genaue Prozentangaben zur Höhe des Anteils der Patientin, die ausstrahlende Schmerzen hatten. [31, 32, 65, 69, 70, 86, 89, 93] Fünf Studien verzichteten auf Angaben zu diesem Aspekt. [26, 48, 62, 76, 91] Zwei Studien berichteten Ergebnisse für Rückenschmerzpatienten und Rückenschmerzpatienten mit Ausstrahlung in die Beine getrennt. [69, 86] (Tab. 4). Und eine Studie schloss Patienten mit Ausstrahlung in die Beine aus. [85]

Von den ausgewählten Studien formulierten elf die Altersgrenzen in ihren Ein- und Ausschlusskriterien [31, 32, 45, 51, 52, 65, 70, 85, 89, 90], dabei lag die untere Altersgrenze zwischen 16-25 und die obere Altersgrenze zwischen 45-66. Vier der Studien setzten nach oben hin keine Alterseinschränkung fest. [31, 45, 51, 52, 69, 70, 78, 89, 90] Drei Studien gaben keine Altersgrenzen an, sondern beschrieben lediglich, welche Spannweite das Alter der eingeschlossenen Patienten hatte. [32, 34, 42, 65] Da Storrø et al. [86] nur krankgeschriebene Patienten erfasst hat, waren alle Studienteilnehmer im arbeitsfähigen Alter. Korff et al. [48] machte keine Angabe zu den Altersgrenzen was entweder bedeutet,

dass alle Patienten unabhängig vom Alter eingeschlossen wurden oder dass die Studie es nicht berichtet hat.

Die Spannweite des durchschnittlichen Alters lag zwischen 36.3-54 Jahren, wobei der Mittelwert aller Studien 40.7 Jahre ergab. (Mediane Angaben nicht berücksichtigt) Der Frauenanteil lag zwischen 39.5-61.5 %.

Alle Studien wurden zwischen 1986 und 2014 publiziert, ein Großteil davon in den letzten 15 Jahren. Der Rekrutierungszeitraum der Patienten lag zwischen 1984 und 2011.

Die Details der Ein- und Ausschlusskriterien der jeweiligen Studien unterscheiden sich deutlich. Häufig eingesetzte Zielkriterien sind: Krankschreibung der Patienten, Beschwerdeintensität, Schmerzlokalisierung, „red flags“, Operationen an der Wirbelsäule, psychische Erkrankung/Behandlung und bekannte spezifische Krankheitsätiologien. (Tab. 6)

Ein- und Ausschlusskriterien

Studien	Einschlusskriterien			Ausschlusskriterien					
	Rücken- schmerz ¹	Unterer Rücken- schmerz ²	Krankge- schrieben	Bereits operiert	RF	Psychisch erkrankt/ behandelt	Interventionsbehandlung bereits zuvor erhalten	leichte Be- schwerden	Sonstiges:
Poole et al. [76]		x		x	x	x	x		Bestehende gravierende Vorerkrankungen, Durchblutungsstörungen der Beine, juristische Auseinandersetzung wegen Rückenschmerz bezogener Ansprüche
Lamb et al. [45, 51, 52]		x			x	x	x	x	Nicht in Behandlung wegen Rückenschmerz beim Hausarzt in den letzten 6 Monaten, bereits an einer kognitiven Verhaltenstherapie teilgenommen
Korff et al. [48]	x					x		x	Kürzlich Physiotherapie erhalten, geplanter Austritt aus der „Group Health Cooperative“, für OP angedacht
Storrø et al. [86]		x	x						>52 Wochen Beschwerden
Van Tulder et al.[93]		x			x				
Van den Hoogen et al.[89, 90]		x							Beim Hausarzt wegen Rückenschmerz in den letzten 6 Wochen vor Rekrutierung
Nyiendo et al.[31, 69, 70]		x		x	x				Spezifischer Rückenschmerz, Analphabet, Kontraindikationen für Manipulationen der Wirbelsäule, OP im vergangenen Jahr
Miller et al. [62]		x			x				>6 Monate Symptome, Röntgen in den letzten 12 Monaten, geplante Schwangerschaft
Rizza et al. [78]	x								Infos zu Geschlecht oder Alter fehlten. Alle Besuche, die kein Krankheitssymptom zum Gegenstand hatten (v.a. administrative Besuche)

Subakuter und chronischer Rückenschmerz in der Primärversorgung. Eine systematische Übersichtsarbeit symptomevaluierender Studien - *Ergebnisse* -

Van Hoeven et al.[91]		x						Kontraindikation für MRT, Trauma erlitten
Hagen et al. [32, 65]	x	x	x		x			aktuelle Therapie durch einen anderen Spezialisten
Haldorsen et al.[34]	x	x	x		x			Aktuell nicht berufstätig, schlechte Sprachkenntnisse, blind oder taub, aktive rheumatische Erkrankung, neurologische Erkrankung, ernste internistischen Erkrankungen, Vortest Fragebögen nicht ausgefüllt
Frølund et al.[26]		x						
Indahl et al. [42]			x					Krankschreibung >12 Wochen
Storheim et al.[85]		x	x					Lumbalgien, Spinalstenose mit neurologischen Ausfällen, Spondylolysis/ Spondylolisthesis Grad 2, Wirbelkörper-Fraktur, Tumor, Infektion, Alkohol-, Drogenmissbrauch, rheumatische Erk., Krankheiten, die die Teilnahme behindern, mehr als 3x/wöchentlich Sport/Aufbauübungen
1=Schmerzlokalisierung im Rücken nicht festgelegt; 2= Schmerzlokalisierung auf unteren Rücken beschränkt; Einschlusskriterium „Alter“ (S.Tab.1), Ausschlusskriterium „Schwangerschaft“ hier nicht aufgeführt, da dies ein Ausschlusskriterium dieser syst. Übersichtsarbeit ist								

Tabelle 6: Ein- und Ausschlusskriterien in den eingeschlossenen Studien

4.1.3. Ergebnisse der Qualitätsbeurteilung

4.1.3.1. Allgemeine Studiendetails (Domäne A und B)

Häufiger Grund für eine eingeschränkte Qualitätsbewertung war die mangelhafte Beschreibung des Symptoms. So blieb in vielen Studien unklar, ob Patienten mit „Ausstrahlung in die Beine“ (I1) eingeschlossen wurden oder nicht. [26, 48, 62, 76, 91] Da diese Patienten jedoch meist schlechtere Verläufe zeigen (siehe Kapitel 4.2.3.5), ist dies eine wichtige Information zur Einschätzung des Patientenkollektivs. Ein weiterer limitierender Faktor hinsichtlich der Qualität ist die Repräsentativität der Studienpopulation. So fanden sich vielfach Studien, die nur krank geschriebene Patienten einschlossen und damit nur ein spezielles Patientenkollektiv widerspiegelten [32, 34, 51, 52, 65, 86] oder Studien die leichter betroffene Patienten ausschlossen. [48, 51, 52] Diese beiden Aspekte führten dazu, dass die Einschätzung der klinischen Heterogenität/Varianz vieler Studien [26, 48, 51, 52, 76, 86, 32, 65, 62, 91, 34, 42, 85] „unklar“ ausfiel. (Tab. 9) Einige Studien waren zudem nicht multizentrisch. [26, 86, 89, 90, 91] In der Domäne B (Datenerhebung) gab es zwei häufige Gründe für eine reduzierte Studienqualität. Einige Studien waren Datenbank-basiert (Tab. 4) und erhoben Daten somit nicht direkt über den Kontakt zum Patienten, andere Studien machten die Art der Datenerhebung nicht ausreichend deutlich. [26, 32, 34, 42, 65, 78, 85, 86, 91, 93] Außerdem fanden sich fünf Studien, die die Höhe der akzeptierten Dropout-Rate überschritten. [31, 48, 69, 70, 76, 85, 89, 90] Für weitere Details siehe Tabelle 7.

Schließlich wurde für jede Domäne einer jeden Studie eine Einschätzung bezüglich des Biasrisikos abgegeben (Tab. 8). Hier zeigte sich, dass einige Studien im Bereich der Patienten- und Arztauswahl (Domäne A) ein „unklares“ Biasrisiko hatten. Dabei wurde vor allem berücksichtigt, ob die Patienten konsekutiv rekrutiert wurden und ob die Studie multizentrisch war. Das Biasrisiko im Bereich der Datenerhebung (Domäne B) viel bei etwa 2/3 der Studien „unklar“ aus. Dabei spielte zum einen eine indirekte oder unklare Datenerhebung eine Rolle, die sich meist bei Datenbankstudien fand (Tab. 1) und zum anderen eine Dropout-Rate über 20 %.

Bewertung allgemeiner Studiendetails									
Auswahl Patienten und Ärzte (Domäne A)						Datenerfassung (Domäne B)			
Studie	I1	I2	I3	I4	I5	I6	I7	I8	I9
Signalfragen	Symptom klar beschrieben?	Auswahlkriterien der Patienten klar beschrieben?	Patienten konsekutiv rekrutiert?	Design multi-zentrisch?	Studienpopulation repräsentativ für das volle Spektrum der Rückenschmerzpopulation?	Ärzte repräsentativ für die Primärversorgung?	Direkte Datenerhebung über das Symptom/Einschlusskriterien?	Datenerfassung für alle Patienten gleich?	Höhe der Dropouts unter 20 %?
Frolund et al.[26]	Rot	Rot	Grün	Rot	Gelb	Grün	Gelb	Grün	Grün
Korff et al.[48]	Rot	Grün	Gelb	Gelb	Rot	Grün	Grün	Grün	Rot
Lamb et al.[45, 51, 52]	Grün	Grün	Grün	Grün	Rot	Grün	Grün	Grün	Grün
Nyiendo [31, 69, 70]	Grün	Grün	Grün	Grün	Grün	Grün	Grün	Grün	Rot
Poole et al.[76]	Rot	Grün	Grün	Grün	Gelb	Grün	Grün	Grün	Rot
Storø et al.[86]	Grün	Grün	Grün	Rot	Rot	Grün	Rot	Grün	Grün
V.d. Hoogen [89, 90]	Grün	Grün	Grün	Rot	Grün	Grün	Grün	Grün	Rot
Tulder [93]	Grün	Grün	Grün	Grün	Grün	Grün	Gelb	Grün	Grün
Hagen et al.[32, 65]	Grün	Grün	Grün	Grün	Rot	Grün	Rot	Grün	Grün
Miller et al.[62]	Rot	Grün	Grün	Grün	Gelb	Grün	Grün	Grün	Grün
Van Hoeven [91]	Rot	Grün	Grün	Rot	Rot	Grün	Rot	Grün	Grün
Rizza et al.[78]	Grün	Grün	Grün	Grün	Grün	Grün	Rot	Grün	Grün
Haldorsen et al.[34]	Grün	Grün	Grün	Grün	Rot	Grün	Rot	Grün	Grün
Indahl et al.[42]	Grün	Grün	Grün	Grün	Rot	Grün	Rot	Grün	Grün
Storheim et al.[85]	Grün	Grün	Gelb	Grün	Rot	Grün	Rot	Grün	Rot

Rot: Nein, Gelb: Unklar, Grün: Ja

Tabelle 7: Qualitätsbeurteilung der Studien

Studie	Biasrisiko Patienten- und Arztpraxenauswahl (Domäne A)	Biasrisiko Datenerhebung (Domäne B)	Biasrisiko Ätiologieergebnisse (Domäne C)	Biasrisiko Prognoseergebnisse (Domäne D)
Frølund et al.[26]	Yellow	Green	White	White
Korff et al.[48]	Yellow	Yellow	White	Green
Lamb et al.[45, 51, 52]	Green	Green	White	Green
Nyiendo et al.[31, 69, 70]	Green	Yellow	White	Green
Poole et al.[76]	Green	Yellow	White	Green
Storrø et al.[86]	Yellow	Yellow	White	Green
V.d. Hoogen et al. [89, 90]	Yellow	Yellow	White	Green
Van Tulder et al. [93]	Green	Green	White	Green
Hagen et al.[32, 65]	Green	Yellow	White	Green
Miller et al.[62]	Green	Green	White	Green
Van Hoeven et al. [91]	Yellow	Yellow	Green	White
Rizza et al.[78]	Green	Yellow	White	White
Haldorsen et al. [34]	Green	Yellow	White	Green
Indahl et al.[42]	Green	Yellow	White	Green
Storheim et al. [85]	Yellow	Yellow	White	Green

Grün: Niedrig, Gelb: Unklar, weiß: Studien liefern kein Ätiologie-/Prognoseergebnis

Tabelle 8: Biasrisiko-Beurteilung

Studien	Risiko für klinische Heterogenität/ Varianz	Studien	Risiko für klinische Heterogenität/ Varianz
Frølund et al.[26]	Yellow	Miller et al.[62]	Yellow
Korff et al.[48]	Yellow	Van Hoeven et al.[91]	Yellow
Lamb et al.[45, 51, 52]	Yellow	Rizza et al.[78]	Green
Nyiendo et al.[31, 69, 70]	Green	Haldorsen et al.[34]	Yellow
Poole et al.[76]	Yellow	Indahl et al.[42]	Yellow
Storrø et al.[86]	Yellow	Storheim et al. [85]	Yellow
V.d. Hoogen et al. [89, 90]	Green		White
Van Tulder et al.[93]	Green		White
Hagen et al.[32, 65]	Yellow		White

Grün: niedrig, Gelb: unklar

Tabelle 9: Risiko für klinische Heterogenität/Varianz

4.1.3.2. Ätiologieergebnisse (Domäne C)

Da nur eine Studie [91] identifiziert werden konnte, die sich mit der Ätiologie von Rückenschmerzen bei chronischen Patienten befasst, gibt es im Folgenden auch nur eine Bias Beurteilung für Domäne C (Tab. 8). Diese fiel positiv aus, mit einem niedrigen Risiko für Verzerrungen. Entsprechend des „niedrigen“ Biasrisikos wurden alle Signalfragen im Bereich Ätiologie mit „ja“ beantwortet. (Tab. 10).

Ätiologieergebnisse (Domäne C)			
Studie	I10	I11	I12
Signalfragen	Ätiologische Kategorie klar definiert?	Diagnostische Methoden ausreichend, um die Ätiologie zu bestimmen?	Gleiche Diagnostik für jeden Patienten?
Van Hoeven et al. [91]	Ja	Ja	Ja

Tabelle 10: Qualitätsbeurteilung der Ätiologieergebnisse

4.1.3.3. Prognoseergebnisse (Domäne D)

Domäne D war wesentlich umfangreicher, da die Studien im Bereich Prognose die meisten Ergebnisse erbrachten. Es wurde für jede Studie jedes extrahierte Ergebnis bewertet. Wenn eine Studie beispielsweise sowohl den Roland-Morris Disability Questionnaire als auch die visuelle Analogskala verwendete, wurden beide mit den Bias Tool bewertet. Wichtig für die die Einschätzung des Biasrisikos war vor allem, ob die eingesetzten Messungen geeignet waren um das jeweilige Prognoseergebnis zu bestimmen und ob für jeden Patienten die gleich Art von Messung durchgeführt wurde. In diesen Punkten ließ sich für alle Prognosekategorien aller untersuchten Studien ein „niedriges“ Biasrisiko vermerken. (Tab. 8) Auffällig war andererseits, dass sich keine Studie finden ließ, die dem Qualitätsmerkmal gerecht wurde eine symptomfreie Kontrollgruppe zu untersuchen. (Tab. 11)

Prognoseergebnisse (Domäne D)					
Studie		I13	I14	I15	I16
Signalfragen		Prognoseergebnis klar definiert?	Vergleichsgruppe ohne Symptom?	Messung geeignet, um das jeweilige Prognoseergebnis korrekt zu bestimmen?	Gleiche Messung für jeden Patienten?
Alle Studien	Alle Prognoseergebnisse	Ja	Nein	Ja	Ja

Tabelle 11: Qualitätsbeurteilung der Prognoseergebnisse

4.2. Studienergebnisse

4.2.1. Ergebnisse im Bereich Ätiologie

Es wurde eine Studie identifiziert, die sich mit der Ätiologie von Rückenschmerzen befasst. Von Hoeven et al. [91] berichtete in seiner Studie über den Anteil von axialer Spondyloarthritis bei chronischen Rückenschmerzpatienten. Dabei wurden die gesamten Datenbankinformationen von 12477 Hausarztpatienten in Rotterdam analysiert. Es wurde nach der ICPC-Codierung L03 für unspezifischen Rückenschmerz gesucht. 1106 Patienten hatten die Codierung in ihren Akten erhalten, und waren zu diesem Zeitpunkt 20-45 Jahre alt. Diese wurden daraufhin kontaktiert und überprüft, ob sie die weiteren Ein- und Ausschlusskriterien erfüllten. Wesentlich war vor allem, dass sie zum Zeitpunkt der Befragung Rückenschmerzen von mehr als 12 Wochen hatten.

Um eine möglicherweise zugrunde liegende axiale Spondyloarthritis zu identifizieren, mussten sich die eingeschlossenen 364 Teilnehmer verschiedenen diagnostischen Tests unterziehen. Dazu gehörten: Screeningfragebögen (Calin, Berlin, ASAS), Bluttests (BSG, CRP, HLA-B27 Typisierung), klinische Untersuchung (Brustkorbbeweglichkeit, modifizierter Schobertest, laterale spinale Beweglichkeit, zervikale Rotationsfähigkeit, intermalleolärer Abstand, Tragus-Wand-Abstand), die Anamnese durch einen erfahrenen Rheumatologen (incl. weiterer Fragebögen: je eine visuelle Analogskala für Schmerz, allgemeine Gesundheit und Müdigkeit) und die Anfertigung eines MRTs des Iliosakralgelenks.

Dies führte bei 86 der 364 Teilnehmer zu einer Diagnose von axialer Spondyloarthritis, was einem Prozentsatz von 23.6 % (KI: 19.4 bis 28.3%) entsprach. Von den 86 Patienten hatten 35 % röntgenologische axiale Spondyloarthritis und 65 % nicht-röntgenologische Spondyloarthritis. Wiederum 22 % der Patienten mit diagnostizierter Spondyloarthritis hatten Spondylitis ankylosans (M. Bechterew).

4.2.2. Ergebnisse im Bereich Prävalenz

Zur Beurteilung der Prävalenz von Rückenschmerzen wurden Studien identifiziert, die die Anzahl der subakuten und chronischen Rückenschmerzpatienten in Bezug zu allen Patientenkontakten in einem bestimmten Zeitraum setzten.

Die Literatursuche erbrachte zwei Studien. Die erste Studie (Frølund et al. [26]) berichtete, dass von 26 Hausärzten in einer Arbeitswoche im April 1984 (kumulativ: 128 Arbeitstage) 2886 Patientenkontakte erfasst wurden. Von diesen 2886 kamen 38 Patienten wegen chronischer Rückenschmerzen in die Praxis. Der Rückenschmerz wurde in dieser Studie als muskuloskelettal charakterisiert und betraf den unteren Rücken. Daraus folgte ein prozen-

tualer Anteil von 1.3 % (13/1000 Patientenkontakten). Es zeigte sich ein Verhältnis von 1:1 im Bezug auf akut und chronisch erkrankte Rückenschmerzpatienten. Da die Studie nicht speziell für Rückenschmerzpatienten ausgelegt war, sondern auf alle Schmerzpatienten in der Hausarztpraxis abzielte, ergaben sich für die Subgruppe der chronischen Rückenschmerzpatienten keine Informationen über Durchschnittsalter oder Frauenanteil.

In der zweiten Studie Rizza et al. [78] wurden chronisch erkrankte Patienten über eine ICPC-Datenbanksuche identifiziert. Dafür legten die Studienautoren insgesamt 147 ICPC-2 Codierungen fest, die für chronisch erkrankte Patienten verwendet werden, unter anderem auch für chronisch erkrankte Rückenschmerzpatienten. Es konnte so der Anteil chronischer Rückenschmerzpatienten an der Gesamtheit chronisch Kranker beim Hausarzt ermittelt werden. Im Endergebnis wurden 66.212 Patienten mit insgesamt 335.117 Hausarztbesuchen eingeschlossen. Von diesen 66.212 Chronikern erhielten 3.2 % (KI 2.17-4.23 %) den ICPC-Code L84 (Rückensyndrom ohne Ausstrahlung) und 2.75 % (KI 1.85-3.64 %) den Code L86 (Rückenschmerz mit Ausstrahlung).

Damit haben beim Hausarzt insgesamt 5.95 % der chronisch Kranken chronischen Rückenschmerz.

4.2.3. Ergebnisse im Bereich Prognose

4.2.3.1. Allgemeines

Die Ergebnisse werden im Folgenden in den Kategorien funktionelle Beeinträchtigung, Schmerz, Lebensqualität, Stimmung, Verlauf und arbeitsbezogene Ergebnisse analysiert. Die in den Studien verwendete Vielfalt an Messinstrumenten ist groß (Tab. 12). Dementsprechend musste der Übersicht wegen entschieden werden, welche Prognosedaten zu welchem Zeitpunkt eingehender dargestellt werden sollen. Da der überwiegende Teil der Studien Ergebnisse zu den Zeitpunkten 6 und 12 Monate berichtet, wurde entschieden die beiden Zeitpunkte zur detaillierten Analyse zu nutzen. Zwei Studien [62, 85] berichteten für diese zwei Zeitpunkte jedoch keine Ergebnisse, sodass wir hier alternative Zeitpunkte analysierten (4, 6.5 und 9 Monate). Des Weiteren wurde entschieden, dass die indirekten Prognosemarker „Medikamenteneinnahme/-verschreibung“ und „Anzahl der Arztbesuche“ nicht weiter ausgewertet werden, sodass Studien die diese Ergebnisse berichteten zwar eingeschlossen wurden, die entsprechenden Ergebnisse jedoch im Folgenden nicht berichtet werden. Einigen Studien konnten mehrere Publikationen zugeordnet werden. Bei manchen wichen die berichteten Ergebnisse marginal voneinander ab. In diesem Fall wurden so viele Ergebnisse wie möglich aus derselben Publikation extrahiert.

Übersicht der verwendeten Messinstrumente:					
Messinstrument:	Bestes Ergebnis:	Skala:	Messinstrument:	Bestes Ergebnis:	Skala:
Nottingham Health Profile	0	0-100	(shortended) Fear Avoidance Beliefs Questionnaire	0	17-68
Visuelle Analogskala	0	0-10 0-50 0-100	Worry rating	0	0-10
(revised) Oswestry Disability Questionnaire	0	0-100	Beck depression inventory II	0	0-63
(modified) VonKorff Questionnaire	0	0-100	SF-12/SF-36	100	0-100
(modified) Roland Morris Disability Questionnaire	0	0-24	Pain Self-Efficacy	60	0-60

Tabelle 12: Übersicht der Messinstrumente

4.2.3.2. Subakuter Rückenschmerz

Vier Studien (Hagen et al. [32, 65], Indahl et al. [42], Storheim et al. [85], van den Hoogen [89]) erbrachten Prognoseergebnisse für subakuten Rückenschmerz. Die drei norwegischen Studien [32, 42, 65, 85] nutzen sehr ähnliche Rekrutierungswege und Designs. Die Patienten wurden eingeschlossen, wenn sie 8-12 Wochen wegen Rückenschmerz krankgeschrieben waren. Indahl et al. [42] und Hagen et al. [32, 65] nutzten Daten des „National Insurance Office“. Dort werden zentral die ICPC-codierten Krankschreibungen der Hausärzte registriert. Storheim et al. [85] nutzte neben der Datenbank auch die direkte Rekrutierung über den Hausarzt. Van den Hoogen et al. [89] hingegen rekrutierte nur über die Praxen. Hagen et al. [32, 65] berichtete im Gegensatz zu den anderen, welche ICPC-Codes genau verwendet wurden: L02, L03, L84, L86, L14 (Beinsymptome, -beschwerden).

Die norwegischen Studien [32, 42, 65, 85] schlossen nur Patienten im arbeitsfähigen Alter ein, während van den Hoogen et al. [89] auch die älteren Patienten berücksichtigte. Das Durchschnittsalter der Teilnehmer der vier Studien lag nahe beieinander, 38.9-43.3 Jahre. Die Anzahl der Teilnehmer der jeweiligen Kontrollgruppe war sehr unterschiedlich, 31 vs. 512. Drei der vier Studien [32, 42, 65, 89] schlossen Patienten mit Ausstrahlung in die Beine ein. Während van den Hoogen et al. [89], Storheim et al. [85] und Indahl et al. [42] nur Patienten mit Schmerzen im unteren Rücken einschlossen, schränkte Hagen et al. [32, 65] die Lokalisation im Rücken nicht ein. Die Studienteilnehmer von Storheim et al. [85], hatten im Schnitt vor 8.5 Jahren ihre erste Episode von Rückenschmerz.

Im Bereich des Biasrisikos (Patienten- und Arztauswahl (Domäne A)/Datenerhebung (Domäne B)) ähneln sich die Studien (Tab. 8). Das Risiko wurde zweimal als „unklar/unklar“ [85, 89, 90] bewertet und zweimal als „niedrig/unklar“ [32, 42, 65]. Wesentliche Aspekte für diese Beurteilung waren eine indirekte Datenerhebung bei Datenbankstudien [32, 42, 65, 85], eine fehlende Multizentrität [89,90], eine unklare konsekutive Rekrutierung [85] und eine zu hohe Dropout-Rate [85, 89, 90]. Im Bereich der klinischen Heterogenität/Varianz erhielten alle Studien die Bewertung „unklar“, dies lag an der eingeschränkten Repräsentativität des Patientenkollektivs aufgrund des Ausschlusses von Patienten mit Ausstrahlung in die Beine [85] sowie dem Ausschluss arbeitsfähiger Patienten (Tab. 7, Tab. 9). [32, 42, 65, 85]

4.2.3.2.1. Lebensqualität

Storheim et al. [85] hatte eine Nachbeobachtungsperiode von 18 Wochen (etwa 4 Monate). In dieser Zeit stieg die mittlere Lebensqualität über alle Subskalen gesehen um 9.8 % (PV) (Abb. 2). Die Höhe der Verbesserung und der absoluten Werte am Ende der Nachbeobachtungszeit streuten jedoch breit. Die geringste Lebensqualität ergab sich im Bereich „Rollenfunktion in körperlicher Hinsicht“ (0.26, PEW), eine mittlere Lebensqualität erreichten die Patienten in den Bereichen „Schmerz/“health transition“/Vitalität“ (0.39-0.49, PEW), die beste Lebensqualität erzielten sie in den Bereichen „körperliche Funktion/allgemeine Gesundheitswahrnehmung/soziale Funktionsfähigkeit/Rollenfunktion in emotionaler Hinsicht/psychisches Wohlbefinden“ (0.61-0.74, PEW). Die prozentuale Veränderung lag zwischen -2.9 und 23.6 % (Tab. 13).

Bei van den Hoogen et al. [85, 89] verbesserte sich die Lebensqualität (mediane Ergebnisse) im Verlauf von 6 Monaten im Schnitt um 14.7 % (PV) (Abb. 2). Dabei wurde „daily functioning“ nicht berücksichtigt, da es eine Sonderstellung inne hat und nicht direkt mit den anderen Lebensqualität-Skalen vergleichbar ist. Sechs der sieben Subkategorien des NHP erreichten den medianen Wert von 0. Dies entsprach einer prozentualen Verbesserung zwischen 7-46 %. Zwei der Subskalen (Energieverlust, soziale Isolation [46]) waren von vornherein bei einem medianen Wert von 0 und blieben auch nach 6 Monaten auf diesem Niveau. Lediglich die Skala für Müdigkeit [46] zeigte sich auch nach 6 Monaten erhöht bei einem medianen Wert von 24 mit einer prozentualen Verbesserung von 13 % (Tab. 13).

Ein Vergleich beider Studien ist nicht möglich, da Storheim et al. [85] Mittelwerte angibt und van den Hoogen et al. [85, 89] im Median berichtet.

Dennoch zeigt Abbildung 2, dass sich die Lebensqualität in beiden Studien ähnlich positiv entwickelt.

Es lässt sich festhalten, dass sich die Lebensqualität arbeitsfähiger Patienten im Verlauf von 6 Monaten durchschnittlich um 14.7 % verbessert. Dabei erreichen nahezu alle Subskalen beste mediane Ergebnisse.

Arbeitsunfähige Patienten mit subakutem Rückenschmerz erreichen im Schnitt eine Verbesserung von 9.8 % (PV) über vier Monate gesehen. Die Ergebnisse der verschiedenen Subskalen fallen jedoch sehr unterschiedlich aus. (Tab. 13)

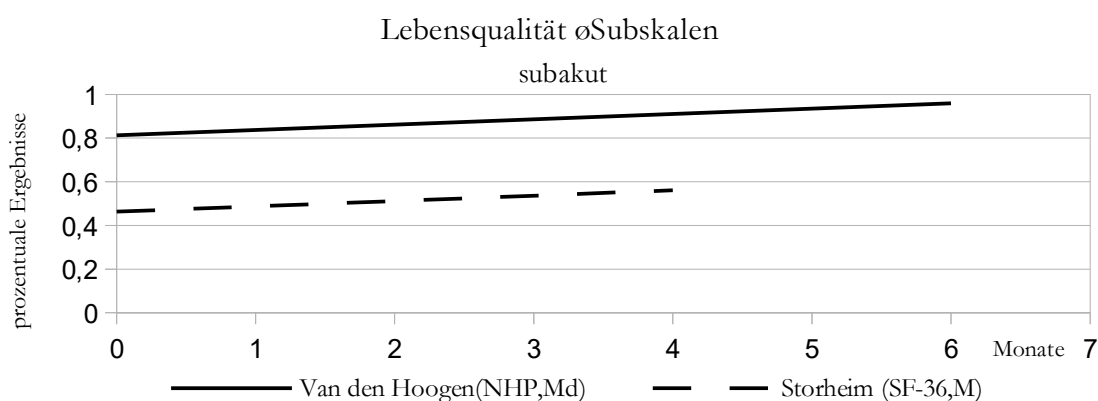


Abbildung 2: Lebensqualität - Durchschnittswert der Subskalen (subakuter Rückenschmerz)

Die Polarität der Messinstrumente wurde in dieser Abbildung angeglichen
 1-bestes Ergebnis 0-schlechtestes Ergebnis
 van den Hoogen: „daily functioning“ nicht mit einbezogen, da es eine Sonderstellung hat

Studienergebnisse im Bereich Lebensqualität							
Studie	Pat. N	Messinstrument	Base-line	PBW	6 Monate ^a	PEW	PV, %
Van den Hoogen et al. [89]	31	NHP (0-100) – Schmerz ^b	Md:46	0.46	Md:0	0	46
Van den Hoogen	31	NHP (0-100) – Mobilität	Md:22	0.22	Md:0	0	22
Van den Hoogen	31	NHP (0-100) – Schlaf	Md:0	0	Md:0	0	0
Van den Hoogen	31	NHP (0-100) – Energieverlust	Md:37	0.37	Md:24	0.24	13
Van den Hoogen	31	NHP (0-100) – Emotionale Reaktion	Md:7	0.07	Md:0	0	7
Van den Hoogen	31	NHP (0-100) – soziale Isolation	Md:0	0	Md:0	0	0
Van den Hoogen	31	NHP (0-7) – „Daily functioning“	Md:3	0.429	Md:0	0	42.9
Storheim et al.[85]	29	SF-36 (0-100) – körperliche Funktion ^c	M:60.9	0.609	M:66.9	0.669	6.0

Storheim	29	SF-36 (0-100) – Rollenfunktion in körperlicher Hinsicht	M:7.8	0.078	M:25.9	0.259	18.1
Storheim	29	SF-36 (0-100) – Schmerz	M:25.8	0.258	M:38.4	0.384	12.6
Storheim	29	SF-36 (0-100) – allgemeine Gesundheitswahrnehmung	M:63.8	0.638	M:60.9	0.609	-2.9
Storheim	29	SF-36 (0-100) – Vitalität	M:40.3	0.403	M:44.2	0.442	3.9
Storheim	29	SF-36 (0-100) – soziale Funktionsfähigkeit	M:63.8	0.638	M:73.3	0.733	9.5
Storheim	29	SF-36 (0-100) – Rollenfunktion in emotionaler Hinsicht	M:62.1	0.621	M:73.6	0.736	11.5
Storheim	29	SF-36 (0-100) – psychisches Wohlbefinden	M:67.7	0.677	M:73.3	0.733	5.6
Storheim	29	SF-36 (0-100) – „Health transition“	M:24.9	0.249	M:48.5	0.485	23.6
a- Der Nachbeobachtungszeitpunkt bei Storheim war 4 Monate (16 Wochen) b-deutsche Übersetzungen aus Studie [46] c- deutsche Übersetzung aus Studie [66]							

Tabelle 13: Lebensqualität (subakuter Rückenschmerz)

4.2.3.2.2. Schmerz

Das mediane Ergebnis einer Studie (van den Hoogen et al. [89]) lag nach 6 Monaten bei 0.1 (PEW, VAS). Dies entsprach einer prozentualen Verbesserung von 32 %. Storheim et al. [85] verblieb im Endergebnis (Mittelwerte) auf einem höheren Schmerzniveau von 0.48 (PEW, VAS) entsprechend einer prozentualen Veränderung von 10 % nach etwa 4 Monaten.

Damit liegt die prozentuale Veränderung für arbeitsunfähige Patienten mit subakutem Rückenschmerz sowohl im Bereich Lebensqualität als auch Schmerz bei 10 %. (Tab. 14)

4.2.3.2.3. Funktionelle Beeinträchtigung

Storheim et al. [85] berichtete einen Baseline-Wert von 9.3 Punkten im RDQ, der sich um 1.6 Punkte nach 4 Monaten verbesserte. Die entsprach einer prozentualen Veränderung von 6.6 %. (Tab. 14)

Studie	Pat. N	Messinstrument	Baseline	PBW	6 Monate ^a	PEW	PV, %
Schmerz							
Van den Hoogen et al.[89]	31	VAS (0-50)	Md:21	0.42	Md:5	0.1	32
Storheim et al.[85]	29	VAS (0-100)	M:58.3	0.583	M:48	0.48	10
Funktionelle Beeinträchtigung							
Storheim et al.	29	RDQ (0-24)	M:9.3	0.388	M:7.7	0.321	6.6
a-Nachbeobachtungszeitpunkt bei Storheim waren 4 Monate (18 Wochen)							

Tabelle 14: Schmerz und funktionelle Beeinträchtigung (subakuter Rückenschmerz)

4.2.3.2.4. Arbeitsbezogene Ergebnisse

Hagen et al. [32, 65] berichtete, dass von 220 krankgeschriebenen Studienteilnehmern nach 6 Monaten 45 % (99 Teilnehmer) wieder Vollzeit arbeiten konnten. Nach 12 Monaten erhöhte sich der prozentuale Anteil um weitere 11.4 % auf 56.4 %. Dies bedeutet im Umkehrschluss allerdings auch, dass mehr als 1/3 der Teilnehmer auch nach einem Jahr noch nicht wieder arbeitsfähig war. Dies entsprach einem Mittelwert von etwa 134 Tagen, die die Teilnehmer krankgeschrieben waren. Indahl et al. [42] berichtete einen ähnlichen Trend. Nach 200 Tagen (etwa 6 1/2 Monaten) waren noch 60 % der Patienten krankgeschrieben. Bei Storheim et al. [85] wurden an 56.3 Tagen innerhalb der 18 Wochen Krankengelder („benefitted days“) ausgezahlt (Tab. 15).

Arbeitsbezogenen Ergebnisse					
Studie	n	Ergebnis	Baseline	6 Monate	12 Monate
Hagen et al.[32, 65]	220	Returning to Full-Duty Work	k.A	45%	56.4%
Hagen et al.	220	Days of sickness compensation	k.A	k.A	M:133.7 KI:118.9 bis 148.5
Hagen et al.	220	Disability pension	k.A	k.A	6.4%
Indahl et al. [42] ^a	512	On sick leave	100%	60%	k.A
Storheim et al. [85] ^b	29	Benefitted Days		56.3	

a- Der Nachbeobachtungszeitpunkt war 200 Tage b- Der Nachbeobachtungszeitpunkt war 18 Wochen

Table 15: Arbeitsbezogene Ergebnisse (subakuter Rückenschmerz)

4.2.3.2.5. Abhängigkeit des Verlaufs vom Ausgangswert

Es wurde ein Diagramm erstellt, das den prozentualen Baseline-Wert in einen Zusammenhang mit der prozentualen Veränderung zum letzten Nachbeobachtungszeitpunkt setzt (4 bzw. 6 Monate, siehe oben). Dies sollte untersuchen, ob Patienten mit ausgeprägter Beeinträchtigung, niedriger Lebensqualität und hoher Schmerzintensität zu Beginn der Untersuchung eine ausgeprägtere relative Verbesserung erzielen als Patienten, die zu Beginn weniger eingeschränkt waren. Abbildung 3 zeigt, dass Patienten mit guter Lebensqualität zum Baseline-Zeitpunkt eine stärkere Verbesserung der Lebensqualität im Verlauf von 6 Monaten aufweisen als Patienten mit niedrigeren Ausgangswerten.

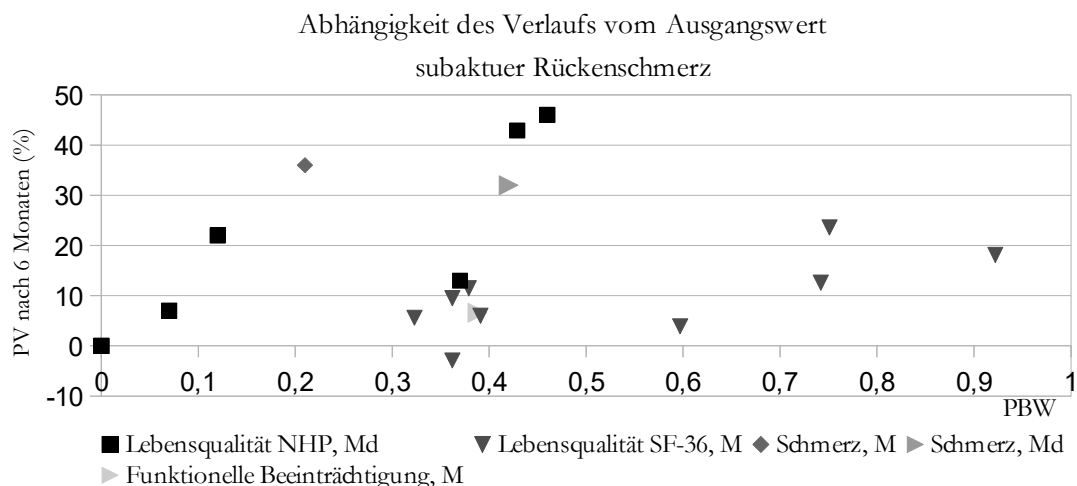


Abbildung 3: Abhängigkeit des Verlaufs vom Ausgangswert (Subakuter Rückenschmerz)

Die Polarität der Messinstrumente wurde in dieser Abbildung angeglichen. Ergebnisse zum Zeitpunkt 4 und 6 Monate.

4.2.3.2.6. Fazit

Arbeitsunfähige Patienten mit subakutem Rückenschmerz können innerhalb von 4 Monaten eine Verbesserung von 10 % in den Bereichen Lebensqualität und Schmerz erwarten sowie eine Verbesserung um 6.6 % im Bereich funktionelle Beeinträchtigung. [85] Außerdem zeigt sich, dass nach 6 Monaten noch zwischen 55-60 % krankgeschrieben sind, nach 12 Monaten noch etwa 43 %. [32, 42, 65, 85].

Die Ergebnisse für ein gemischtes Patientenkollektiv (bezüglich der Arbeitsfähigkeit) zeigt sehr gute mediane Ergebnisse, die in den Bereichen Lebensqualität und Schmerz sehr nahe am bestmöglichen Ergebnis liegen. Die prozentuale Veränderung liegt hier bei 14.7 % und 32 % nach 6 Monaten. [89]

4.2.3.3. Chronischer Rückenschmerz

Fünf Studien ([Korff et al. [48], Poole et al. [76], Storrø et al. [86], Van den Hoogen et al. [89], Van Tulder et al. [93]) lieferten Daten für chronisch erkrankte Rückenschmerzpatienten. Das Biasrisiko (Patienten- und Arztauswahl (Domäne A)/Datenerhebung (Domäne B)) war dreimal „unklar/unklar“ [48, 76, 86], einmal „niedrig/niedrig“ [93] und einmal „niedrig/unklar“ [89] (Tab. 8). Dabei zeigten sich Qualitätseinschränkungen in den Bereichen Dropout-Rate [48, 76, 90] und Multizentrität [48, 76, 86, 90]. Zudem blieb einmal unklar, ob die Patienten konsekutiv rekrutiert wurden [48]. Im Bereich der klinischen Heterogenität/Varianz wurden die Studien dreimal als „unklar“ bewertet [48, 76, 86] und zweimal als „niedrig“ [89, 93] (Tab. 9). Grund für diese Bewertung war vor allem der Aspekt „Ausstrahlung in die Beine“. Die Studien machten nicht deutlich, ob diese Patienten eingeschlossen wurden. [48, 76] Außerdem schloss eine Studie Patienten mit leichten Beschwerden aus [48] und eine andere untersuchte nur Patienten mit Krankschreibung. [86]

Wie man bereits in der Qualitätsbeurteilung sehen konnte sind die Studien abgesehen von ihren grundlegenden Gemeinsamkeiten, die sich durch die Ein- und Ausschlusskriterien dieser Übersichtsarbeit ergeben, nur eingeschränkt miteinander vergleichbar. Im Detail fanden sich Unterschiede vor allem im Bezug auf die Art der Rekrutierung, bei der Auswahl der Patienten sowie bei dem Aspekt „Schmerzausstrahlung in die Beine“. Zwei Studien rekrutierten ihre Teilnehmer direkt (Poole et al., van den Hoogen et al. [76, 89]), zwei Studien indirekt über ICPC-Codes (van Tulder et al., Storrø et al. [86, 93]). Korff et al. [48] beschrieb den Rekrutierungsprozess nicht im Detail zudem wurden nur Patienten eingeschlossen, deren funktionelle Beeinträchtigung im RMQ >7 war. Innerhalb dieser stärker beeinträchtigten Gruppe zeigte sich dennoch eine recht gleichmäßige Verteilung im Bezug auf die Schmerzintensität und die funktionelle Beeinträchtigung der Studienteilnehmer. Messinstrument für diese Einschätzung war der „Chronic pain grade“ (CPG), dieser zeigte folgende Aufteilung der Kontrollgruppe: CPG I: 24 %, CPG II: 22.3 %, CPG III: 24.8 %, CPG IV: 28.9 %. Bei Storrø et al. [86] galt es zu beachten, dass nur Patienten mit einer Krankschreibung zwischen 16 und 52 Wochen eingeschlossen wurden und somit für die Gesamtheit der Rückenschmerzpatienten nur sehr eingeschränkt repräsentativ sind.

Zwei Studien (Poole et al., Tulder et al. [76, 93]) gaben zudem an, wie lange die Beschwerden zum Zeitpunkt der Rekrutierung bestanden. Im Mittel waren es 114.7 Monate (SD:106.7) bei Poole et al. [76] und bei van Tulder et al. [93] im Median 53.5 Monate (Spannweite: 4-492 Monate). Das Fehlen dieser Informationen bei den anderen Studien er-

schwerte die Einschätzung des Patientenkollektivs und deren Vergleichbarkeit. Über die Schmerzausstrahlung gaben drei der fünf Studien Auskunft (Storrø et al., van Tulder et al., van den Hoogen et al. [86, 89, 93]). Der Anteil der Patienten mit Schmerzausstrahlung lag zwischen 36.1 %-66 %. Nur van Tulder et al. [93] machte darüber hinaus die Angabe wie hoch der Anteil der Patienten war, die wegen ihrer ersten Episode von Rückenschmerzen in Behandlung waren. Dieser lag bei 59 %. Die Anzahl der Teilnehmer der fünf Studien lag zwischen 72-368. Die untere Altersgrenze lag zwischen 16-25, die obere Altersgrenze zwischen 60-66 Jahren, während sie bei van den Hoogen et al. [89] nach oben offen war. Das Durchschnittsalter der Teilnehmer lag zwischen 41.1-49.8 Jahren.

4.2.3.3.1. Funktionelle Beeinträchtigung

Der prozentuale Baseline-Wert der Teilnehmer lag zwischen 0.396-0.496 und damit auf einem mittleren Beeinträchtigungsniveau. Damit verbesserte sich die funktionelle Beeinträchtigung zu den letzten Nachbeobachtungszeitpunkten um 5.7-6.8 % (6 Monate) bzw. um 5.8-10 % (12 Monate).

Die Ausgangswerte und die Entwicklung über die Zeit sind somit bei allen Studien recht ähnlich. Sie sprechen für eine leichte Verbesserung der funktionellen Beeinträchtigung über 6-12 Monate. [48, 76, 93] (Abb. 4)

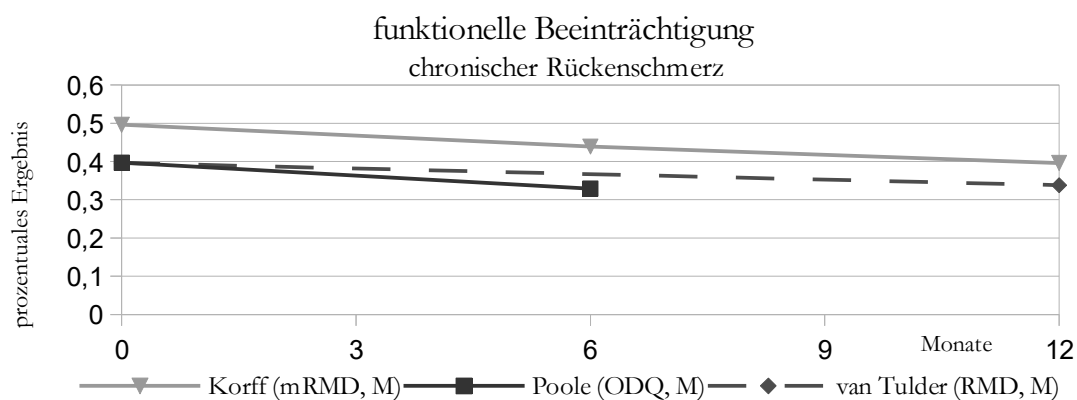


Abbildung 4: Funktionelle Beeinträchtigung (chronischer Rückenschmerz)

1- schlechtestes Ergebnis 0- bestes Ergebnis

4.2.3.3.2. Lebensqualität

Zwei der Studien (Korff et al., Poole et al. [48, 76]) verwendeten den SF-36 als Messinstrument, die anderen zwei (van den Hoogen et al., van Tulder et al. [89, 93]) nutzten das Nottingham Health Profile.

Während Korff et al. [48] nur von zwei Subskalen (Soziale Funktionsfähigkeit und psychisches Wohlbefinden) des SF-36 berichtete, präsentierte Poole et al. [76] alle 8 Subskalen des SF-36. Um beide Ergebnisse besser miteinander vergleichen zu können, wurden die beiden oben erwähnten Subskalen von Korff et al. [48] sowie die beiden entsprechenden Subskalen bei Poole et al. gemittelt. [76] (Abb. 5, vgl. Korff und „Poole teil“, Tab. 16)

Korff, SF-36 (soziale Funktionsfähigkeit und psychisches Wohlbefinden)	PBW: 0.697	6 Monate (PEW): 0.716	12 Monate (PEW) 0.728
„Poole teil“, SF-36 (soziale Funktionsfähigkeit und psychisches Wohlbefinden)	PBW: 0.597	6 Monate (PEW): 0.646	

Tabelle 16: Gemittelte Teilergebnisse, Lebensqualität Subskalen (chronischer Rückenschmerz)

Die Betrachtung der gemittelten Ergebnisse der zwei Subskalen zeigte, dass die Teilnehmer bei Korff et al. [48] auf einem höheren Ausgangswert starteten. Die Teilnehmer von Poole et al. [76] erreichten stattdessen nach 6 Monaten eine höhere prozentuale Verbesserung. Im Endergebnis verblieben die Patienten von Korff et al. [48] allerdings auf einem leicht höheren Niveau der Lebensqualität als die Teilnehmer von Poole et al. [76]. Vergleicht man dieses Ergebnis mit dem gemittelten Gesamtergebnis aller 8 Subskalen von Poole et al. [76] fiel die Lebensqualität über alle Subskalen gesehen jedoch deutlich niedriger aus. (Abb. 5, Tab. 17) Dies liegt unter anderem daran, dass die Subskalen „soziale Funktionsfähigkeit“ und „psychisches Wohlbefinden“ zu den denen zählen, die gut ausfielen. (Tab. 18)

Poole gesamt	Lebensqualität	Baseline: 0.479 (PBW)	6 Monate: 0.53 (PEW)
--------------	----------------	-----------------------	----------------------

Tabelle 17: Gemitteltes Gesamtergebnis, Subskalen Lebensqualität (chronischer Rückenschmerz)

Das das Gesamtergebnis deutlich weniger gut ausfällt liegt vor allem an den Bereichen „Rollenfunktion in körperlicher Hinsicht“, „Schmerz“ und „Vitalität“. (Tab.18) Eine entsprechende Beobachtung wurde bereits bei der Lebensqualität subakuten Patienten gemacht. (Siehe Kapitel 4.2.3.2.1)

Zusammengefasst ist die Tendenz der Lebensqualität allerdings in beiden Studien positiv [48, 76]. Die gemessenen prozentualen Verbesserungen über die Zeit sind jedoch nicht allzu hoch und die Teilnehmer verbleiben nach 6-12 Monaten auf einem Niveau zwischen 0.378 und 0.744 (PEW).

Van den Hoogen et al. [89] und van Tulder et al. [93] berichteten über die Lebensqualität mit dem Messinstrument NHP. Die Daten von van Tulder et al. [93] werden an dieser Stelle nur teilweise ausgewertet, da 4 der 6 Subskalen zum Baseline-Zeitpunkt als Mittelwert angegeben wurden, zum Nachbeobachtungszeitpunkt jedoch als Median, da sie aus unterschiedlichen Publikationen stammen.

Die verbleibenden gemeinsamen medianen Subskalen „Schmerz“ und „Mobilität“ zeigten beide eine positive Tendenz. Die prozentuale Veränderung war bei van den Hoogen et al. [89] jedoch deutlicher ausgeprägt, obwohl der Endpunkt mit 6 Monaten früher lag als bei van Tulder et al. [93]. (Abb. 5, siehe: „van den Hoogen teil“ vs. „van Tulder“)

Betrachtet man alle Subskalen van den Hoogen et al. [89] in ihrer Gesamtheit, ist das auffälligste Ergebnis, dass 5 der 7 Subskalen zum Endpunkt 6 Monate den medianen Wert von 0 annahmen. Dies betraf die Bereiche Schlaf, Energieverlust, emotionale Probleme, soziale Isolation und „social functioning“. Aber auch in den verbleibenden zwei Kategorien „Schmerz“ und „Mobilität“ waren die erzielten Verbesserungen ausgeprägt. Besonders die prozentuale Verbesserung im Bereich Schmerz stach mit 43 % heraus (Tab. 18).

Es fällt auf, dass in beiden Studien [89, 93] zum jeweiligen Endpunkt nur noch die Subskalen „Schmerz“ und „Mobilität“ erhöht waren, alle anderen nehmen den medianen Wert 0 an.

Sowohl Studien mit subakuten als auch Studien mit chronischen Rückenschmerzpatienten machen deutlich, dass bestimmte Bereiche der Lebensqualität durch die Beschwerden besonders betroffen sind, während andere Bereiche weniger eingeschränkt sind. Dies zeigte sich unabhängig vom verwendeten Messinstrument. Bei subakuten [85] und chronischen [76] Rückenschmerzpatienten sind psychisches Wohlbefinden, Rollenfunktion in emotionaler Hinsicht und soziale Funktionsfähigkeit nach 4 bzw. 6 Monaten gut, während die Subskalen Schmerz, Vitalität und Rollenfunktion in körperlicher Hinsicht schlecht ausfallen.

Ebenso ergibt sich beim Vergleich innerhalb der Gruppe der chronischen Rückenschmerzpatienten [89, 93], dass nach 6 bzw. 12 Monaten die medianen Ergebnisse der Subskalen

Schmerz und Mobilität weiter erhöht sind, während alle anderen Subskalen (Schlaf, Energieverlust, emotionale Reaktion, soziale Isolation) beim medianen Wert 0 liegen.

In diesem Detail kristallisiert sich ein Unterschied zu dem Ergebnis der subakuten Patienten heraus, bei diesen bleibt nur die Subskala Energieverlust nach 6 Monaten noch erhöht, während alle anderen Subskalen den medianen Wert 0 annehmen.

Die beiden Studien [76, 89], die alle Subskalen des jeweiligen Messinstrumentes analysiert haben, zeigen, dass die zu erwartende Verbesserung der Lebensqualität nach 6 Monaten über alle Subskalen gemittelt zwischen 5,1 % (SF-36, Mittelwerte) [76] und 16 % (NHP, mediane Werte) [89] liegt. Der weitere Vergleich von Studienergebnissen und Studiencharakteristika sowie Qualitätsbeurteilung führte zu keinen weiteren Erkenntnissen in diesem Bereich.

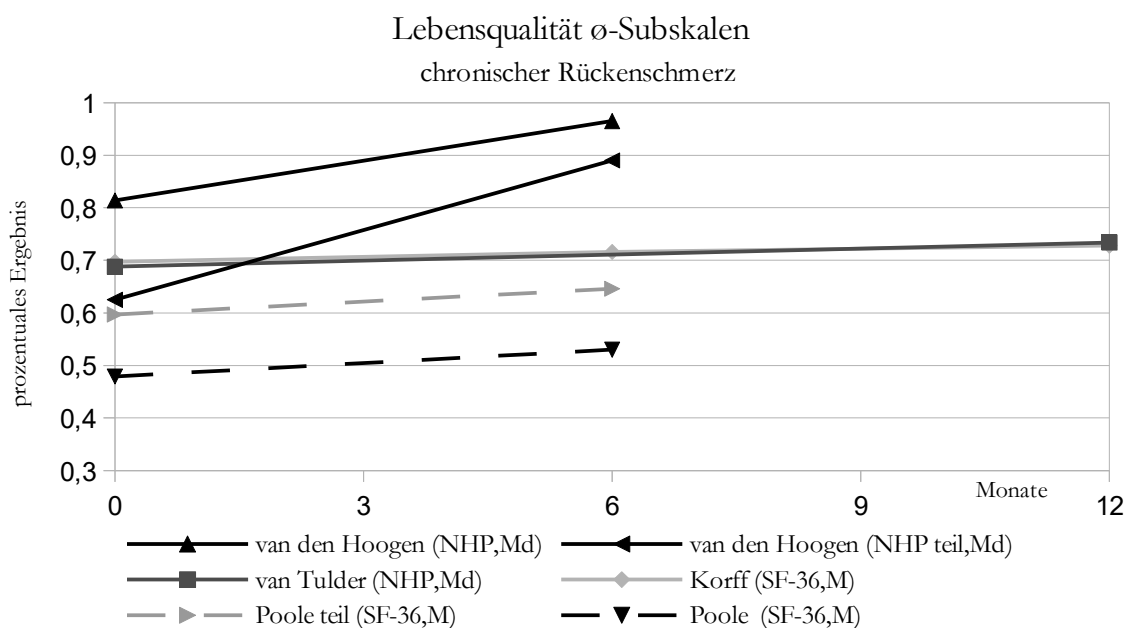


Abbildung 5: Lebensqualität (chronischer Rückenschmerz) gemittelte Werte der Subskalen

Die Polarität der Messinstrumente wurde angeglichen, 1- bestes Ergebnis 0- schlechtestes Ergebnis
Beim NHP wurde die Subskala "Daily functioning" nicht einberechnet, da sie eine Sonderform darstellt

Ergebnisse im Bereich Lebensqualität										
Studie	Pat. N	Messinstrument	Baseline	PBW	6 Monate	PEW	PV, %	12 Monate	PEW	PV, %
Korff et al. [48]	121	SF-36 (0-100), soziale Funktionsfähigkeit ^a	M:70.4 s:27.0	0.704	M:73.6 s:27.8	0.736	3.2	M:74.4 s:24.0	0.744	4.0
Korff et al.	121	SF-36 (0-100), psychisches Wohlbefinden	M:68.9 s:16.9	0.689	M:69.5 s:19.1	0.695	0.6	M:71.1 s:18.4	0.711	2.2
Poole et al. [76]	75	SF-36 (0-100), körperliche Funktionen	M:45.4 s:27.7	0.454	M:52.2 s:29.5	0.522	6.8	k.A.	k.A.	k.A.
Poole et al.	75	SF-36 (0-100), soziale Funktionsfähigkeit	M:58.2 s:29.7	0.582	M:61.5 s:30.8	0.615	3.3	k.A.	k.A.	k.A.
Poole et al.	75	SF-36 (0-100), Rollenfunktion in körperlicher Hinsicht	M:29.1 s:39.8	0.291	M:37.8 s:42.5	0.378	8.7	k.A.	k.A.	k.A.
Poole et al.	75	SF-36 (0-100), Rollenfunktion in emotionaler Hinsicht	M:58.1 s:46.0	0.581	M:62.0 s:44.0	0.62	3.9	k.A.	k.A.	k.A.
Poole et al.	75	SF-36 (0-100), Schmerz	M:37.5 s: 20.3	0.375	M:44.4 s:28.5	0.444	6.9	k.A.	k.A.	k.A.
Poole et al.	75	SF-36 (0-100), psychisches Wohlbefinden	M:60.2 s:18.4	0.602	M:67.7 s:18.5	0.677	7.5	k.A.	k.A.	k.A.
Poole et al.	75	SF-36 (0-100), Vitalität	M:39.7 s:23.2	0.397	M:43.4 s:21.8	0.434	3.7	k.A.	k.A.	k.A.
Poole et al.	75	SF-36 (0-100), allgemeine Gesundheitswahrnehmung	M:55.0 s:23.1	0.55	M:55.0 s:24.1	0.55	0	k.A.	k.A.	k.A.
Van den Hoogen et al. [89]	72	NHP (0-100), Schmerz ^b	Md:53.0	0.53	Md:10.0	0.1	43	k.A.	k.A.	k.A.
Van den Hoogen et al.	72	NHP (0-100), Mobilität	Md:22	0.22	Md:11	0.11	11	k.A.	k.A.	k.A.
Van den Hoogen et al.	72	NHP (0-100), Schlaf	Md:6	0.06	Md:0	0	6	k.A.	k.A.	k.A.
Van den Hoogen et al.	72	NHP (0-100), Energieverlust	Md:24	0.24	Md:0	0	24	k.A.	k.A.	k.A.
Van den Hoogen et al.	72	NHP (0-100), Emotionale Reaktion	Md:7	0.07	Md:0	0	7	k.A.	k.A.	k.A.
Van den Hoogen et al.	72	NHP (0-100), soziale Isolation	Md:0	0.0	Md:0	0	0	k.A.	k.A.	k.A.
Van den Hoogen et al.	72	NHP (0-7), Daily functioning	Md:3	0.429	Md:0	0	37.5	k.A.	k.A.	k.A.
Van Tulder et al. [92, 93]	368	NHP (0-100), Schmerz	Md:40.5 IQR:10.5 bis 69.8	0.405	k.A.	k.A.	k.A.	Md:31.4 IQR:0 bis 64.4	0.314	9.1
Van Tulder et al.	368	NHP (0-100), Mobilität	Md:21.8 IQR:11.2 bis 34.4	0.218	k.A.	k.A.	k.A.	Md:21.8 IQR:11.2 bis 34.4	0.218	0
Van Tulder et al.	368	NHP (0-100), Schlaf	M:14.90 s:24.10	0.149	k.A.	k.A.	k.A.	Md:0	0	k.A.

Ergebnisse im Bereich Lebensqualität										
Studie	Pat. N	Messinstrument	Baseline	PBW	6 Monate	PEW	PV, %	12 Monate	PEW	PV, %
Van Tulder et al.	368	NHP (0-100), Energieverlust	M:30.80 s:35.30	0.308	k.A.	k.A.	k.A.	Md:0	0	k.A.
Van Tulder et al.	368	NHP (0-100), Emotionale Reaktion	M:12.20 s:18.70	0.122	k.A.	k.A.	k.A.	Md:0	0	k.A.
Van Tulder et al.	368	NHP (0-100), soziale Isolation	M:7.40 s:18.40	0.074	k.A.	k.A.	k.A.	Md:0	0	k.A.
a-deutsche Übersetzung aus Studie [46] b- deutsche Übersetzung aus Studie[66]										

Tabelle 18: Lebensqualität (chronischer Rückenschmerz)

4.2.3.3.3. Stimmung

In die Kategorie Stimmung wurden die Messinstrumente „shortened fear-avoidance-beliefs“, „worry rating“ und „beck depression inventory II“ eingeordnet. Alle diese Messinstrumente befassen sich letztlich mit ganz unterschiedlichen Aspekten, sodass ein Vergleich der Ergebnisse nicht möglich ist. Festhalten kann man jedoch auch hier, dass in allen Messinstrumenten eine leichte Verbesserung über die Zeit messbar ist. Diese liegt zwischen 2.5-14 % (PV) nach 6 Monaten und zwischen 5.74-17 % (PV) nach 12 Monaten. Es gibt somit auch im Bezug auf psychische Aspekte der Patienten eine leicht positive Tendenz. (Tab. 19)

Studie	Pat. N	Messinstrument	Baseline	Pat. N	PBW	6 Monate	PEW	PV, %	12 Monate	PEW	PV, %
Stimmung											
Korff et al.[48]	121	Shortened Fear-avoidance-beliefs (17-68)	M:41.30 s:8.20	121	0.607	M:39.2 s:9.7	0.576	3.09	M:37.4 s:9.50	0.55	5.74
Korff et al.	121	Worry rating (0-10)	M:6.2 s:2.7	121	0.62	M:4.8 s:2.7	0.48	14	M:4.5 s:2.7	0.45	17
Poole et al.[76]	75	Beck depression inventory II, (0-63)	M:14.4 s:9.8	45	0.229	M:12.8 s:9.2	0.203	2.54	k.A.	k.A.	k.A.

Tabelle 19: Stimmung (chronischer Rückenschmerz)

4.2.3.3.4. Schmerz

Bei Korff et al. [48] zeigte sich eine Schmerzreduktion um 11 % (PV) nach 6 Monaten, die sich nach 12 Monaten nicht weiter verbesserte. Das Endergebnis lag bei 0.47 (PEW), sodass die Schmerzen auf einem mittleren Schmerzniveau stagnierten. In einem ähnlichen Bereich befanden sich auch die Ergebnis von van Tulder et al. [93].

Hier lag die Schmerzreduktion bei etwa 8 % (PV) mit einem absoluten prozentualen Wert von 0.48 (12 Monate, PEW).

Sie war somit etwas geringer, wohin gegen der absolute Wert fast identisch mit der Studie von Korff et al. war. [48]

Auch die medianen Ergebnisse von van den Hoogen et al. [89] zeigten eine Verbesserung der Schmerzintensität. So reduzierten sich die Schmerzen um 30 % (PV) und verblieben bei einem medianen Ergebnis von 0.1 (6 Monate, PEW).

Poole et al. [76] fiel hier mit einer Verschlechterung um 2.1 % (PV) aus dem Rahmen. Der Abgleich mit der Qualitätsbeurteilung und den Studienspezifika ließ für diese Abweichung keinen Grund erkennen. Mit dem absoluten Ergebniswert von 0.43 (PEW) lag die Studie jedoch letztlich in einem ähnlichen Bereich wie Korff et al. [48] und van Tulder et al. [93].

Zusammenfassend zeigt sich, dass sich die durchschnittliche Schmerzintensität im Verlauf von 6 Monaten im Schnitt um 4,5 % (-2,1 bis 11 %) und um 9,5 % (8 bis 11 %) nach 12 Monaten verbessert. Die absoluten Werte liegen in einem ähnlichen Bereich um die 0.45 (6 Monate) bzw. 0.48 (12 Monate). (Abb. 6)

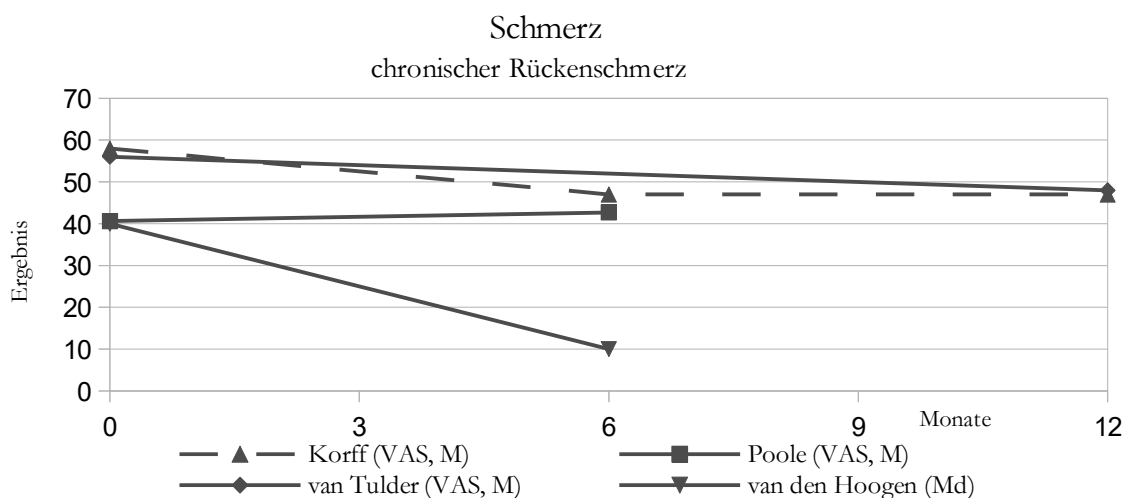


Abbildung 6: Schmerz (chronischer Rückenschmerz)

100- schlechtestes Ergebnis, 0- bestes Ergebnis

4.2.3.3.5. Arbeitsbezogene Ergebnisse

Zwei Studien [86, 93] lieferten Ergebnisse für arbeitsbezogene Fragen. Van Tulder et al. [93] berichtete, dass chronische Rückenschmerzpatienten den Medianwert von 6.5 Wochen krank geschrieben waren (Spannweite 0.5-48 Wochen). Insgesamt waren 29 % der 232 Studienteilnehmer innerhalb von 12 Monaten ab irgendeinem Zeitpunkt arbeitsunfähig. Storror et al. [86] beobachtete über ein Jahr krank geschriebene Rückenschmerzpatienten. Sie stellten fest, dass nach 6 Monaten 41.3 % wieder in den Beruf einsteigen konnten. Dies steigerte sich auf 48 % nach 12 Monaten. Zugleich bedeutete dies, dass jeder zweite Patient mit chronischen Rückenschmerzen auch nach 1 Jahr noch nicht wieder arbeitsfähig war.

Vergleicht man dieses Ergebnis mit dem Kollektiv der subakuten arbeitsunfähigen Patienten [32, 42, 65, 85] wird deutlich, dass die Ergebnisse sehr ähnlich ausfallen. Nach 6 Monaten sind im Schnitt 40-45 % wieder arbeitsfähig während es nach 12 Monaten 57 % sind. Somit sind nach 1 Jahr etwa 10 % mehr der subakuten Patienten wieder arbeitsfähig im Vergleich zu den chronischen Rückenschmerzpatienten. [86] (Tab. 20)

4.2.3.3.6. Verlauf

Van Tulder et al. [93] lieferte außerdem Daten zum Verlauf der Rückenschmerzen. Über ein Jahr gesehen hatten die Teilnehmer im Medianwert von 16 Wochen Rückenschmerzen und insgesamt 3 Rückenschmerzepisoden. Nach 12 Monaten hatten 6 % der Teilnehmer keine Beschwerden mehr, während auf der anderen Seite 10 % der Teilnehmer ununterbrochen von Beschwerden berichteten. Die restlichen Teilnehmer zeigten intermittierend Beschwerden von unterschiedlicher Dauer. (Tab. 20)

Studie	Pat. N	Messinstrument	Base-line	Pat. N	6 Monate	12 Monate
Arbeitsbezogene Ergebnisse						
Storrø et al.[86]	46	Return to work,n (%), (Number of patients who were taken off the sick list)	k.A.	k.A.	19 (41.3)	22 (48)
Van Tulder et al.[93]	232	Absentism from work (%)	k.A.	k.A.	k.A.	29%
Van Tulder	232	Duration of sick leave (weeks)	k.A.	k.A.	k.A.	Md: 6.5; range: 0.5-48 Wochen
Verlauf						
Van Tulder	368	Duration of LBP Number of Episodes	k.A.	k.A.	k.A.	Md: 16 Wochen Md: 3
Van Tulder	368	-Any complaint -Complaints in one follow up period -Complaints in two follow up periods -Complaints in all 3 follow up periods -Continuously	k.A.	k.A.	k.A.	6% 14.5% 22.5% 57% 10%

Tabelle 20: Arbeitsbezogene Ergebnisse und Verlaufsergebnisse (chronischer Rückenschmerz)

4.2.3.3.7. Abhängigkeit des Verlaufs vom Ausgangswert

Im Bereich der chronischen Rückenschmerzpatienten zeigt sich die leichte Tendenz, dass primär stärker eingeschränkte/erkrankte Patienten im Verlauf von 6 Monaten eine höhere prozentuale Verbesserung erzielen als solche die zu Beginn weniger eingeschränkt sind. Am deutlichsten wird dies im Bereich der Lebensqualität (mediane Ergebnisse). (Abb. 7)

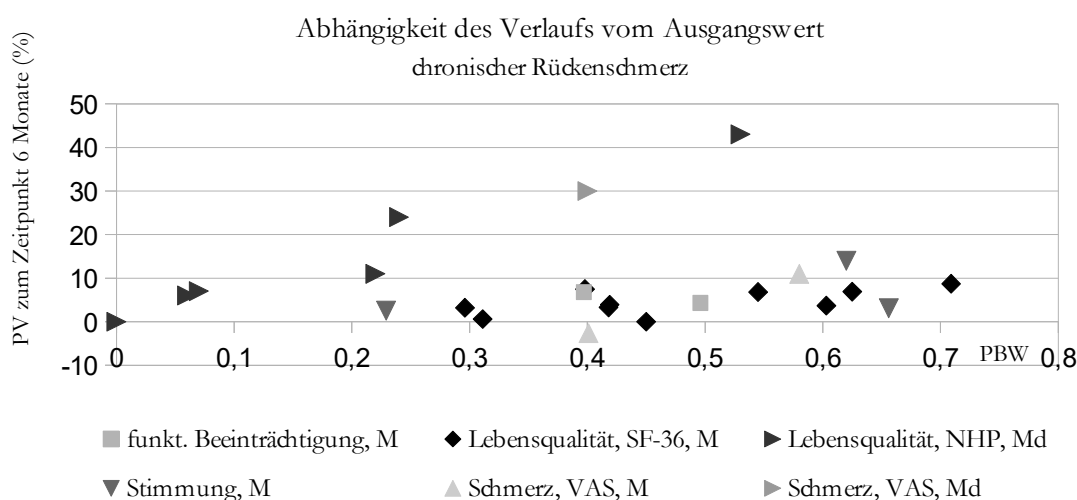


Abbildung 7: Abhängigkeit des Verlaufs vom Ausgangswert (chronischer Rückenschmerz)

Die Polarität der Messinstrumente wurde in dieser Abbildung angeglichen
Ergebnisse in dieser Abbildung sind sowohl als Mittelwert als auch als Median angegeben

4.2.3.3.8. Fazit

Chronische Rückenschmerzpatienten haben nach 12 Monaten eine Verbesserung zwischen 5,8-10 % im Bereich funktionelle Beeinträchtigung zu erwarten und eine Verbesserung zwischen 8-11 % im Bereich Schmerzintensität. Dabei liegen die absoluten Schmerzergebnisse im Schnitt bei 0,48 (VAS, PEW). Die Verbesserung der medianen Schmerzergebnisse beträgt nach 6 Monaten 30 %.

Berücksichtigt man bei der Analyse der Lebensqualität nur Studien, die alle Subskalen analysieren wird deutlich, dass die zu erwartende Verbesserung nach 6 Monaten über alle Subskalen gemittelt zwischen 5,1 % (SF-36, Mittelwerte) [76] und 16 % (NHP, mediane Werte) [89] liegt. Zudem zeigen sowohl Studien mit subakuten als auch Studien mit chronischen Rückenschmerzpatienten, dass bestimmte Bereiche der Lebensqualität eingeschränkter sind als andere. Schlechter fallen vor allem die Bereiche Schmerz, Vitalität und Rollenfunktion in körperlicher Hinsicht (SF-36) sowie Schmerz, Mobilität und Energieverlust (NHP) aus. Das psychische Wohlbefinden, das auf sehr unterschiedliche Weise gemessen wird, ergibt auch eine leichte positive Tendenz über die Zeit mit einer Verbesserung von 5,7-17 % nach 12 Monaten.

Der Vergleich der subakuten und chronischen arbeitsunfähigen Patienten zeigt, dass die Ergebnisse für beide Gruppen nach 6 Monaten sehr ähnlich ist. So sind zu diesem Zeitpunkt 40-45 % der ursprünglich krank geschriebenen wieder arbeitsfähig, während nach 12 Monaten 57 % der subakuten Patienten wieder arbeiten und nur 47 % der chronischen Patienten.

4.2.3.4. Subakuter und chronischer Rückenschmerz (gemischte Gruppe)

Insgesamt sechs Studien lieferten Ergebnisse sowohl für Patienten mit subakuten als auch mit chronischen Rückenschmerzen (Nyiendo et al. [31,70, 80], Lamb et al. [45, 51, 52], Haldorsen et al.[34], Miller et al. [62], van den Hoogen et al. [90]). Dieses Kapitel umfasst all jene Patienten, die zum Zeitpunkt der Rekrutierung mindestens 6 Wochen Beschwerden hatten, jedoch auch bereits länger als 12 Wochen Symptome haben konnten.

Die Studien waren in ihren Details erneut sehr heterogen. So wies die Studie von Lamb et al. [45, 51, 52] ein recht hohes Durchschnittsalter von 54 Jahren auf und schloss Patienten mit leichten Beschwerden aus. Es wurde berichtet, dass der Rückenschmerz zu Beginn der Studie für 54 % der Patienten mäßig beschwerlich war („moderately troublesome back pain“), während dieser für die anderen 46 % der Patienten sehr bis extrem beschwerlich

war („very or extremely troublesome“). Die Teilnehmer hatten im Mittel 13 Jahre (s: 12,7, Spanne: 0,25-53 Jahre) lang Beschwerden, bevor sie in die Studie eingeschlossen wurden. Etwa die Hälfte der eingeschlossenen Patienten war zum Zeitpunkt der Rekrutierung Teil- oder Vollzeit beschäftigt. Bei Miller et al. [62] lag die Dauer der Rückenschmerzen zum Zeitpunkt der Rekrutierung bei einem Medianwert von 10 Wochen (Q1,Q3:7,14) und weicht damit deutlich vom Patientenkollektiv von Lamb et al. [45, 51, 52] ab. Dies liegt daran, dass Miller et al. [62] Patienten mit einer Symptombdauer von mehr als 6 Monaten ausgeschlossen hat. Haldorsen et al. [34] bezog nur Patienten mit Krankschreibung zwischen 8 Wochen und 6 Monaten mit ein.

Zwei Studien hatten mehr als 20 % Dropouts (van den Hoogen et al., Nyiendo et al. [31, 69, 70, 90]). Van den Hoogen et al. [90] wählte zudem kein multizentrisches Design. Dies führte zu folgender Biasrisikobewertung. (Patienten- und Arztauswahl (Domäne A)/Date-nerhebung (Domäne B)): zweimal „niedrig/niedrig“ [45, 51, 52, 62], zweimal „niedrig/unklar“ [31, 34, 69, 70] und einmal „unklar/unklar“ [90]. (Tab. 8) Die klinische Heterogenität/Varianz war zweimal „niedrig“ [31, 69, 70, 90] und dreimal „unklar“ [34, 45, 51, 52, 62]. Diese Bewertung kam zustande, da Patienten mit leichten Beschwerden [45, 51, 52] und arbeitsfähige Patienten [34] ausgeschlossen wurden. Zudem blieb in einer Studie unklar, ob Patienten mit Ausstrahlung in die Beine eingeschlossen wurden, sodass nicht sicher war, ob das Studienkollektiv repräsentativ war. [62]

Eine gute Vergleichbarkeit erreichten die Studien zum einem im Bereich Symptomlokalisation, da alle Studien Patienten mit Schmerzen im unteren Rücken rekrutierten und nur Haldorsen et al. [34] zusätzlich Patienten mit Schmerzen im oberen/mittleren Rücken. Zum anderen sind zwei Studien Kohortenstudien [31, 69, 70, 90] und drei Studien randomisiert kontrollierte Studien [34, 45, 51, 52, 62]. Die Teilnehmerzahl lag zwischen zwischen 81 und 330. Fünf der sechs Studien [31, 34, 45, 51, 52, 69, 70, 90] schlossen Patienten mit Ausstrahlung ins Bein ein, sie sind deshalb besonders gut vergleichbar.

4.2.3.4.1. Funktionelle Beeinträchtigung

Miller et al. [62] berichtete zum Baseline-Zeitpunkt und nach 9 Monaten Werte als Median. Die Teilnehmer starten mit einem Wert von 8 Punkten (Q1,Q3: 4, 12) im RMQ und verbesserten sich auf einen Wert von 2 (IQR: 0,2). Dies entsprach einer prozentualen Verbesserung von 25 %. Die Ergebnisse von van den Hoogen et al. [90] waren nur stark eingeschränkt verwertbar, da sie unvollständig angegeben waren. Die Studie berichtete über die

funktionelle Beeinträchtigung mit der VAS (0-50). Die Patienten verbesserten sich nach 4 Wochen um den Median von 1 Punkt (IQR: 2-0) und nach weiteren 4 Wochen um einen weiteren Punkt (IQR:3-0). Ein Baseline-Wert wurde nicht angegeben. Dennoch wurde deutlich, dass sich die funktionelle Beeinträchtigung leicht verringerte.

Zwei Studien (Lamb et al., Nyiendo et al. [31, 45, 51, 52, 69, 70]) zeigten über den Verlauf von 12 Monate eine Verringerung der funktionellen Beeinträchtigung, angegeben als Mittelwert. Bei Lamb et al. [45, 51, 52] war die Verbesserung insgesamt weniger ausgeprägt und lag bei 4.5 % (PV, Ø:mod. Korff, RDQ) nach 6 Monaten bei etwa 6.5 % (PV, Ø:mod. Korff, RDQ) nach 12 Monaten. Bei Nyiendo et al. [31, 69, 70] war die prozentuale Verbesserung mehr als doppelt so hoch und lag bei 14.4 %. Dagegen zeigten die prozentualen Ergebniswerte, dass die Ergebnisse von Lamb et al. [45, 51, 52] um das Ergebnis von Nyiendo et al. [31, 69, 70] streuten. (Abb. 8)

Der Abgleich der Studiencharakteristika zeigt mögliche Einflussgrößen für die geringere prozentuale Verbesserung von Lamb et. al [45, 51, 52]. Diese Studie schloss Patienten mit leichten Beschwerden aus und hatte ein Patientenkollektiv mit hohem Durchschnittsalter.

Die Abbildung 8 unterstreicht folgendes: Alle Studien zeigen eine Verbesserung der funktionellen Beeinträchtigung, die allerdings unterschiedlich stark ausgeprägt ist.

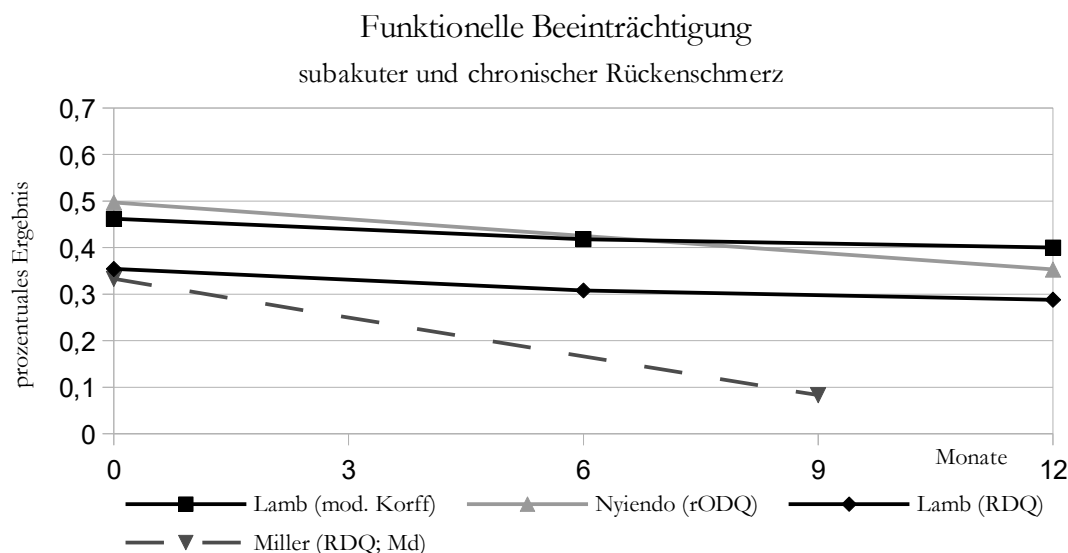


Abbildung 8: Funktionelle Beeinträchtigung (subakuter/chronischer Rückenschmerz)

1- schlechtestes Ergebnis 0- bestes Ergebnis

Ein Vergleich der gemischten Gruppe mit der Gruppe der subakuten und chronischen Rückenschmerzpatienten zeigt, dass die Ergebnisse der drei Gruppen in einem ähnlichem Spektrum streuen. Die Ergebnisse der subakuten Patienten sind etwas besser, als die der chronischen Gruppe. Das Ergebnis der gemischten Gruppe spiegelt das gemischte Patientenkollektiv wider. (Abb. 9) Es zeigt sich, dass sich die funktionelle Beeinträchtigung in diesem Zeitraum über alle Gruppen hinweg um 5.8- 14.4 % verbessert. Der Ergebniskorridor liegt zwischen 0.29-0.4 (PEW). [31, 45, 51, 52, 70, 80]

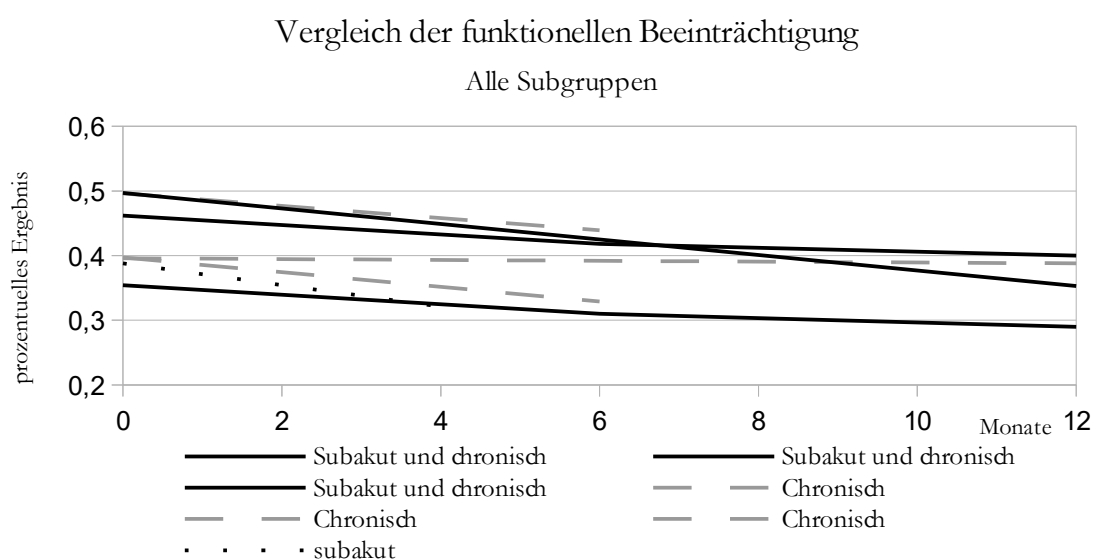


Abbildung 9: Vergleich der funktionellen Beeinträchtigung aller Subgruppen

1- schlechtestes Ergebnis 0- bestes Ergebnis, Nur Studien mit Mittelwerten in die Grafik mit aufgenommen

4.2.3.4.2. Schmerz

Miller et al. [62] berichtete über sein Ergebnis mit der VAS (0-5) als Median. Zum Baseline-Zeitpunkt lag der Wert bei 2 Punkten (Q1,Q3: 1,2) und erreichte nach 9 Monaten den Wert 1 (IQR: 0.2). Auch Lamb et al. [45, 51, 52] stellte nach 6 Monaten eine Verbesserung um- 6.4 % (PV) und um 8.4 % (PV) nach 12 Monaten fest. Nyiendo et al. [31, 69, 70] zeigte auch im Bereich Schmerz eine deutliche Verbesserung um 18.9 % (PV, 12 Monate). Wie bereits im Bereich funktionelle Beeinträchtigung wird auch hier grafisch deutlich, dass die Studien eine Zustandsverbesserung zeigen und das Nyiendo et al. eine höhere prozentuale Verbesserung erreichte. (Abb. 10)

Der Vergleich der gemischt subakut-chronischen Rückenschmerzgruppe mit der Gruppe der subakuten und der Gruppe der chronischen Rückenschmerzpatienten zeigt, dass die Ergebnisse der drei Gruppen in einem ähnlichem Spektrum streuen. Zudem ist keine klare Tendenz zu erkennen, dass die Ergebnisse einer Gruppe besser oder schlechter ausfallen, als die einer anderen. (Abb. 10) Es zeigt sich zudem, dass sich die Schmerzintensität in diesem Zeitraum zwischen -2 und 19 % verändert. Der Ergebniskorridor liegt zwischen 0.35- 0.51 (PEW). (Abb. 11) [31, 45, 51, 52, 70, 80]

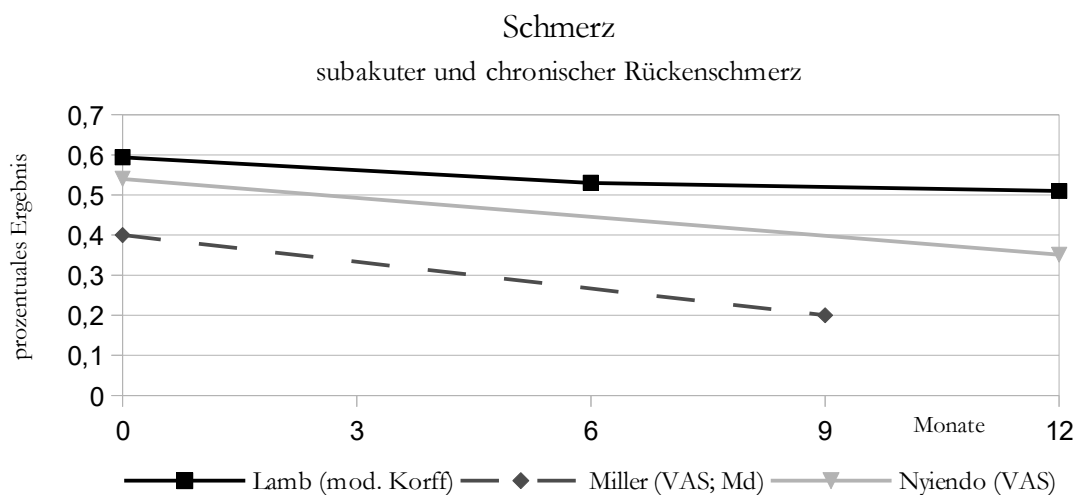


Abbildung 10: Schmerz - gemischte Rückenschmerzpatienten (subakut/chronisch)

1- schlechtestes Ergebnis 0- bestes Ergebnis

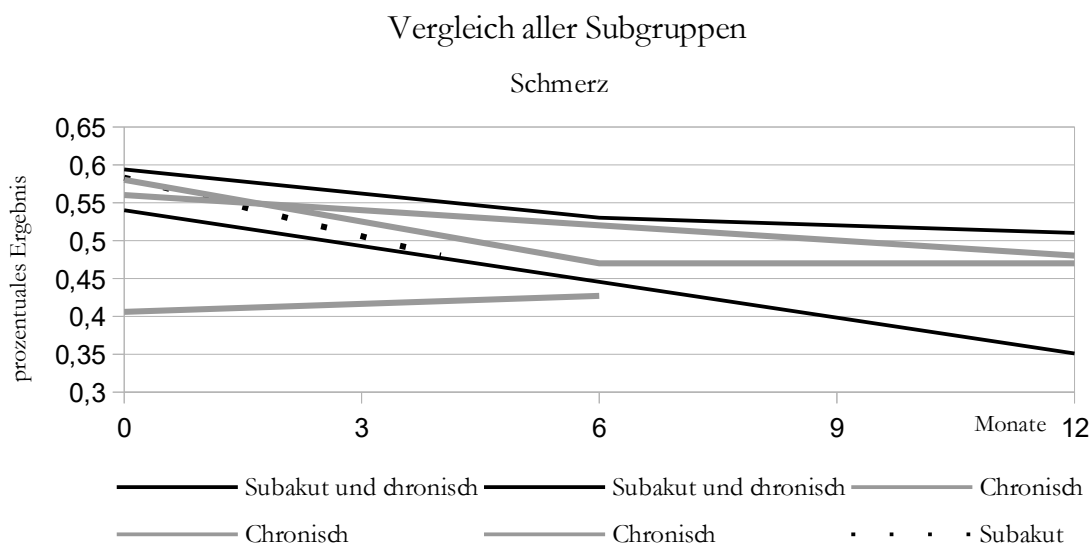


Abbildung 11: Schmerz - Analyse aller Subgruppen

1-schlechtestes Ergebnis 0- bestes Ergebnis

4.2.3.4.3. Lebensqualität

Zwei Studien nutzen als Subskalen den SF-12 (Lamb et al., Nyiendo et al. [31, 45, 51, 52, 69, 70]). Die Ergebnisse von Lamb et al. [45, 51, 52] blieben auf einem mittleren Niveau über die Zeit stabil. Bei Nyiendo et al. lag die prozentuale Veränderung zwischen 4.9-15.8 % mit einem Endergebnis von etwa 0.6. [31, 69, 70] (Abb. 12) Bei Nyiendo et al. [31, 69, 70] fielen die Ergebnisse im mentalen („men“) Bereich besser aus, während bei Lamb et al. [45, 51, 52] die physischen („phy“) Werte besser waren.

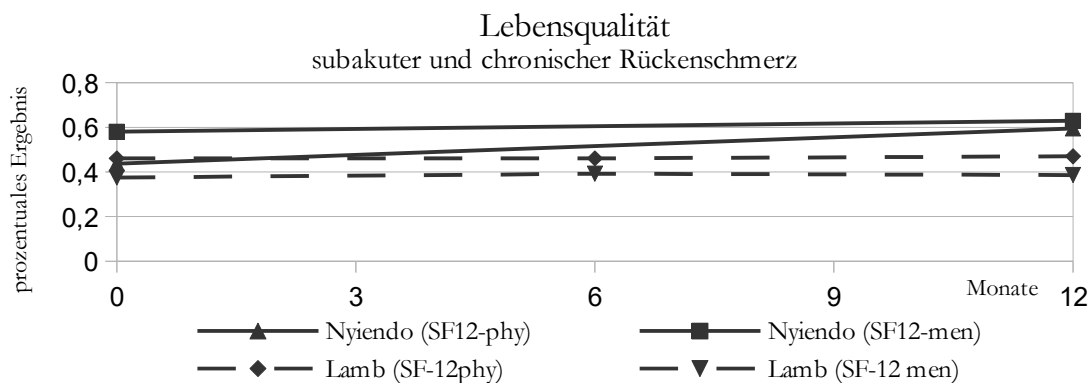


Abbildung 12: Lebensqualität – gemischte Rückenschmerzpatienten (subakut/chronisch)

1- bestes Ergebnis 0- schlechtestes Ergebnis

4.2.3.4.4. Arbeitsbezogene Ergebnisse

Lamb et al. [45, 51, 52] berichtete wie viele Teilnehmer im Verlauf der 12 Monate wegen ihrer Rückenschmerzen nicht arbeiten konnten. („Days off sick from work“). Der Anteil der Krankschreibungen stieg von zunächst 14.8 % auf 29.6 %. Die Dauer der Krankschreibung belief sich nach 6 Monaten auf 6.1 Tage und auf 10.2 Tage nach 12 Monaten.

Haldorsen et al. [34] beobachtete den Verlauf von krankgeschrieben Patienten. Diese waren zu Beginn der Beobachtung mindestens 8 Wochen krankgeschrieben. Nach 12 Monaten Nachbeobachtung waren 58 % der Teilnehmer wieder arbeitsfähig.

Dieses Ergebnis ähnelt den bereits berichteten arbeitsbezogenen Ergebnissen der subakuten Subgruppe und chronischen Subgruppe. (Kapitel 4.2.3.2.5/4.2.3.3.5)

4.2.3.4.5. Stimmung

Im Bereich „fear-avoidance-beliefs“ blieben die Teilnehmer fast konstant auf dem selben Niveau mit einer maximalen Verbesserung von 4.6 % (PV) nach 12 Monaten. [45, 51, 52]

4.2.3.4.6. Verlauf

Lamb et al. [45, 51, 52] beschrieb nach 6 und 12 Monaten wie der Verlauf der Rückenschmerzen war. Den 7 Verlaufskategorien („vollständig erholt“ bis „extrem verschlimmert“) wurden zur besseren Auswertung Zahlenwerte zugeordnet (0-6). Dies ergab einen Mittelwert von 3.1 nach 12 Monaten. Waren nach 6 Monaten nur 1 % der Teilnehmer vollständig gesund („completely recovered“) stieg dieser Wert nach 12 Monaten auf 3 %. Am anderen Ende der Skala hatten sich nach 6 Monaten 2 % der Patienten extrem verschlimmert („vastly worsened“) und nach 12 Monaten 3 %. Wenn man diese 7 Kategorien in 3 Gruppen aufteilt (verbessert („completely recovered“, „much improved“, „slightly improved“), gleichbleibend („no change“) und verschlimmert („slightly worsened“, „much worsened“, „vastly worsened“)), ergab sich folgende Verteilung: Nach 12 Monaten hatte sich bei 31 % die Symptomatik verbessert, bei 45 % waren die Schmerzen gleichbleibend, während es sich bei 24 % verschlimmert hatte. Van den Hoogen et al. [90] konnte berichten, dass sich die Patienten im Median nach etwa 10 Wochen erholten. (Tab. 21)

Studie	Pat. N	Messinstrument	Base-line	Pat. N	PBW	6 Monate	PEW	PV, %	12 Monate	PEW	PV, %
Stimmung											
Lamb et al. [45, 51, 52]	233	Fear-avoidance-beliefs questionnaire, 0-24	M:14.2 s:6.22	219	0.592	M:13.8 s:6.09	0.575	1.67	M:13.1 s:6.18	0.546	4.58
Arbeitsbezogene Ergebnisse											
Lamb et al.	233	Days off sick from work: yes, n(%)	k.A.	k.A.	k.A.	13(14.8)	k.A.	k.A.	29(29.6)	k.A.	k.A.
Lamb et al.	233	No. of days sick	k.A.	k.A.	k.A.	M:6.1 s:4.4	k.A.	k.A.	M:10.2 s:12.9	k.A.	k.A.
Haldorsen et al. [34]	81	Return to work	0%	k.A.	k.A.	k.A.	k.A.	k.A.	58%	k.A.	k.A.
Verlauf											
Lamb et al.	233	Course of LBP: <ul style="list-style-type: none"> • Completely recovered (6) • Much improved (5) • Slightly improved(4) • No change(3) • Slightly worsened(2) • Much worsened(1) • Vastly worsened(0) 	k.A.	k.A.	k.A.	M:3.04 2(1%) 19(10%) 35(19%) 81(44%) 27(15%) 18(10%) 4(2%)	k.A.	0.507	M:3.1 6(3%) 24(12%) 32(16%) 89(45%) 24(12%) 17(9%) 5(3%)	0.517	1

Lamb et al.	233	Some/complete recovery	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.	31%	k.A.	k.A.
v.d. Hooogen et al. [90]	121	Time till recovery (weeks)	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.	Md:10 IQR 5-25	k.A.	k.A.

Tabelle 21: Ergebnisse gemischter Rückenschmerzpatienten (subakut und chronisch)

4.2.3.4.7. Abhängigkeit des Verlaufs vom Ausgangswert

Das Diagramm setzt prozentuale Baseline-Werte in Zusammenhang mit der prozentualen Abweichung nach 12 Monaten. Hier zeigt sich, dass Patienten mit ausgeprägter Beeinträchtigung und hoher Schmerzintensität zu Beginn der Untersuchung keine deutlich höhere Verbesserung im Vergleich zu Patienten erzielen, die zu Beginn weniger eingeschränkt sind. (Abb. 13)

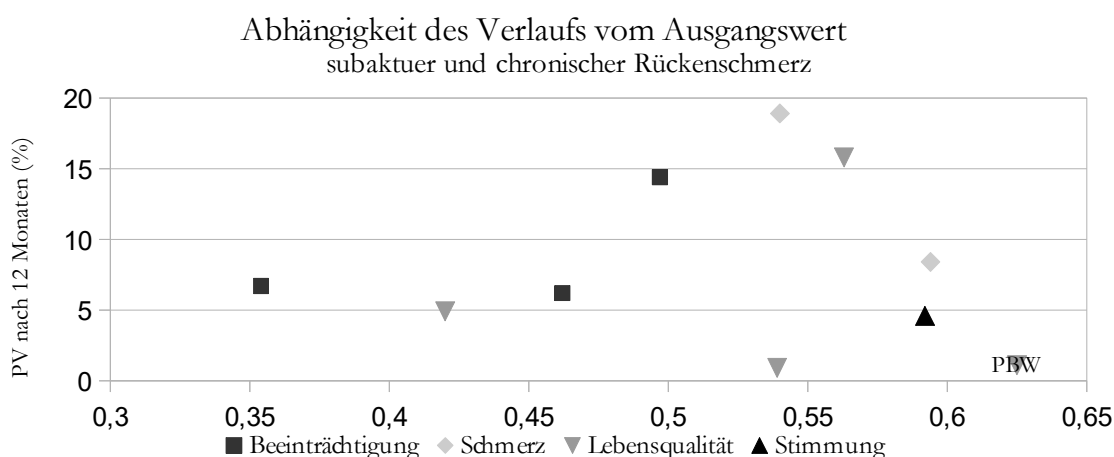


Abbildung 13: Abhängigkeit des Verlaufs vom Ausgangswert gemischte Rückenschmerzpatienten (subakut/chronisch)

4.2.3.4.8. Fazit

Im Bereich funktionelle Beeinträchtigung berichten die Studien eine prozentuale Veränderung zwischen 6-15 % (Studienergebnisse, Mittelwert) und 25 % (Studienergebnis, Median). Der Bereich Schmerzintensität weißt prozentuale Veränderung von 6-19 % (Studienergebnisse, Mittelwert) und 20 % (Studienergebnis, Median) auf.

Dabei zeigt ein Vergleich der drei Subgruppe „subakut“, „chronisch“ und „gemischt subakut und chronisch“, dass die Ergebnisse der drei Gruppen in einem ähnlichen Korridor streuen. Jedoch ohne deutliche Tendenz, dass die Ergebnisse der subakuten Patienten signifikant besser ausfallen als die der chronischen Rückenschmerzpatienten. Diese Beobachtung konnte im Bezug auf den Aspekt „Rückkehr zum Arbeitsplatz“ bei arbeitsunfähigen Patienten wiederholt werden.

4.2.3.5. Rückenschmerz mit Ausstrahlung in die Beine

Zwei Studien (Nyiendo et al., Storro et al. [69, 86]) lieferten Ergebnisse getrennt für Patienten, die nur Rückenschmerz haben und für die, die Rückenschmerz mit Ausstrahlung in die Beine haben. Besonders bemerkenswert war in diesem Abschnitt der Unterschied zwischen den Ergebnissen von Patienten, deren Schmerzausstrahlung bis unterhalb des Knies reichte und den anderen beiden Gruppen („Nur Rückenschmerz“ und „Rückenschmerz mit Ausstrahlung in den Oberschenkel“).

Nyiendo et al. [69] beschäftigte sich mit Teilnehmern, die sowohl subakuten als auch chronischen Rückenschmerz hatten (>6 Wochen Beschwerden). Insgesamt 310 Patienten nahmen an der Studie teil. Mit einer Dropout-Rate von 38 % erhielt die Studie eine Biasrisikobewertung von „niedrig/unklar“. (Tab. 8)

Storro et al. [86] schloss nur Patienten mit einer Krankschreibung zwischen 16 und 52 Wochen ein. Identifiziert wurden sie über den ICPC-Code für Rückenschmerz auf den Krankschreibungen. Dies führte zu einer Biasrisikobeurteilung von „unklar/unklar“. (Tab. 8) Insgesamt 36 % der 72 Teilnehmer hatten Rückenschmerz mit Schmerzausstrahlung. Storro et al. erhielt aufgrund der reduzierten Repräsentativität die Bewertung „unklar“ im Bereich klinische Heterogenität/Varianz. (Tab. 9)

4.2.3.5.1. Schmerz

Besonders auffällig ist die unterschiedliche Entwicklung von Patienten mit Schmerzausstrahlung unterhalb der Knie im Vergleich zu Patienten, die ausschließlich Rückenschmerz hatten oder deren Schmerzen nur in den Oberschenkel ausstrahlen. Obwohl Patienten mit Ausstrahlung oberhalb des Knies und solche mit Schmerzausstrahlung unterhalb der Knie zum Baseline-Zeitpunkt fast identische Schmerzwerte erzielten (56.1, 57.1), erreichte die erste Gruppe nach 12 Monaten eine prozentuale Verbesserung von 32 %, während die zweite Gruppe nur 14 % erzielt. Patienten, die ausschließlich Schmerzen im Rücken haben, haben von vornherein ein niedrigeres Schmerzniveau (48.5) und verbessern sich um 21 % (PV). (Abb. 14)

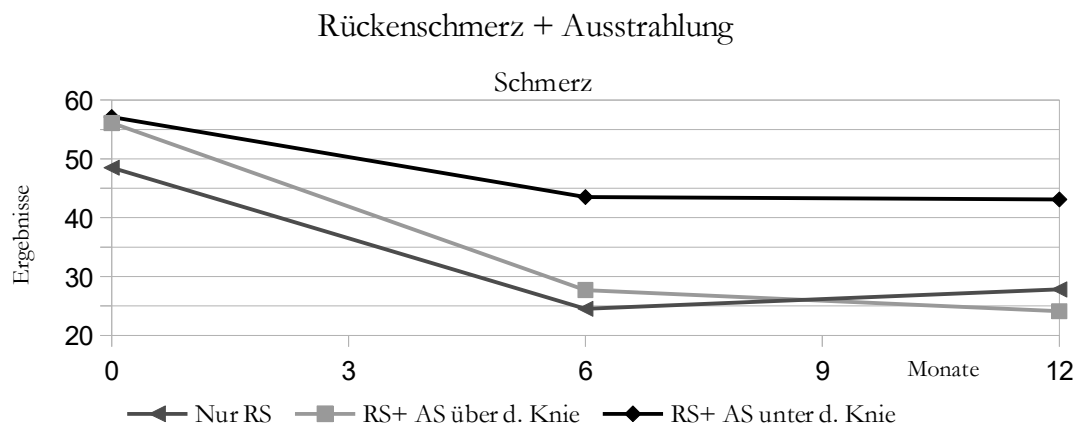


Abbildung 14: Schmerz (Rückenschmerz mit Ausstrahlung)

1- schlechtestes Ergebnis 0- bestes Ergebnis

4.2.3.5.2. Funktionelle Beeinträchtigung

Ein sehr ähnliches Bild zeichnet sich auch im Bereich der funktionellen Beeinträchtigung ab. Die schlechteste Prognose haben Patienten mit Schmerzausstrahlung unterhalb des Knies mit einer Verbesserung um 14 % (PV) nach 12 Monaten bei einem Ausgangswert von 56.3. Danach folgen die Teilnehmer mit Schmerzausstrahlung oberhalb des Knies, mit einer Verbesserung von 23 % bei einem Ausgangswert von 48.1. Nochmals leicht besser schneiden Patienten ab, die ausschließlich Rückenschmerzen haben. Diese haben zwar eine etwas geringere Verbesserung, starten allerdings auf einem niedrigeren Schmerzniveau (42.7) und verbessern sich im Verlauf der 12 Monate um 18.9 % (PV). (Abb. 15)

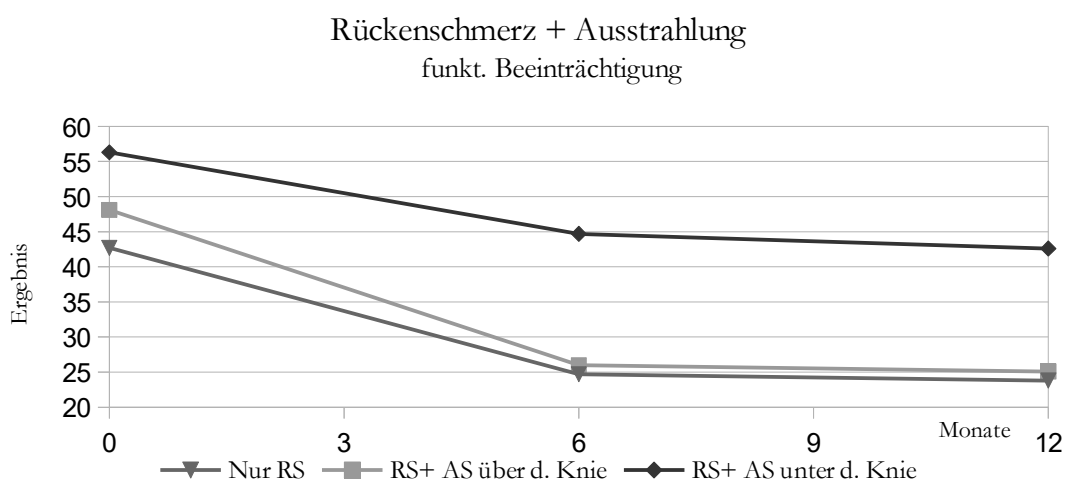


Abbildung 15: Funktionelle Beeinträchtigung (Rückenschmerz mit Ausstrahlung)

0- bestes Ergebnis, 100- schlechtestes Ergebnis, RS- Rückenschmerz, AS- Ausstrahlung

4.2.3.5.3. Arbeitsbezogene Ergebnisse

Im Bereich „Krankheitsausfälle aufgrund von Rückenschmerzen“ sind die Unterschiede zwischen Patienten mit Ausstrahlung und ohne Ausstrahlung in die Beine weniger stark ausgeprägt.

Nach 12 Monaten sind in 46.2 % der Patienten mit Ausstrahlungssymptomatik wieder im Berufsleben, während es in der anderen Gruppe 48 % sind.

5. Diskussion

5.1. Zusammenfassung der Ergebnisse

Die Zusammenfassung der Ergebnisse findet sich im Kapitel der jeweiligen Subgruppen-diskussion, siehe Kapitel 5.3.4.1-5.3.4.4.

5.2. Methoden

Entsprechend der „Cochrane Collaboration“ [39] und den „PRISMA Richtlinien“ [55, 63] zeichnen verschiedene Qualitätsmerkmale (siehe unten) gute systematische Übersichtsarbeiten aus. Dabei sollte allerdings beachtet werden, dass diese Checkliste entwickelt wurde, um systematische Übersichtsarbeiten von Diagnose- und Therapiestudien zu erstellen. Dementsprechend war eine Anpassung der Kriterien für diese systematische Übersichtsarbeit symptomevaluierender Studien nötig.

- Verfassen eines Studienprotokolls
- Festlegung der Forschungsfrage a priori
- Festlegung der Ein- und Ausschlusskriterien a priori
- Breit angelegte Suchstrategie
- Transparente Entscheidungen über Ein- und Ausschluss von Studien
- Bias- bzw. Qualitätsbeurteilung der eingeschlossenen Studien
- Beurteilung durch mindestens zwei unabhängige Personen
- Datenextraktion nach standardisierten Kriterien
- Beurteilung der Heterogenität der Studien [6, 28, 39, 55, 63, 96]

Ein Studienprotokoll ist vor allem in Hinsicht auf die mögliche Kontrolle nachträglicher Abweichungen vom Studienprotokoll relevant. Für diese Arbeit wurde das Vorgehen und die Struktur zuvor festgelegt, jedoch nicht in dem formellen Rahmen eines Studienprotokolls. Dies schmälert etwas die Möglichkeit einer externen Qualitätskontrolle.

Ein wesentlicher Aspekt von systematischen Übersichtsarbeiten ist des Weiteren, die genaue Festlegung der zu untersuchenden Forschungsfragen, bevor mit der Suche und Bewertung der Studien begonnen wird. Dieses Qualitätsmerkmal wurde bei der Umsetzung dieser Ar-

beit beachtet und sorgte dafür, dass die Literatur mit einer festgelegten Zielsetzung durchsucht wurde. Dies fördert die systematische Auswahl von Studien und verhindert, dass Studien aus einem subjektiven Interesse des Bearbeiters ausgewählt werden. Zudem kann die identifizierte Literatur nur dann hinreichend bewertet und ausgewertet werden, wenn zuvor feststeht auf welche Daten die Suche abzielt.

Für diese Arbeit wurde Pubmed/Medline als Quelle genutzt. Diese Datenbank deckt etwa 4800 internationale Zeitschriften ab mit über 24 Millionen Publikationen. [97] Über eine Schneeballsuche konnten weitere Publikationen identifiziert werden. Kritisch zu betrachten ist in diesem Zusammenhang der Entschluss, die Datenbank Embase nicht in die Suche aufzunehmen. Relevant ist diese Einschränkung im Bezug auf die möglichst vollständige Identifikation von Studien. Es ist davon auszugehen, dass durch dieses Vorgehen einige einschlägige Studien nicht gefunden wurden. So zeigten zwei Dissertationen die systematische Übersichtsarbeiten zu anderen Symptomen symptomevaluierender Studien durchgeführt haben, dass in Embase noch einmal die 1.5 bis 3-fache Menge an Abstracts identifiziert werden konnten. [20a, 84a] Darüber hinaus wäre es wünschenswert gewesen auch nach nicht publizierten Studien zu suchen, da es denkbar ist, dass die Ergebnisse publizierter und nicht-publizierter Studien voneinander abweichen. [96] Ein Publikationsbias ist aufgrund des gewählten Vorgehens, in dieser systematischen Übersichtsarbeit nicht völlig auszuschließen. In ähnlichem Maße ist auch die Schneeballsuche eine Quelle möglicher Verzerrungen, da Autoren dazu neigen, eher Literatur zu zitieren, welche die eigenen Ergebnisse unterstützt, anstatt solche, die gegenteilige Ergebnisse wiedergibt. [39]

Ein weiterer Punkt ist die Auswahl der zulässigen Sprachen. Zunächst wurden Publikationen in allen Sprachen identifiziert. Im Verlauf der Datenanalyse wurden jedoch nur noch Publikationen in Englisch und Deutsch extrahiert, da ein verlässliches Verständnis der anderen Sprachen nicht zu gewährleisten war. Dies ist bedauerlich, da auf diese Weise 25 Studien ausgeschlossen wurden, die möglicherweise einschlägig gewesen wären. Andererseits konnte so verhindert werden, dass durch mangelnde Sprachkenntnisse möglicherweise falsch interpretierte Daten ausgewertet wurden. Bei einer Untersuchung zum Thema Sprachen in systematischen Übersichtsarbeiten/Metanalysen wurde zudem festgestellt, dass der Ausschluss nicht-englischer Studien häufig keinen relevanten Einfluss auf das Ergebnis von Metaanalysen hat. [43, 64]

Wie von der Checkliste gefordert, wurde die Suchsyntax mindestens einer Suche in dieser Dissertation dargelegt, um eine Replizierung der Suche zu ermöglichen.

Ein weiteres Qualitätsmerkmal ist die strukturierte Bewertung der Studien anhand zuvor festgelegter Ein- und Ausschlusskriterien. Zudem muss der Grund für den Ausschluss einer Studie dokumentiert werden. Dies wurde entsprechend der Richtlinien umgesetzt.

Die darauf folgende Qualitätsbewertung der eingeschlossenen Studien war bei bestimmten Einzelaspekten mit einigen Schwierigkeiten verbunden. Dazu gehörte z.B. die Differenzierung zwischen dem Beratungsanlass (Rückenschmerz) und dem Beratungsergebnis (spezifischer/nichtspezifischer Rückenschmerz oder eine spezifische Diagnose). Die Unterscheidung zwischen diesen beiden Zuständen war für die vorliegende Arbeit wichtig. Zum besseren Verständnis dieses Aspektes sollte man wissen, dass Rückenschmerzen nur selten auf eine konkrete Pathologie zurückgeführt werden können, sodass über 85 % [3] der Patienten, die Klassifikation „nichtspezifischer Rückenschmerz“ erhalten und keine weitere Diagnostik angestrebt wird. Die Unterscheidung dieser beiden Zustände war jedoch nicht in jeder Publikation ausreichend exakt in Hinblick auf die Belange der vorliegenden Arbeit. Die mangelnde Differenzierung ist zum einen der Grauzone zwischen der Diagnose „nichtspezifischer Rückenschmerz“ und dem Symptom „Rückenschmerz“ geschuldet, zum anderen aber auch abhängig von der jeweiligen Publikation. Vor allem in Studien, die ihre Patienten über Aktendurchsicht oder Datenbanksuchen identifizierten, fiel es schwer dieses Kriterium zu prüfen, da die zugrunde gelegten Kriterien nicht ausreichend dargelegt wurden.

Ein weiterer Gesichtspunkt betrifft die Beurteilung des Aspekts „usual-care“. Damit ist die übliche Therapie in der Hausarztpraxis gemeint, die gemäß der Leitlinien [13] verschiedene Dinge beinhalten kann und soll: Anamnese, körperliche Untersuchung und ein zurückhaltender Umgang mit bildgebenden Verfahren. Patienten sollten schmerzadaptiert aktiv bleiben und Bettruhe meiden. Standardmedikation sind nicht-steroidale Antiphlogistika. Nach Ermessen des Arztes werden Überweisungen und weitere Behandlungen eingeleitet wie z.B.: Wärmetherapie, Entspannungsverfahren, Physiotherapie, Massagen, etc. sowie Konsultationen bei Fachärzten (z.B. Orthopäden, Rheumatologen, etc.). In vielen Studien musste entschieden werden, ob die Details der Behandlung der Definition von „usual-care“ entsprachen. Dieser Entscheidungsprozess war schwierig und an vielen Stellen eine Einzelfallentscheidung. Um diese Fehlerquelle bzw. Verzerrungsquelle zu entgehen, hätten von vornherein Therapiestudien ausgeschlossen werden müssen. Das hätte allerdings die Aussagekraft dieser systematischen Übersichtsarbeit deutlich verringert, da weniger Studien eingeschlossen worden wären, die das Gesamtbild durch ihre Informationen bereichert haben.

Der heterogenen Definition und Umsetzung von „usual-care“ in den Kontrollgruppen widmete sich bereits Sommerville et al. [84]. Sie erschwert den Vergleich von einzelnen Studien und den Ergebnissen ihrer Kontrollgruppen. Ein Versuch „usual-care“ zu standardisieren, zeigte bisher nur einen geringen Effekt im Verhalten der Ärzte. [19] Es liegt deshalb in der Verantwortung des jeweiligen Wissenschaftlers, die Heterogenität bestmöglich zu berücksichtigen.

Häufig erwies sich auch die Beurteilung des Ein- und Ausschlusskriteriums „Selektion“ (Tab. 3) als schwierig. Immer wieder gingen Autoren nicht ausreichend detailliert darauf ein auf Basis welchen Ablaufs sie Patienten rekrutierten. Ein bedenklicher Rekrutierungsprozess wäre es, wenn Patienten aufgrund einer Verdachtsdiagnose an Dritte überwiesen würden und die Patienten dann dort aufgrund dieser „Vorselektion“ rekrutiert würden. Dieses Vorgehen wäre mit den Einschlusskriterien dieser Arbeit nicht in Einklang zu bringen. Fand jedoch eine Vorauswahl entsprechend der Ein- und Ausschlusskriterien der jeweiligen Studie statt, führte dies zwar ggf. zur einer geringeren Repräsentativität der Studie, widersprach jedoch nicht den Einschlusskriterien der vorliegenden Arbeit. Eine sichere Differenzierung zwischen diesen beiden Arten der Patientenselektion war aufgrund fehlender Angaben in den Publikationen nicht immer mit absoluter Gewissheit möglich. Daraus ergibt sich ein Risiko für Verzerrungen.

Wie bereits von Ziegler et al. formuliert, lässt sich trotz der Bemühung zur Standardisierung die Auswahl der Studien nicht vollständig objektivieren, da die in Publikationen dargestellten Informationen häufig nicht alle Aspekte hinreichend abbilden. [96] Indem jedoch die Einschlägigkeit aller Studien von zwei unabhängigen Beurteilern bewertet wurde, wurde versucht das Risiko einer Fehlbeurteilung zu minimieren. Der prozentuale Anteil an übereinstimmenden Beurteilungen wurde in dieser Arbeit nicht protokolliert und hätte die Transparenz der Übersichtsarbeit weiter erhöhen können.

Zur Qualitäts- bzw. Biasbeurteilung der eingeschlossenen Publikationen wurde ein momentan in Entwicklung befindliches Instrument der Abteilung Allgemeinmedizin, Präventive und Rehabilitative Medizin des Fachbereichs Medizin der Philipps-Universität Marburg verwendet. Um die Qualität der Studien besser beurteilen zu können erfolgte die Beurteilung durch zwei unabhängige Personen. Die Qualität wurde dabei sowohl im Bereich der allgemeinen Studiencharakteristika als auch für die jeweiligen Ergebniskategorien bewertet. Bisher gibt es keine evaluierten Messinstrumente zur Beurteilung symptomevaluierender Studien. Das hier eingesetzte Instrument muss seine Validität und Reliabilität erst noch unter

Beweis stellen. Dennoch wurde durch den Einsatz dieses Instruments eine strukturierte und nachvollziehbare Bewertung aller Studien ermöglicht.

Um eine selektive Datenauswahl zu verhindern, sollten Daten in Formulare übertragen werden, die ein Schema vorgeben nach dem Informationen aus allen Studien gleichermaßen extrahiert werden. Dadurch wird erreicht, dass nicht individuell entschieden wird welche Daten relevant sind und welche nicht. Dies wurde in dieser Arbeit erreicht, indem die Daten in ein zuvor angelegtes SPSS Dokument übertragen wurden.

Da keine Metaanalyse der eingeschlossenen Studien durchgeführt wurde, kamen keine statistischen Mittel zum Einsatz, um die Heterogenität der Studien zu evaluieren. Stattdessen wurden die Studien einzeln analysiert und deren Besonderheiten hervorgehoben und zu einem Gesamtbild zusammengefasst. Diese Schritte liefern bereits ein recht umfassendes Bild der eingeschlossenen Studien, können jedoch die Auswertung durch weitere statistische Mittel nicht völlig ersetzen.

Durch die Verwendung prozentualer Ergebnisse und der prozentualen Veränderungen wurde in dieser Arbeit ein guter Weg gefunden mit der Heterogenität von Messinstrumenten umzugehen. Dies könnte auch für zukünftige Übersichtsarbeiten geeignet sein, die in der Analyse deskriptiv bleiben und keine Metaanalyse den den dazugehörigen statistischen Mitteln verwenden.

5.2.1. Schlussfolgerung

Wesentliche methodische Merkmale einer systematischen Übersichtsarbeit wurden berücksichtigt und umgesetzt. Dies stellt alle weiteren Ergebnisse dieser Arbeit auf ein solides Fundament. Schwierigkeiten, die mit dem Thema und Design der Arbeit zusammenhängen, wurden bestmöglich gelöst. Das größte Risiko von möglichen Verzerrungen liegt im eingeschränkten Umfang der Literatursuche und im Ausschluss nicht-englisch oder deutschsprachiger Literatur. Die Literatursuche in Medline war trotz dieses Risikos umfassend und ermöglichte den Zugang zu einem Großteil der publizierten Literatur.

5.3. Diskussion der Ergebnisse

5.3.1. Grundlegendes

Die eingeschlossenen Studien sind in vielerlei Hinsicht inhomogen und erschweren deshalb die Formulierung einer Gesamtaussage. Dies liegt neben den eingeschlossenen Publikationen auch am Thema dieser Übersichtsarbeit. Zunächst kann man darüber diskutieren, wie man Rückenschmerz definieren will und wie man seine Lokalisation beschreibt oder begrenzt. Auch der Versuch Rückenschmerzpatienten in homogene Subgruppen zu unterteilen, ist weniger trivial als zunächst gedacht. Ein gängiges Vorgehen in der Wissenschaft ist die Unterteilung der Patienten in akut, subakut und chronisch sowie ggf. in den Aspekt „Ausstrahlung in die Beine“. Einige Arbeiten weisen allerdings darauf hin, dass die Unterteilung der Patienten ausschließlich auf Basis der Dauer der Beschwerden nicht zielführend ist. [38] Es zeigte sich, dass Rückenschmerz sich vor allem über seinen episodischen Charakter definiert. [49, 84] Die wissenschaftliche Einteilung in Subgruppen scheint nicht den klinischen Alltag widerzuspiegeln, in denen wiederkehrende akute Episoden nur schwer von episodisch auftretenden chronischen Beschwerden zu unterscheiden sind. [84, 94] Erschwert wird die Situation darüber hinaus noch, indem sich über die Zeit viele Ansätze und Definitionen für die Unterteilung in Subgruppen entwickelt haben. Für die Subgruppe der chronischen Patienten gibt es folgende Definitionen: mehr als 7 Wochen Beschwerden, mehr als 12 Wochen Beschwerden, mehr als 6 Monate Beschwerden, mehr als die Hälfte der Tage Beschwerden im Verlauf eines Jahres. [13, 14, 94]. Schließlich wurde für diese Arbeit die Definition der nationalen Versorgungsleitlinie übernommen. Diese besagt, dass Beschwerden „subakut“ sind ab einer Beschwerdedauer von 6 Wochen und „chronisch“ ab einer Beschwerdedauer von 12 Wochen. [13] Dennoch sollte man festhalten, dass das Symptom Rückenschmerz in seinem ureigenen Erscheinungsbild bereits Herausforderungen für die wissenschaftliche Arbeit mit sich bringt.

Ausdrückliches Ziel dieser systematischen Übersichtsarbeit war es, symptomevaluierende Studien zum Thema „subakuter und chronischer Rückenschmerz“ zu identifizieren und zu analysieren. Nur wenige Studien legten in ihrem Design Wert auf die Unterscheidung zwischen Beratungsergebnis/Diagnose und Beratungsanlass/Symptom. Damit waren wie bereits in der Diskussion der Methodik erörtert verschiedene Probleme vergesellschaftet. Insbesondere in der Subgruppe der chronisch Kranken ist der Begriff „chronischer Rückenschmerz“ sowohl Ausdruck ihres Symptoms als auch ihrer Diagnose/Klassifikation. Dies

machte die Bewertung des Aspekts „symptomevaluierend“ bei vielen Publikationen zu einer Herausforderung.

Ein Ansatz von Studien ist es, Patienten über die Codierung zu identifizieren oder zu analysieren. Dieser Ansatz muss kritisch diskutiert werden, insbesondere bei der Zielsetzung Patienten über ihr Symptom einzuschließen. Es gibt drei große Codierungssysteme, die in der Primärversorgung zum Einsatz kommen: ICD-10, Read Codes und ICPC. Während ICD-10 für detaillierte Mortalitätsstatistiken entwickelt wurde, wurde der ICPC speziell für den Einsatz in der Primärversorgung entwickelt und ist so besser geeignet den Praxisalltag abzubilden und gleichzeitig für die epidemiologische Forschung eingesetzt zu werden. [15, 54] Insbesondere erfüllt der ICPC das Kriterium neben der Diagnose auch den Beratungsanlass bzw. das vorgetragene Symptom zu verschlüsseln. Der ICPC hat eine biaxiale Struktur durch die Symptome, Diagnose und Procedere verschlüsselt werden. Im Detail setzt sich der ICPC aus Kapiteln (Buchstaben) und Komponenten (Zahl) zusammen. Hierbei bestimmt der Buchstabe das Organsystem und die Zahl, ob es sich um ein Symptom, eine Diagnose oder ein Procedere handelt. [7, 9, 54] Bei Studien, die Codierungen für die Rekrutierung oder Erhebung von Daten nutzen ist die Validität der Daten zweifelhaft, da die zugrunde liegende Intention der Verschlüsselung meist die Abrechnung ist. Ein nicht zu vernachlässigender Aspekt ist außerdem, wie genau die jeweilige Publikation auf die Rekrutierungs- bzw. Datenerhebungsstrategie eingeht. Für das Ziel dieser Arbeit wäre es vor allem wichtig gewesen, dass die Studien angeben, welche Codierungen sie genau verwendet haben und ob die Teilnehmer aufgrund ihrer Diagnose und/oder ihres Beratungsanlasses rekrutiert wurden.

Erler et al. hat die Übereinstimmung von Beratungsanlässen und Codierung (ICD-10) bei Hausärzten in Deutschland untersucht. Er kam zu dem Schluss, dass leichte Erkrankungen auf Symptomebene eher häufiger vorkommen als sie codiert werden („underreported“), während ätiologisch gesicherte Erkrankungen mit höherem Schweregrad seltener vorkommen als die codiert werden („overreported“). Der Grad der Übereinstimmung zwischen Codierung und Beratungsanlass variierte für verschiedene Symptome. Am besten war die Übereinstimmung bei arterieller Hypertonie und Diabetes, doch auch im Bereich Rückenschmerz war die Übereinstimmung gut. [24] Es wird außerdem berichtet, dass der Grad des „overreporting/underreporting“ sich von Land zu Land unterscheidet. Erler et al. vermutet, dass dies zum einen mit der Abrechnungsart und zum anderen mit der Verwendung von elektronischen Patientenakten zusammenhängt. Dies ist ein weiterer Aspekt, der die

Beurteilung von Ergebnissen von Datenbankstudien beeinflusst. Britt et al. [12] überprüfte die Übereinstimmung zweier Beurteiler, die auf Videoband gezeigte Patientenfälle codierten (ICPC). Die Übereinstimmung war am höchsten, wenn es darum ging, das betroffene Organsystem zu codieren, und nahm ab, je weiter es ins Detail ging. Dies lässt die Vermutung zu, dass die Übereinstimmung von Codierungen ausreichend übereinstimmt, wenn mehrere Ärzte in eine Studie involviert sind. Dies gilt vor allem, wenn es weniger Codierung von Details geht, sondern um gröbere Kategorisierungen. [12] Inwiefern die Datenbankeinträge die Realität widerspiegeln, hängt somit zum einen davon ab auf welcher Detailebene man abzielt und zum anderen, um welches Symptom/Diagnose es geht. Darüber hinaus ist davon auszugehen, dass die Intensität der Schulung der beteiligten Codierer eine wesentliche Rolle spielt.

Die Vielzahl an eingesetzten Messinstrumenten sowohl für „funktionelle Beeinträchtigung“ als auch für „Lebensqualität“ und „Schmerz“ erschwerte zudem die Zeichnung eines Gesamtbildes. Dies führt dazu, dass im Bereich Prognose, wo die meisten Ergebnisse generiert wurden, die Vergleichbarkeit erschwert war. Ähnliche Probleme berichten auch andere systematische Übersichtsarbeiten im Bereich Rückenschmerz. [84] Mit dem Einsatz der prozentualen Ergebnisse und Veränderungen wurde in dieser Arbeit ein Weg gefunden, mit der Heterogenität der Messinstrumente umzugehen. Dieser Ansatz könnte auch für zukünftige Übersichtsarbeiten eine Option sein, die Studien zusammenfassen wollen, die eine Vielzahl an Messinstrumenten verwenden.

Die methodische Qualität der eingeschlossenen Studien erwies sich als recht homogen, sowohl was die Stärken betrifft als auch was die Schwächen betrifft. Methodisch einwandfrei zeigten sich die Studien bei der Bewertung der Ergebniserhebung (Ätiologie/Prognose). Einschränkungen fanden sich vor allem bei der Auswahl der Patienten und der Datenerfassung. Häufig blieb unklar, ob Studien Patienten mit Ausstrahlung in die Beine einschlossen und damit auch wie repräsentativ das Patientenkollektiv war. Außerdem fanden sich viele Studien, die nur arbeitsunfähige Patienten untersuchten oder Patienten mit leichten Beschwerden ausschlossen. Etwa ein Drittel erfüllte zudem nicht das Kriterium der Multizentrität. Da viele Studien Patienten über Datenbanken identifizierten, blieb die Datenerhebung des Symptoms und der Einschlusskriterien häufig indirekt. Zudem fanden sich bei etwa einem Drittel der Studien Dropouts von mehr als 20 %. Alles in allem ergab sich allerdings kein Anlass zur Annahme, dass die Qualität der Einzelstudien die Aussagekraft dieser systematischen Übersichtsarbeit gefährdete. Im diesem Zusammenhang ist auch die Rede

von „garbage in – garbage out“, damit ist gemeint, dass eine systematische Übersichtsarbeit nur so gut sein kann, wie die Studien die ihre Basis bilden. [6, 24]

Zusammenfassend kann man sagen, dass sich die eingeschlossenen Studien zwar in grundlegenden Aspekten gleichen, sich aber in wesentlichen Details unterscheiden. Dies liegt neben dem Design, Ein- und Ausschlusskriterien, Patientenkollektiv und Symptomdetails auch an den besonderen Eigenschaften des Symptoms Rückenschmerz und der Schwierigkeit der Wissenschaft, damit umzugehen. All diese Punkte erschweren eine Zusammenfassung der Einzelergebnisse.

5.3.2. Diskussion und Interpretation der Prävalenzdaten

Die Literatursuche im Bereich Prävalenz subakuter und chronischer Rückenschmerzen ergab eine Periodenprävalenz (1 Woche) von 1.3 % für Patienten, die wegen chronischer muskuloskelettaler Beschwerden ihren Hausarzt konsultieren. Das Verhältnis von akut zu chronisch lag bei 1:1. [26]

5.95 % (2.75 % (L86), 3.2 % (L84), ICPC) der Patienten mit chronischen Erkrankungen suchten den Hausarzt wegen chronischem Rückenschmerz auf. Eine vergleichbare Studie [37], die anhand der ICD-Codes Rückenschmerzpatienten aus der Gesamtheit der Patientenkontakte identifizierte, beschrieb eine Prävalenz von 2.2 % (1 Jahr). Damit liegt sie in einem Bereich zwischen den Ergebnissen der zwei hier eingeschlossenen Studien. Dies stützt wiederum die Ergebnisse anderer Arbeiten, die zeigen, dass Rückenschmerz zu den fünf häufigsten chronischen Erkrankungen zählt, die zu Behandlungsanlässen in Hausarztpraxen führen. [37, 78] Im Rahmen dieser Arbeit wurden keine Ergebnisse für die Subgruppe der subakuten Rückenschmerzen gefunden. Stattdessen fand sich eine schwedische, ICD-9 basierte, Studie, die vermuten lässt, dass die Prävalenz von subakuten Rückenschmerz beim Hausarzt um 0.8 % liegt. [37]

Die gewonnenen Ergebnisse dieser systematischen Übersichtsarbeit zeigen auch, dass es insgesamt wenige Studien zur Prävalenz von chronischen Rückenschmerzen in der Hausarztpraxis gibt und keine Studien, die sich explizit mit der Prävalenz von subakutem Rückenschmerz beschäftigen. Da Studien zur Prognose des Rückenschmerzes regelhaft Subgruppen wie akut, subakut und chronisch bilden, stellt sich die Frage warum dies im Bereich der Prävalenz nicht getan wird. Ein Großteil der publizierten Studien beschäftigt sich nur mit der Prävalenz der Gesamtgruppe der Rückenschmerzpatienten.

Um trotz der dünnen Datenlage eine genauere Vorstellung über die Prävalenz von chronischem Rückenschmerz in der Hausarztpraxis zu erhalten, können als Anhaltspunkt die Bevölkerungsprävalenz und die Prävalenz chronischer Schmerzen herangezogen werden.

Die Bevölkerungsprävalenz chronischer Rückenschmerzen wurde in einer systematischen Übersichtsarbeit von Meucci et al. (2015) [60] untersucht. Sie fanden heraus, dass die Prävalenz zwischen 1-25 % liegt. Die überwiegende Zahl der Studien erbrachte allerdings Ergebnisse zwischen 3.9-10.2 %. Nur ein kleiner Teil der Studien zeigte Ergebnisse zwischen 13.1-20.3 %. In der älteren Bevölkerung wuchs der Anteil auf bis zu 25.4 % an, während er bei den unter 40-jährigen bei 4.2 % lag. Etwa die Hälfte der Studien fand heraus, dass Frauen etwa 50 % häufiger betroffen sind, als Männer. [60] Dies zeigt eindrucklich, welche Relevanz Rückenschmerz für Menschen weltweit hat. Des Weiteren erscheint es bei dieser hohen Prävalenz nicht verwunderlich, dass chronischer Rückenschmerz eine häufige oder sogar die häufigste Ursache für chronischen Schmerz in der Bevölkerung darstellt. [11, 23, 27, 29, 35, 67, 80] Es ist anzunehmen, dass ein großer Teil der Patienten zu irgendeinem Zeitpunkt aufgrund der Beschwerden den Hausarzt kontaktiert.

Nutzt man zudem die Häufigkeit von Schmerzkonsultation als Richtwert für die Prävalenz von Rückenschmerz in der Hausarztpraxis zeigt sich, dass dort der Rückenschmerz eine der häufigsten Diagnosen ist und die häufigste unter allen Schmerzkonsultationen. [36, 37] Auch diese Werte untermauern die Annahme, dass chronischer Rückenschmerz eine große Rolle im hausärztlichen Alltag spielt. Chronische Rückenschmerzen machen etwa 38 % der Konsultationen wegen Rückenschmerz beim Hausarzt aus. [37] Dabei liegt die Periodenprävalenz von Rückenschmerz dort zwischen 7.9-13 % (1 Woche) [8, 59], 2.8-7.1 % (1 Jahr) [5, 36, 37, 73] und 2.2-2.9 % (10 Jahre). [25, 61]

Ein wesentliches Problem der epidemiologischen Forschung im Bereich Rückenschmerz ist die hohe Variationsbreite im methodischen Aufbau, Umfang, Beobachtungszeitraum und Ursprungsland. Dies erschwert den Vergleich und die Zusammenfassung der Literatur.

Die Zusammenschau der Ergebnisse zeigt eindrucklich die Bedeutung des Symptoms und zudem die Notwendigkeit guter Prävalenzstudien im Hausarzt-Setting. Diese sollten neben der Gesamtgruppe der Rückenschmerzpatienten auch die üblichen Subgruppen analysieren, um das Wissen rund um Rückenschmerz beim Hausarzt zu komplettieren.

5.3.3. Diskussion und Interpretation der Ätiologiedaten

Die Literatursuche ergab eine Prävalenz von 23.6 % für axiale Spondyloarthritis bei Patienten mit chronischem Rückenschmerz im Alter von 20-45 Jahren. Von diesen diagnostizierten Patienten hatten 22 % die Subform Morbus Bechterew. [91] Vergleicht man diese Ergebnisse mit anderer Literatur zeigt sich, dass die Prävalenz deutlich höher liegt als die von Underwood et al. berichteten 1 %. [88] Die Ursache dafür kann darin liegen, dass die eingeschlossene Studie durch die Auswahl der Teilnehmer bereits eine Selektion in Richtung axialer Spondyloarthritis vorgenommen hat. Sowohl das gewählte Alter der Patienten als auch die Dauer der Rückenschmerzen sind Teil des Risikoprofils für diese Erkrankung.

Zwei andere Studien beschreiben hingegen eine noch höhere Prävalenz von 45.4 % und 43.7 % [10, 75]. Beide Studien haben allerdings nur Patienten mit chronischem Rückenschmerz untersucht, die bereits durch Screeningfragebögen identifiziert wurden.

Unsere Daten zeigen, dass axiale Spondyloarthritis bei etwa einem Viertel der Patienten mit chronischem Rückenschmerz unter 45 Jahren diagnostiziert werden kann. Hausärzte sollten deshalb in dieser Altersgruppe ein besonderes Augenmerk auf diese Erkrankung richten. Eine rechtzeitige Diagnose der Erkrankung kann das Behandlungsergebnis deutlich verbessern. [82]

Abgesehen von axialer Spondyloarthritis ist die aktuelle Studienlage zur Ätiologie von chronischem Rückenschmerz in der Hausarztpraxis unergiebig. Dies mag dem Umstand geschuldet sein, dass 85 % der Patienten nichtspezifischen Rückenschmerz haben, sodass eine weitreichende Diagnostik nicht zielführend ist und den aktuellen Leitlinien widerspricht. [2] Insbesondere im Sinne der Patienten sind umfangreiche Studien zur Diagnostik von chronischem Rückenschmerz schwierig umzusetzen, da strahlungsreiche oder interventionlastige Diagnostik nötig wäre, um Ergebnisse zu erhalten. Zudem zeigt sich immer wieder, dass organische Veränderungen nicht zwangsläufig mit Beschwerden einher gehen. [13]

5.3.4. Diskussion und Interpretation der Prognosedaten

5.3.4.1. Subakuter Rückenschmerz

Die drei norwegischen Studien [32, 42, 65, 85] dieser Übersichtsarbeit analysierten Rückenschmerzpatienten mit Krankschreibung. Davon zeigt eine, dass sich die Lebensqualität und Schmerz im Verlauf von 4 Monaten um etwa 10 % (Subskalen-Lebensqualität: Spannweite: -2.9 bis 23.6 (PV)) verbessert. Die absoluten Ergebnisse liegen nach 4 Monaten zwischen 0.26-0.74 (PEW, SF-36). Die positive Entwicklung im Bereich funktionelle Beeinträchtigung ist weniger deutlich. Die prozentuale Verbesserung liegt bei 6.6 % (PEW: 0.32, RDQ). Zwei der Studien [32, 42, 65] zeigen, dass nach 6 Monaten noch zwischen 55-60 % krankgeschrieben sind, nach 12 Monaten immerhin noch etwa 43 %. Im Schnitt dauert es 19 Wochen, bis die Patienten wieder arbeiten können.

Die niederländische Studie [89], berichtet mediane Ergebnisse und zeigt hier hohe prozentuale Verbesserungen in den Bereichen Lebensqualität (\emptyset Subskalen:14.7 %, Spannweite: 0-46 %) und Schmerzintensität (32 %) nach 6 Monaten.

Einschränkend muss bei den norwegischen [32, 42, 65, 85] Studien beachtet werden, dass diese in vielerlei Hinsicht eine Ausnahme darstellen und deshalb nur sehr eingeschränkt mit anderen Studien vergleichbar sind. So schmälert die Tatsache, dass nur Patienten mit einer Krankschreibung zwischen 8-12 Wochen analysiert wurden, die Repräsentativität. Inwieweit die Subgruppe der krankgeschriebenen Patienten überhaupt mit dem durchschnittlichen Patientenkollektiv zu vergleichen ist, bleibt unklar. Zudem wurden die Teilnehmer der Gruppe möglicherweise fehlklassifiziert, da man davon ausgehen muss, dass die Rückenschmerzepisode einiger Patienten bereits vor der Krankschreibung begonnen hat. Dies wiederum würde bedeuten, dass die Symptombdauer von 12 Wochen in einigen Fällen überschritten wurde und eine Einteilung in die Subgruppe subakut/chronisch (gemischt) für diese Arbeit nötig gewesen wäre.

Insgesamt muss man feststellen, dass nur zwei Studien Daten im Bereich Lebensqualität, Schmerz und funktionelle Beeinträchtigung lieferten. Dabei analysierte eine Studie arbeitsunfähige Patienten und eine Studie ein zufällig durchmischtes Patientenkollektiv, im Bezug auf dem Arbeitsaspekt. Es gibt somit nur recht wenig Prognoseergebnisse für subakute Rückenschmerzpatienten, die aufgrund ihres Beratungsanlasses in der Hausarztpraxis untersucht werden.

Doch trotz der berichteten Unterschiede zwischen den eingeschlossenen Studien, ist die Tendenz sowohl für Krankgeschriebene als auch für das durchschnittliche Patientenkollektiv mit subakutem Rückenschmerz positiv. Dies untermauert auch eine nicht eingeschlossene Studie von Karjalainen et al. [44] (Hausarzt, Beschwerdedauer 4-12 Wochen, Mittelwerte) der zeigte, dass in 6 Monaten die Schmerzintensität um etwa 20 % abnimmt, dass sich die funktionelle Beeinträchtigung um etwa 13-15 % verbessert sowie sich die Lebensqualität um etwa 3 % steigert. Im Vergleich zu der hier eingeschlossenen Arbeit, die arbeitsunfähige Patienten untersucht [85], liegt das Ergebnis also doppelt so hoch. [79] Dies könnte darin begründet sein, dass die Symptombdauer etwas kürzer war und die Patienten außerdem nicht krankgeschrieben waren. Beides spricht dafür, dass die Patienten der norwegischen Studie [85] krankheitsbedingt stärker eingeschränkt waren und deshalb schlechtere Ergebnisse erzielten.

Studien, die die Krankschreibungshäufigkeit und Dauer bei primär nicht krankgeschriebenen Patienten mit subakutem Rückenschmerz in der Hausarztpraxis untersuchen, konnten nicht gefunden werden. Sie wären im Zusammenhang mit den hier eingeschlossenen Studien interessant gewesen und hätten die Datenlage abrunden können.

Die eingeschlossenen und analysierten Studien erlauben jedoch einen Rückschluss darauf wie hoch der Effekt einer untersuchten Intervention sein muss, um den Nutzen der üblichen hausärztlichen Behandlung zu übertreffen. Diese sollten bei arbeitsunfähigen, subakuten Rückenschmerzpatienten nach 4 Monaten 10 % im Bereich Schmerz und Lebensqualität überschreiten und 6.6 % im Bereich funktionelle Beeinträchtigung. Berichten Studien ihre Ergebnisse im Median und analysieren ein durchschnittliches Patientenkollektiv, sollte man eine prozentuale Verbesserung um 15 % im Bereich Lebensqualität und um 32 % im Bereich Schmerzintensität nach 6 Monaten erwarten. Außerdem zeigte sich, dass bei hausärztlicher Behandlung arbeitsunfähige Patienten nach 6 Monaten zu 40-45 % an den Arbeitsplatz zurück gekehrt sind und zu 57 % nach 12 Monaten. [32, 42, 65, 85, 89]

5.3.4.2. Chronischer Rückenschmerz

Zusammengefasst erbringen die Studien mit chronischen Rückenschmerzpatienten folgende Ergebnisse: Man kann eine Verbesserung der funktionelle Beeinträchtigung um 5.8-6.8 % nach 6 Monaten und eine Verbesserung um 5.8-10 % nach 12 Monaten erwarten. Damit zeigt sich eine leichte Verbesserung, bei einem mittleren Beeinträchtigungsniveau zwischen 0.329-0.439 (PEW, 6 Monate).

Die Ergebnisse im Bereich Lebensqualität streuen breit, abhängig von der beobachteten Subskala und Studie. Die Tendenz über alle Bereich hinweg ist jedoch positiv. Die prozentuale Veränderung liegt zum letzten Nachbeobachtungszeitpunkt über alle Subskalen hinweg zwischen 2.2 % und 43 %. Am niedrigsten war die Lebensqualität in den Bereichen „Rollenfunktion im Bereich körperlicher Funktionen“, „Schmerz“, „Vitalität“ [76] sowie „Mobilität“ [89].

Die Stimmung der chronischen Rückenschmerzpatienten ist mit verschiedensten Instrumenten gemessen worden (shortened fear-avoidance-beliefs, worry rating, beck depression inventory II). Alle sind nicht mit einander vergleichbar. Dennoch kann man eine leicht positive Tendenz für die Parameter der geistigen Gesundheit bescheinigen und zwar mit einer prozentualen Veränderung zwischen 5.74-17 % nach 12 Monaten.

Zwei Studien [48, 93] zeigen eine positive Entwicklung im Bereich „Schmerz“, mit einer Verbesserung um 8-11 % (PV) nach 12 Monaten. Eine Studie hingegen berichtet eine leichte Verschlechterung um 2.9 % (PV) nach 6 Monaten. [76] Der absolute Wert der drei Studien liegt jedoch auf einem ähnlichen Niveau um 0.45 (6 Monate). Die medianen Ergebnisse von van den Hoogen et al. [89] erreichen nach 6 Monaten Werte von 0.1 (PEW) und liegen damit in einem sehr guten Bereich mit einer prozentualen Verbesserung von 30 %. [89]

Chronische Rückenschmerzpatienten sind im Median 6.5 Wochen (Spannweite: 0.5-48) im Jahr krank geschrieben. Etwa 1/3 der Studienteilnehmer ist von Krankschreibungen betroffen. Beobachtete man ein Kollektiv, deren Teilnehmer bereits krank geschrieben sind, fällt die Prognose deutlich schlechter aus. Jeder zweite Patient ist auch noch nach 12 Monaten krankgeschrieben. Betrachtet man den Verlauf der Rückenschmerzen, stellt man fest, dass von etwa drei Episoden pro Jahr berichtet werden, mit einer medianen Länge von 16 Wochen. 6 % der Chroniker haben nach 12 Monaten gar keine Beschwerden mehr, 10% haben sie ununterbrochen.

Ein Vergleich der Studien wird zusätzlich dadurch erschwert, dass die Gesamtdauer der Rückenbeschwerden stark variiert. Patienten mit ihrer ersten 12-wöchigen Episode werden mit Patienten verglichen, die seit Jahren oder Jahrzehnten intermittierend oder permanent

Beschwerden haben. Vergleicht man die Ergebnisse von Poole et al. [76], dessen Patienten im Schnitt mehr als zwei Jahre Beschwerden hatten, mit einer Studie, deren Teilnehmer gerade erst die 12 Wochen Marke überschritten haben [16], sieht man diese Unterschiede.

Im Bereich funktionelle Beeinträchtigung ähnelt sich zwar die prozentuale Veränderung (6.8 % (Poole) vs. 6.6 % (Costa)), die absoluten Werte der prozentualen Baseline-Werte und des letzten Nachbeobachtungszeitpunktes weichen jedoch deutlich von einander ab (0.808/0.874 vs. 0.603/0.671). Eine ähnliche Tendenz findet sich im Bereich Schmerz. Das Niveau liegt näher beisammen (0.322/0.25 vs. 0.406/0.427), dafür zeigt die prozentuale Veränderung die schlechtere Prognose für Langzeitchroniker (-2.1 %/7.2 %).

Dieser Vergleich verdeutlicht, dass solchen Patienten, die erst seit kurzem der Definition „chronischer Rückenschmerz“ entsprechen, eine bessere Prognose zukommt als denen, die bereits seit Jahren Beschwerden haben.

Andererseits analysiert eine der eingeschlossenen Studien [89] den Unterschied zwischen subakuten und chronischen Patienten unmittelbar und findet nur einen geringen Unterschied. So haben beide Gruppen nach 6 Monaten einen medianen Schmerzwert von 0.1 und viele der medianen Lebensqualität-Subskalen liegen beim besten Wert.

Die Gruppe der chronisch Kranken bleibt in der Regel inhomogen, sodass eindeutige Ergebnisse selten sind. Cedraschi et al. [14] (Chiropraktiker-/Rheumatologie-Setting) konnte zeigen, dass 46.6 % der zum Baseline-Zeitpunkt als chronisch definierten Patienten auch noch nach 3 Jahren in diese Kategorie fallen, während 53.4 % nicht mehr als chronisch zu definieren sind. 15.6 % von vormals als nicht-chronisch klassifizierten Patienten entsprechen nach 3 Jahren der Definition. [14]

Thomas et al. [87] zeigte, dass Patienten mit einer Schmerzdauer von mehr als 13 Wochen ein Odds Ratio von 5.34 haben, auch noch nach einem Jahr „andauernden“ Schmerz zu haben. (vs. Symptombdauer 4-12 Wochen: OR 1.43). Die Sinnhaftigkeit einer Subgruppenbildung zwischen subakuten und chronischen Rückenschmerzen kann daraus ersehen werden. Außerdem wird der große Einfluss der initialen Symptombdauer bei Behandlungsbeginn auf die Prognose von Rückenschmerz klar. Hierbei geht es primär um die Wahrscheinlichkeit, dass ein Patient zukünftig unter Rückenschmerzen leidet und weniger darum, wie schmerzhaft oder einschränkend diese sind.

Von praktischer Bedeutung ist auch die Frage, inwieweit sich das Kollektiv der Hausarztpatienten vom Kollektiv der Gesamtbevölkerung unterscheidet. Hier zeigte eine Studie der deutschen Bevölkerung, dass Menschen mit chronischen Rückenschmerzen einen Baseline-Wert von 0.44 (PEW) im Bereich Schmerz haben. Im Vergleich dazu zeigten Patienten, der

hier eingeschlossenen Studien, einen Baseline-Wert zwischen 0.4-0.58 (PEW).

Im Bereich Lebensqualität (SF-36, Skala 0-100) erreichte die Subskala „allgemeine Gesundheitswahrnehmung“ einen Wert von 55 im hier analysierten Hausarztkollektiv und einen Wert von 62.7 in der Gesamtbevölkerung. Die Subskala „Vitalität“ lag im Hausarztbereich bei 39.7 und in der Bevölkerung bei 53.5.

Daraus ergibt sich, dass insbesondere im Bereich Lebensqualität Hausarztpatienten stärker beeinträchtigt sind als Personen mit chronischem Rückenschmerz in der allgemeinen Bevölkerung. [27] Dies leuchtet insofern ein, als dass Patienten sich erst dann an den Arzt wenden, wenn Schmerzen und vor allem die damit verbundenen Einschränkungen ein nicht mehr tolerables Level erreicht haben.

Die detaillierte Analyse der eingeschlossenen Studien erbrachte zudem die Erkenntnis, dass bestimmte Bereiche der Lebensqualität sowohl bei subakuten Rückenschmerzpatienten als auch bei chronischen Rückenschmerzpatienten schlechter ausfallen. Gute Ergebnisse werden in den Gebieten psychisches Wohlbefinden, Rollenfunktion in emotionaler Hinsicht und soziale Funktionsfähigkeit nach 4 bzw. 6 Monaten gefunden, während die Subskalen Schmerz, Vitalität und Rollenfunktion in körperlicher Hinsicht schlecht ausfallen. [76, 85]

5.3.4.3. Subakuter und chronischer Rückenschmerz (gemischte Gruppe)

Die Zusammenschau der Ergebnisse der sechs eingeschlossenen Studien, ergibt eine leichte Verbesserung der Symptomatik für Patienten, die mit subakuten und chronischen Rückenschmerzen zum Hausarzt kommen. [31, 34, 45, 51, 52, 62, 70, 90]

Die funktionelle Beeinträchtigung verbessert sich am letzten Nachbeobachtungszeitpunkt (9 und 12 Monate) um einen Wert zwischen 6.2-25 % (PV). Die absoluten Werte verbleiben auf einem Niveau zwischen 0.288-0.418 (PEW) und 0.083 (Median).

Alle Studien bestätigen zudem eine Verbesserung der Schmerzsymptomatik. Die prozentuale Verbesserung liegt in diesem Bereich zwischen 8.4-20 %. Die prozentualen Ergebnisse bewegen sich zum Ende der Nachbeobachtungszeit im Mittel zwischen 0.351-0.51 und 0.2 (Median).

Im Bereich Lebensqualität ergibt sich eine prozentuale Verbesserung zwischen 0.9- 15.8 %, wobei die Höhe der prozentualen Ergebnisse zwischen 0.386-0.629 liegt. Im Bezug auf „fear-avoidance-beliefs“ bleibt die prozentuale Veränderung über die Zeit gering bei 4.6 % mit einem Endwert von 0.546 (PEW).

29 % der Teilnehmer brauchen im Verlauf von 12 Monaten eine Krankschreibung, was zu einem Krankenstand von etwa 10 Tagen führt. Der Verlauf der Rückenschmerzen wird von einer Studie analysiert [45, 51, 52]: Hier verbessert sich bei 31% die Beschwerden, bei 45% bleiben die Beschwerden gleich und bei 24% verschlimmert sich die Symptomatik.

Die hier analysierten Studien unterstreichen besonders eindrücklich, wie viel vom Design und den Rekrutierungskriterien abhängt. Miller et al. [62] sticht mit seinen deutlich besseren Ergebnissen im Bereich funktionelle Beeinträchtigung und Schmerz im Vergleich zu den anderen beiden Studien heraus. [52, 70] Dies erklärt sich durch den Umstand, dass Miller et al. [62] nur Patienten mit einer Schmerzdauer zwischen 10 Wochen und 6 Monaten einschloss und die Ergebnisse zudem als Median angegeben sind. Die anderen drei Studien [31, 45, 51, 52, 69, 70, 90] haben jedoch im Bezug auf die Symptombdauer nach oben keine Beschränkung, sodass auch Patienten eingeschlossen wurden, die im Mittel 13 Jahre lang Beschwerden [45, 51, 52] hatten. Dies zeigt beispielhaft, mit welcher großen Inhomogenität umgegangen werden muss, wenn man sich mit dem Patientenkollektiv der chronisch Kranken befasst.

Vergleicht man die eingeschlossenen Studien (trotz aller Inhomogenität) im Bereich subakuter und chronischer Rückenschmerz, haben beide Patientengruppen über die Zeit eine leichte Verbesserung zu erwarten. Des weiteren zeigt sich, dass die Ergebnisse der subakuten, chronischen und gemischt subakut/chronischen Rückenschmerzpatienten in einem ähnlichem Spektrum streuen. Dabei fallen die Ergebnisse im Bereich Schmerz für die Gruppe der subakuten Rückenschmerzpatienten leicht besser aus als die der Chroniker. Diese leichte Tendenz zu Gunsten der subakuten Rückenschmerzpatienten findet sich bezüglich der funktionellen Beeinträchtigung nicht mehr. [31, 70, 80, 45, 48, 51, 52, 76, 85, 93] Schwierig einzuschätzen bleibt jedoch, wie relevant die gemessenen Verbesserungen für den Alltag des Patienten sind und ab welcher numerischen Veränderung in den Messinstrumenten die Verbesserung subjektiv wahrgenommen wird. Bereits zuvor stellte Korff et al. zu diesem Thema fest, dass allein das Fortbestehen von Schmerzen nicht immer mit einem schlechten Ergebnis für den Patienten gleichzusetzen ist. [49] Ab welchen Werten in Sinne des Patienten von einem schlechten oder guten Ergebnis gesprochen werden kann, ist nicht einfach zu beantworten. Dennoch sollen im Folgenden Lösungsansätze präsentiert werden.

Einige Arbeiten befassen sich mit der „minimalen klinisch relevanten Veränderung“ („minimale important change“) und legen den Wert in einem Bereich um die 30 % des Ausgangswertes der jeweiligen Messinstrumente für Schmerz und funktionelle Beeinträchtigung fest.

Als absolute Werte wurden festgesetzt: VAS (0-100) 15, numerische Skala (0-10) 2, RDQ (0-24) 5, ODI (0-100) 10. [72] Verwendet man diese Richtlinien exemplarisch für die Studien der chronischen Rückenschmerzpatienten in der vorliegenden Arbeit, so wurde nur bei 1/7 der Prognoseergebnisse ein „minimale important change“ erreicht. Auch wenn der „minimale important change“ klassischerweise nicht für Beobachtungsstudien, sondern für Therapiestudien eingesetzt wird, gibt er einen ersten Anhaltspunkt dafür, wie bedeutend die gemessenen Verbesserungen für die Prognose der Patienten sein könnten.

Einen anderen Weg nutzt die subjektive Bewertung der Patienten für die Bewertung des Ergebnisses. Ein solches anker-basiertes System ist das PASS (Patient acceptable symptomatic state). PASS beschreibt das höchste Ergebnis, unterhalb dessen die Patienten zufrieden sind (diese Beschreibung gilt für Skalen, deren höchster Wert das schlechteste Ergebnis widerspiegelt). Dieser Wert ist im Vergleich zum „minimale important change“ unabhängig vom Baseline-Wert. Er zeigt nur geringe Abweichungen im Bezug auf Alter, Symptombdauer und Geschlecht der Befragten. [74] Es folgen einige Beispielwerte aus der Literatur:

Für alle chronische Erkrankungen des rheumatischen Formenkreises wurde ein ähnlicher Wert um die 30-40 mm (VAS) als „Cutoff-Wert“ beschrieben. [74] Eine marokkanische Studie berichtete über Patienten mit rheumatologischen Erkrankungen, u.a. auch degenerativen Rückenschmerzen, einen „Cutoff-Wert“ von 50mm (VAS) [95]. Eine multinationale Studie [89a] berichtete einen „Cutoff-Wert“ von 42mm (VAS) im Bereich Schmerz und 43mm (VAS) für funktionelle Beeinträchtigung in einer Gruppe aus Patienten mit unterschiedlichen rheumatoiden Erkrankungen, u.a. auch degenerativen Rückenschmerzen.

Im Rahmen der hier vorliegenden Arbeit unterschritt keine der eingeschlossenen Studien, die chronische Rückenschmerzpatienten untersuchten, diese beispielhaften „Cutoff-Werte“. Dies legt nahe, dass der überwiegende Teil der Patienten keine für sie zufriedenstellende Ergebnisse erzielte.

Obwohl es insbesondere in Bereichen, in denen vorwiegend symptomatisch behandelt wird, sinnvoll erscheint sogenannte PROs (patient-reported outcome measures) einzusetzen, fand dies in keiner der hier analysierten Studien statt. Dabei sollte gerade bei chronischen Krankheitsbildern der Patientenzufriedenheit eine besondere Aufmerksamkeit geschenkt werden. [74]

In diesem Zusammenhang konnte Hush et al. [41], in einer systematischen Übersichtsarbeit zeigen, dass in der subjektiven Beurteilung der Beschwerden insbesondere Aspekte der Lebensqualität ausschlaggebend sind. Zu diesen Aspekten gehörten: Haushalt, Beruf, Fähigkeit zur sozialen Teilhabe, emotionales Wohlbefinden und einige mehr. Auch die Schmerzstärke, weniger die Schmerzfreiheit sind wichtig. [41] Darüber hinaus gilt es zu beachten, dass Genesung keine statische Größe ist, sondern entsprechend dem periodischen Charakter von Rückenschmerz, über die Zeit fluktuiert.

Für die Zukunft der Rückenschmerzforschung wäre es hilfreich Daten zur Hand zu haben, die eine Interpretation und Bewertung der Ergebnisse der jeweiligen Messinstrumente erlauben. Bisher fällt diese im Sinne einer Kategorisierung in „gut“, „mittel“ oder „schlecht“ eher schwer. Ebenso wäre ein intensivierter Einsatz von PROs interessant, um den subjektiven Erfolg bei den teilnehmenden Patienten besser einschätzen zu können.

5.3.4.4. Rückenschmerz mit Ausstrahlung in die Beine

Zwei Studien differenzierten ihre Teilnehmer nach Ausstrahlung der Schmerzen. [69, 86] In diesen zeigt sich besonders eindrucksvoll, dass Patienten mit Schmerzen unterhalb der Knie (RS+UK) eine schlechtere Prognose haben als solche, die nur Schmerzen im Rücken (RS) oder zusätzlich ausstrahlende Schmerzen in die Oberschenkel haben (RS+OK).

In den Subgruppen „RS“ und „RS+OK“ verringern sich die Schmerzen um 21 % (PV) und 32 % (PV). Während sich die Gruppe „RS+UK“ nur um 14 % (PV) verbessert. Das gleiche Bild zeichnet sich im Bereich funktionelle Beeinträchtigung ab. Die Gruppen verbessern sich um 14 % (RS+UK), 18,9 % (RS) und 23 % (RS+OK).

Kein Einfluss zeigt sich jedoch auf die Rückkehr zur Arbeit. 46,2 % arbeiten nach 12 Monaten wieder in der Gruppe mit Ausstrahlung, während es 48 % in der Gruppe ohne Ausstrahlung sind.

Zu diesem Ergebnis kamen bereits viele andere Studien, die Rückenschmerz in anderen Settings oder Rückenschmerz als Diagnose untersucht haben. Unter anderem eine systematische Übersichtsarbeit, die die Auswirkung von Rückenschmerz mit Ausstrahlung in die Beine in verschiedenen Settings untersucht hat. Sie kommt zu dem Schluss, dass Patienten mit Schmerzausstrahlung schlechtere Ausgangswerte im Bereich Schmerz, funktionelle Beeinträchtigung und Subbereichen der Lebensqualität haben. Auch die Verbesserung der Symptomatik über die Zeit ist geringer. [47]

Verbunden mit der schlechteren Prognose für Patienten mit Ausstrahlung in die Beine ist auch eine längere „Recovery“-Dauer (Zeitraum bis zur Gesundung), die Andersson et al. (1999) untersucht hat. Er berichtet, dass 60 % der Patienten ohne Ausstrahlungssymptomatik innerhalb von 10 Tagen wiederhergestellt („recovered“) sind, während es in der Gruppe mit Ausstrahlungssymptomatik nur 40 % sind. [4] Ebenso ist bei diesen Patienten das Risiko erhöht, dauerhaft funktionell einschränkende Rückenschmerzen zu entwickeln. Dies zeigt sich in einem Odds Ratio von 1.89 (0.8 bis 4.4). Im Vergleich dazu haben Patienten ohne Ausstrahlungssymptomatik ein Odds Ratio von 1 und damit kein erhöhtes Risiko. [87] Dies wirkt sich nicht nur auf persönlicher Ebene für den Patienten aus, sondern hat auch Konsequenzen für die Volkswirtschaft. Die Inzidenz von Krankengeldern („sickness benefits“) liegt bei 89,6 pro 1000-Personenjahre versus 60,4 per 1000-Personenjahre bei Patienten ohne ausstrahlende Schmerzen. Dies zeigte jedoch keine Auswirkung auf die Dauer der Krankschreibungen. [50]

Der Vergleich der eingeschlossenen Studien mit ähnlichen Studien aus anderen Settings und Patienten unterschiedlicher Symptombdauer lassen den Schluss zu, dass ausstrahlende Schmerzen über alle anderen Faktoren hinweg die Prognose der Patienten verschlechtern.

5.4. Fazit

Subakuter und chronischer Rückenschmerz ist ein häufiges Symptom in der Bevölkerung und ebenso ein häufiges Symptom in der Hausarztpraxis. Insbesondere beim Gedanken an den demografischen Wandel und den Umstand, dass chronischer Rückenschmerz mit zunehmendem Altern immer häufiger wird, wird seine Bedeutung in Zukunft eher noch zunehmen. Diese Arbeit konnte zeigen, dass Patienten mit subakutem und chronischem Rückenschmerz, die beim Hausarzt in Behandlung sind, im Verlauf eines Jahres im den Bereichen Schmerz, funktionelle Beeinträchtigung, Lebensqualität und Stimmung eine Verbesserung ihrer Beschwerden zu erwarten haben. Das Ausmaß der Verbesserung variiert jedoch stark von Studie zu Studie. Zudem bleibt es schwierig das individuelle Ausmaß der subjektiven Verbesserungen der Patienten einzuschätzen. Studien mit arbeitsunfähigen chronischen Patienten zeigten, dass die arbeitsbezogene Prognose weniger positiv ist, da nach 1 Jahr noch immer etwa 40% arbeitsunfähig sind. Strategien diese hohen Ausfälle in Zukunft zu verringern, sind im Sinne der Patienten und aus volkswirtschaftlicher Sicht enorm wichtig. Deutlich wurde auch, dass Patienten mit ausstrahlenden Schmerzen bis unterhalb der Knie eine schlechtere Prognose haben.

Im Bereich Ätiologie, konnten die Daten zeigen, dass Hausärzte bei jungen Patienten mit Rückenschmerzen, an das Vorliegen einer Spondyloarthritis denken sollten. Dennoch muss man sagen, dass die Datenlage im Bereich Ätiologie und Prävalenz sehr dünn ist. Es wäre wünschenswert, dass zukünftige Studien diese Aspekte genauer beleuchten. Insgesamt fällt eine große Inhomogenität zwischen den Studien auf, insbesondere im Bezug auf Ein- und Ausschlusskriterien, Studiendesign und Verwendung von Messinstrumenten. Es ist unerlässlich, die Erforschung in so vielen Aspekten wie möglich zu standardisieren, um zusammenfassende Betrachtungen der Literatur zu ermöglichen.

6. Literaturverzeichnis:

1. (2015) Nurse practitioner (NP) : MedlinePlus Medical Encyclopedia.
<http://www.nlm.nih.gov/medlineplus/ency/article/001934.htm> (15.03.2015).
2. Aerztliches Zentrum für Qualität in der Medizin (2013) Nationale VersorgungsLeitlinie Kreuzschmerz. Langfassung (12.03.2015).
3. Airaksinen O, Brox J I, Cedraschi C, Hildebrandt J, Klüber-Moffett J, Kovacs F, Mannion A F, Reis S, Staal J B, Ursin H, Zanolini G (2006) Chapter 4 European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J* 15: s192.
4. Andersson G B J (1999) Epidemiological features of chronic low-back pain. *The Lancet* 354: 581–585.
5. Andersson H I, Ejlertsson G, Leden I, Schersten B (1999) Musculoskeletal chronic pain in general practice. Studies of health care utilisation in comparison with pain prevalence. *Scand J Prim Health Care* 17: 87–92.
6. Antje Timmer B R (2008) Systematische Übersichtsarbeiten zu Fragen der Therapie und Prävention Eine Einführung in Frage und Antwort. Teil 2 – Was macht eine gute Übersichtsarbeit aus? *Arzneimitteltherapie* 26: 252–255.
7. Arbeitsgruppe Informatics der Schweizerischen Gesellschaft für Allgemeinmedizin (2008) ICPC-2 Deutsch/German International Classification of Primary Care – 2nd Edition (15.03.2015).
8. Becker A, Kögel K, Donner Banzhoff N, Basler H D, Chenot J.F, Maitra R, Kochen M M (2003) Low back pain patients in general practice: Complaints, therapy expectations and care information. *Z Allgemeinmed*: 126–131.
9. Brage S, Bentsen B G, Bjerkedal T, Nygard J F, Tellnes G (1996) ICPC as a standard classification in Norway. *Fam Pract* 13: 391–396.
10. Brandt H C, Spiller I, Song I-H, Vahldiek J L, Rudwaleit M, Sieper J (2007) Performance of referral recommendations in patients with chronic back pain and suspected axial spondyloarthritis. *Ann Rheum Dis* 66: 1479–1484.
11. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D (2006) Survey of chronic pain in Europe: Prevalence, impact on daily life, and treatment (14.10.2015).
12. Britt H, Angelis M, Harris E (1998) The reliability and validity of doctor-recorded morbidity data in active data collection systems. *Scand J Prim Health Care* 16: 50–55.
13. Bundesärztekammer (BÄK) K B (A d W M F (Nationale VersorgungsLeitlinie Kreuzschmerz – Langfassung. Version 4. 2010. zuletzt verändert: August 2013. Available from:
<http://www.kreuzschmerz.versorgungsleitlinien.de>.
14. Cedraschi C, Robert J, Goerg D, Perrin E, Fischer W, Vischer T L (1999) Is chronic non-specific low back pain chronic? Definitions of a problem and problems of a definition. *Br J Gen Pract* 49: 358–362.
15. Chmiel C, Bhend H, Senn O, Zoller M, Rosemann T (2011) The FIRE project: a milestone for research in primary care in Switzerland. *Swiss Med Wkly* 140: w13142.

16. Costa L d C M, Maher C G, McAuley J H, Hancock M J, Herbert R D, Refshauge K M, Henschke N (2009) Prognosis for patients with chronic low back pain: inception cohort study. *BMJ* 339: b3829.
17. Dagenais S, Caro J, Haldeman S (2008) A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine J* 8: 8–20.
18. Depont F, Hunsche E, Abouelfath A, Diatta T, Addra I, Grelaud A, Lagnaoui R, Molimard M, Moore N (2010) Medical and non-medical direct costs of chronic low back pain in patients consulting primary care physicians in France. *Fundam Clin Pharmacol* 24: 101–108.
19. Dey P, Simpson C W R, Collins S I, Hodgson G, Dowrick C F, Simison A J M, Rose M J (2004) Implementation of RCGP guidelines for acute low back pain: a cluster randomised controlled trial. *Br J Gen Pract* 54: 33–37.
20. Donner-Banzhoff N, Kunz R, Rosser W (2001) Studies of symptoms in primary care. *Fam Pract* 18: 33–38.
- 20a. Dornieden, K. (2014) Kopfschmerz in der Primärversorgung – eine systematische Übersichtsarbeit symptomevaluierender Studie
21. Doust J (2009) Diagnosis in General Practice. Using probabilistic reasoning. *BMJ* 339: b3823.
22. Ekman M, Jonhagen S, Hunsche E, Jonsson L (2005) Burden of illness of chronic low back pain in Sweden: a cross-sectional, retrospective study in primary care setting. *Spine (Phila Pa 1976)* 30: 1777–1785.
23. Elliott A M, Smith B H, Penny K I, Smith W C, Chambers W A (1999) The epidemiology of chronic pain in the community. *Lancet* 354: 1248–1252.
24. Erler A, Beyer M, Muth C, Gerlach F M, Brennecke R (2009) Garbage in - garbage out? Validity of coded diagnoses from GP claims records. *Gesundheitswesen* 71: 823–831.
25. Fink W, Haidinger G (2007) Prevalence of Health Problems in a Family Practice Observed Over 10 Years. *Z Allg Med* 83: 102–108.
26. Frolund F, Frolund C (1986) Pain in general practice. Pain as a cause of patient-doctor contact. *Scand J Prim Health Care* 4: 97–100.
27. Gerhardt A, Hartmann M, Blumenstiel K, Tesarz J, Eich W (2014) The prevalence rate and the role of the spatial extent of pain in nonspecific chronic back pain--a population-based study in the south-west of Germany. *Pain Med* 15: 1200–1210.
28. Greenhalgh T (1997) Papers that summarise other papers (systematic reviews and meta-analyses). *BMJ* 315: 672–675.
29. Gureje O, Simon G E, Korff M von (2001) A cross-national study of the course of persistent pain in primary care. *Pain*: 195–200.
30. Haasenritter J, Viniol A, Becker A, Bosner S, Hullermeier E, Senge R, Donner-Banzhoff N (2013) Diagnosis in context - broadening the perspective. *Z Evid Fortbild Qual Gesundhwes* 107: 585–591.
31. Haas M, Sharma R, Stano M (2005) Cost-effectiveness of medical and chiropractic care for acute and chronic low back pain. *J Manipulative Physiol Ther* 28: 555–563.
32. Hagen E M, Eriksen H R, Ursin H (2000) Does early intervention with a light mobilization

- program reduce long-term sick leave for low back pain? *Spine (Phila Pa 1976)* 25: 1973–1976.
33. Hagen E M, Svensen E, Eriksen H R (2005) Predictors and modifiers of treatment effect influencing sick leave in subacute low back pain patients. *Spine (Phila Pa 1976)* 30: 2717–2723.
34. Haldorsen, Kronholm, Skouen, Ursin (1998) Predictors for outcome of a multi-modal cognitive behavioural treatment program for low back pain patients-a 12-month follow-up study. *Eur J Pain* 2: 293–307.
35. Hardt J, Jacobsen C, Goldberg J, Nickel R, Buchwald D (2008) Prevalence of chronic pain in a representative sample in the United States. *Pain Med* 9: 803–812.
36. Hart L G, Deyo R A, Cherkin D C (1995) Physician office visits for low back pain. Frequency, clinical evaluation, and treatment patterns from a U.S. national survey. *Spine (Phila Pa 1976)* 20: 11–19.
37. Hasselstrom J, Liu-Palmgren J, Rasjo-Wraak G (2002) Prevalence of pain in general practice. *Eur J Pain* 6: 375–385.
38. Hestbaek L, Leboeuf-Yde C, Engberg M, Lauritzen T, Bruun N H, Manniche C (2003) The course of low back pain in a general population. Results from a 5-year prospective study. *J Manipulative Physiol Ther* 26: 213–219.
39. Higgins JPT G S ((2011) *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2001].
40. Hillman M, Wright A, Rajaratnam G, Tennant A, Chamberlain M A (1996) Prevalence of low back pain in the community: implications for service provision in Bradford, UK. *J Epidemiol Community Health* 50: 347–352.
41. Hush J M, Kamper S J, Stanton T R, Ostelo R, Refshauge K M (2012) Standardized measurement of recovery from nonspecific back pain. *Arch Phys Med Rehabil* 93: 849–855.
42. Indahl A, Velund L, Reikeraas O (1995) Good prognosis for low back pain when left untampered. A randomized clinical trial. *Spine* 20: 473–477.
43. Jüni P, Holenstein F, Sterne J, Bartlett C, Egger M (2002) Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *Int J Epidemiol* 31: 115–123.
44. Karjalainen K, Malmivaara A, Mutanen P, Roine R, Hurri H, Pohjolainen T (2004) Mini-intervention for subacute low back pain: two-year follow-up and modifiers of effectiveness. *Spine (Phila Pa 1976)* 29: 1069–1076.
45. Knox C R, Lall R, Hansen Z, Lamb S E (2014) Treatment compliance and effectiveness of a cognitive behavioural intervention for low back pain: a complier average causal effect approach to the BeST data set. *BMC Musculoskelet Disord* 15: 17.
46. Kohlmann T, Bullinger M, Kirchberger-Blumstein I (1997) Die deutsche Version des Nottingham Health Profile (NHP): Übersetzungsmethodik und psychometrische Validierung. *Soz Präventivmed* 42: 175–185.
47. Konstantinou K, Hider S L, Jordan J L, Lewis M, Dunn K M, Hay E M (2013) The impact of low back-related leg pain on outcomes as compared with low back pain alone: a systematic review of the literature. *Clin J Pain* 29: 644–654.
48. Korff M von, Balderson B H K, Saunders K, Miglioretti D L, Lin E H B, Berry S, Moore J E, Turner J A (2005) A trial of an activating intervention for chronic back pain in primary care and

physical therapy settings. *Pain* 113: 323–330.

49. Korff M von, Saunders K (1996) The course of back pain in primary care. *Spine (Phila Pa 1976)* 21: 2833-7; discussion 2838-9.
50. Kristman V L, Hartvigsen J, Leboeuf-Yde C, Kyvik K O, Cassidy J D (2012) Does Radiating Spinal Pain Determine Future Work Disability? A Retrospective Cohort Study of 22,952 Danish Twins (14.10.2015).
51. Lamb S E, Hansen Z, Lall R, Castelnuovo E, Withers E J, Nichols V, Potter R, Underwood M R (2010) Group cognitive behavioural treatment for low-back pain in primary care: a randomised controlled trial and cost-effectiveness analysis. *Lancet* 375: 916–923.
52. Lamb S E, Lall R, Hansen Z, Castelnuovo E, Withers E J, Nichols V, Griffiths F, Potter R, Szczepura A, Underwood M (2010) A multicentred randomised controlled trial of a primary care-based cognitive behavioural programme for low back pain. The Back Skills Training (BeST) trial. *Health Technol Assess* 14: 1-253, iii-iv.
53. Lamb S E, Mistry D, Lall R, Hansen Z, Evans D, Withers E J, Underwood M R (2012) Group cognitive behavioural interventions for low back pain in primary care: extended follow-up of the Back Skills Training Trial (ISRCTN54717854). *Pain* 153: 494–501.
54. Laux G, Koerner T, Rosemann T, Beyer M, Gilbert K, Szecsenyi J (2005) The CONTENT project: a problem-oriented, episode-based electronic patient record in primary care. *Inform Prim Care* 13: 249–255.
55. Liberati A, Altman D G, Tetzlaff J, Mulrow C, Gotzsche P C, Ioannidis J P A, Clarke M, Devereaux P J, Kleijnen J, Moher D (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 339: b2700.
56. Lie S A, Eriksen H R, Ursin H, Hagen E M (2008) A multi-state model for sick-leave data applied to a randomized control trial study of low back pain. *Scand J Public Health* 36: 279–283.
57. Lindstrom I, Ohlund C, Eek C, Wallin L, Peterson L E, Fordyce W E, Nachemson A L (1992) The effect of graded activity on patients with subacute low back pain: a randomized prospective clinical study with an operant-conditioning behavioral approach. *Phys Ther* 72: 279-90; discussion 291-3.
58. Manchikanti L, Singh V, Datta S, Cohen S P, Hirsch J A (2009) Comprehensive review of epidemiology, scope, and impact of spinal pain. *Pain Physician* 12: E35-70.
59. Mantyselka P, Kumpusalo E, Ahonen R, Kumpusalo A, Kauhanen J, Viinamaki H, Halonen P, Takala J (2001) Pain as a reason to visit the doctor: a study in Finnish primary health care. *Pain* 89: 175–180.
60. Meucci R D, Fassa A G, Faria N M X (2015) Prevalence of chronic low back pain: systematic review. *Rev Saude Publica* 49: 1–10.
61. Michaleff Z A, Harrison C, Britt H, Lin C-W C, Maher C G (2012) Ten-year survey reveals differences in GP management of neck and back pain. *Eur Spine J* 21: 1283–1289.
62. Miller P, Kendrick D, Bentley E, Fielding K (2002) Cost-effectiveness of lumbar spine radiography in primary care patients with low back pain. *Spine (Phila Pa 1976)* 27: 2291–2297.
63. Moher D, L A T J A D (2011) Bevorzugte Report Items für systematische Übersichten und Meta-Analysen: Das PRISMA-Statement. *Dtsch med Wochenschr*: e9–e15.

64. Moher D, Pham B, Lawson M L, Klassen T P (2003) The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technol Assess* 7: 1–90.
65. Molde Hagen E, Grasdal A, Eriksen H R (2003) Does early intervention with a light mobilization program reduce long-term sick leave for low back pain: a 3-year follow-up study. *Spine (Phila Pa 1976)* 28: 2309-15; discussion 2316.
66. Monika Bullinger I K J W (1995) Der deutsche SF-36 Health Survey. (Übersetzung und psychometrische Testung eines krankheitstibergreifenden Instruments zur Erfassung der gesundheitsbezogenen Lebensqualität. *Zeitschrift für Gesundheitswissenschaften*: 21–36.
67. Neuhauser H, Ellert U, Ziese T (2005) Chronic back pain in the general population in Germany 2002/2003: prevalence and highly affected population groups. *Gesundheitswesen* 67: 685–693.
68. Nyiendo J, Haas M, Goldberg B, Lloyd C (2001) A descriptive study of medical and chiropractic patients with chronic low back pain and sciatica: management by physicians (practice activities) and patients (self-management). *J Manipulative Physiol Ther* 24: 543–551.
69. Nyiendo J, Haas M, Goldberg B, Sexton G (2001) Pain, disability, and satisfaction outcomes and predictors of outcomes: a practice-based study of chronic low back pain patients attending primary care and chiropractic physicians. *J Manipulative Physiol Ther* 24: 433–439.
70. Nyiendo J, Haas M, Goldberg B, Sexton G (2001) Patient characteristics and physicians' practice activities for patients with chronic low back pain: a practice-based study of primary care and chiropractic physicians. *J Manipulative Physiol Ther* 24: 92–100.
71. O'Halloran J, Miller G C, Britt H (2004) Defining chronic conditions for primary care with ICPC-2. *Fam Pract* 21: 381–386.
72. Ostelo R W J G, Deyo R A, Stratford P, Waddell G, Croft P, Korff M von, Bouter L M, Vet H C de (2008) Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976)* 33: 90–94.
73. Papageorgiou A C, Croft P R, Ferry S, Jayson M I, Silman A J (1995) Estimating the prevalence of low back pain in the general population. Evidence from the South Manchester Back Pain Survey. *Spine (Phila Pa 1976)* 20: 1889–1894.
74. Pham T, Tubach F (2009) Patient acceptable symptomatic state (PASS). *Joint Bone Spine* 76: 321–323.
75. Poddubnyy D, Brandt H, Vahldiek J, Spiller I, Song I-H, Rudwaleit M, Sieper J (2012) The frequency of non-radiographic axial spondyloarthritis in relation to symptom duration in patients referred because of chronic back pain: results from the Berlin early spondyloarthritis clinic. *Ann Rheum Dis*.
76. Poole H, Glenn S, Murphy P (2007) A randomised controlled study of reflexology for the management of chronic low back pain. *Eur J Pain* 11: 878–887.
77. Ressing M, Blettner M, Klug S J (2009) Systematic literature reviews and meta-analyses: part 6 of a series on evaluation of scientific publications. *Dtsch Arztebl Int* 106: 456–463.
78. Rizza A, Kaplan V, Senn O, Rosemann T, Bhend H, Tandjung R (2012) Age- and gender-related prevalence of multimorbidity in primary care: the Swiss FIRE project. *BMC Fam Pract* 13: 113.
79. Rossignol M, Allaert F A, Rozenberg S, Valat J P, Avouac B, Peres G, Le Teuff G, Bannwarth B (2005) Measuring the contribution of pharmacological treatment to advice to stay active in patients with subacute low-back pain: a randomised controlled trial. *Pharmacoepidemiol Drug Saf* 14: 861–

867.

80. Rustøen T, Wahl A K, Hanestad B R, Lerdal A, Paul S, Miaskowski C (2004) Prevalence and characteristics of chronic pain in the general Norwegian population (14.10.2015).
81. Schmidt C O, Kohlmann T (2005) What do we know about the symptoms of back pain? Epidemiological results on prevalence, incidence, progression and risk factors. *Z Orthop Ihre Grenzgeb* 143: 292–298.
82. Sieper J (2012) How to screen for axial spondyloarthritis in primary care? *Curr Opin Rheumatol* 24: 359–362.
83. Smith M, Davis M A, Stano M, Whedon J M (2013) Aging baby boomers and the rising cost of chronic back pain: secular trend analysis of longitudinal Medical Expenditures Panel Survey data for years 2000 to 2007. *J Manipulative Physiol Ther* 36: 2–11.
84. Somerville S, Hay E, Lewis M, Barber J, van der Windt D, Hill J, Sowden G (2008) Content and outcome of usual primary care for back pain: a systematic review. *Br J Gen Pract* 58: 790-7, i-vi.
- 84a. Stadjé, R. (2014) Müdigkeit in der Primärversorgung : eine systematische Übersichtsarbeit
85. Storheim K, Brox J I, Holm I, Koller A K, Bo K (2003) Intensive group training versus cognitive intervention in sub-acute low back pain: short-term results of a single-blind randomized controlled trial. *J Rehabil Med* 35: 132–140.
86. Storro S, Moen J, Svebak S (2004) Effects on sick-leave of a multidisciplinary rehabilitation programme for chronic low back, neck or shoulder pain: comparison with usual treatment. *J Rehabil Med* 36: 12–16.
87. Thomas E, Silman A J, Croft P R, Papageorgiou A C, Jayson M I, Macfarlane G J (1999) Predicting who develops chronic low back pain in primary care: a prospective study. *BMJ* 318: 1662–1667.
- 87a. Tubach F, Ravaud P, Martin-Mola E, Awada H, Bellamy N, Bombardier C, Felson D T, Hajjaj-Hassouni N, Hochberg M, Logeart I, Matucci-Cerinic M, van de Laar M, van der Heijde D, Dougados M (2012) Minimum clinically important improvement and patient acceptable symptom state in pain and function in rheumatoid arthritis, ankylosing spondylitis, chronic back pain, hand osteoarthritis, and hip and knee osteoarthritis: Results from a prospective multinational study. *Arthritis Care Res (Hoboken)* 64: 1699–1707.
88. Underwood M R, Dawes P (1995) Inflammatory back pain in primary care. *Br J Rheumatol* 34: 1074–1077.
89. van den Hoogen H J, Koes B W, van Eijk J T, Bouter L M, Deville W (1997) Pain and health status of primary care patients with low back pain. *J Fam Pract* 44: 187–192.
90. van den Hoogen H J, Koes B W, van Eijk J T, Bouter L M, Deville W (1998) On the course of low back pain in general practice: a one year follow up study. *Ann Rheum Dis* 57: 13–19.
91. van Hoeven L, Luime J, Han H, Vergouwe Y, Weel A (2014) Identifying axial spondyloarthritis in Dutch primary care patients, ages 20-45 years, with chronic low back pain. *Arthritis Care Res (Hoboken)* 66: 446–453.
92. van Tulder M W, Koes B W, Bouter L M, Metsemakers J F (1997) Management of chronic nonspecific low back pain in primary care: a descriptive study. *Spine (Phila Pa 1976)* 22: 76–82.
93. van Tulder M W, Koes B W, Metsemakers J F, Bouter L M (1998) Chronic low back pain in primary care: a prospective study on the management and course. *Fam Pract* 15: 126–132.

94. van Tulder M, Koes B, Bombardier C (2002) Low back pain. Best Pract Res Clin Rheumatol 16: 761–775.
95. Wariaghli G, Allali F, Berrada K, Idrissi Z, Hmamouchi I, Abouqal R, Hajjaj-Hassouni N (2013) The patient acceptable symptom state of chronic musculoskeletal pain measured on a visual analog scale in Moroccan patients. Pain Med 14: 103–109.
96. Ziegler A, Lange S, Bender R (2007) Systematische Übersichten und Meta-Analyse. Artikel Nr. 19 der Statistik-Serie in der DMW. Dtsch Med Wochenschr 132 Suppl 1: e48-52.
97. Zentrales Deutsches Institut für medizinische Dokumentation und Information Medline.
<https://www.dimdi.de/static/de/db/dbinfo/me66.htm> (25.12.2015)

7. Anlagen

Bias-Tool

Domain A: Selection of patients and GPs (refers to all studies regardless the review question)		
Item	Refers to	
I1	<p>Was the symptom to be investigated clearly described? (Yes, No)</p> <p>Yes(2): Study states information about sciatica/radiation into the leg AND length of symptoms AND location of the symptoms (e.g. "low" back pain)</p> <p>No(0): Study does not state these information</p>	Information
I2	<p>Were the selection criteria of the patients clearly described?(Yes, No)</p> <p>Yes(2): It is obvious that all patients with the symptom were recruited OR the selection criteria were clearly stated (mean age, percentage of men/women, red flags)</p> <p>No(0): Study does not state these information</p>	Information
I3	<p>Was a consecutive or random sample of patients enrolled? (Yes, Unclear, No)</p> <p>Yes(2): Explicitly mentioned "all patients"/"consecutively", all patients recruited through databank search</p> <p>Unclear(1)</p> <p>No(0): Explicitly mentioned (convenience sample, ad-hoc-sample)</p>	Bias
I4	<p>Was it a multi-centre study? (Yes, Unclear, No)</p> <p>Yes(2): At least 3 greater areas within 1 country OR Several practices in one greater area (country, region)</p> <p>Unclear(1)</p> <p>No(0): One practice, institution, several practices within one city</p>	Bias
I5	<p>Did the selection criteria of the patients permit the study population to represent the full spectrum of those presenting with the symptom in the respective setting? (Yes, Unclear, No)</p> <p>Yes(2): Study population represents the whole spectrum of patients with acute, subacute or chronic back pain; No restrictions mentioned OR restrictions probably not relevant; Needs to be included: sciatica/radiation in to the legs; Restrictions that were accepted: age restrictions (e.g. 18-65), red flags excluded</p> <p>Unclear(1): Possible relevant restrictions; sciatica/radiation into the legs not stated</p> <p>No(0):Definite relevant restrictions e.g. only patients on sick-leave, only patients with severe symptoms</p>	Variation
I6	<p>Were the participating health care professionals/ institutions representative for the setting to be investigated in the review?</p> <p>Yes(2): No specialisation mentioned</p> <p>Unclear(1): Small amount (<50%) of OR no specialized GPs, other health care professionals (nurse practitioner), out-of-hour service</p>	Variation

	No(0): Substantial amount (>50%) specialised GP, other health care professionals, out-of-hour service	
	<p><u>Concern that the selection of patients introduced substantial variation: low, unclear, high</u></p> <p>Low(0): Consider low if I5 and I6 are answered with yes Unclear(1): Consider unclear if I5 and I6 are answered as follows: yes/unclear, unclear/unclear, yes/no High(2): Consider high if I5 or I6 are answered as follows: no/no, no/unclear</p>	Overall assessment 1
	<p><u>Concern that the selection of patients introduced substantial bias: low, unclear, high</u></p> <p>Low(0): Consider low if I3 and I4 are answered with yes Unclear(1): Consider unclear if I3 and I4 are answered as follows: yes/unclear, unclear/unclear, yes/no High(2): Consider high if I3 or I4 are answered as follows: no/no, no/unclear</p>	Overall assessment 2
Domain B: Data collection and patient flow (refers to all studies regardless of the review question)		
Item	Refers to	
I7	<p>Were data about the symptom and the inclusion criteria collected directly from the patients (as opposed to a proxy like a register, routine documentation)</p> <p>Yes(2): Data were collected during a prospective primary study Unclear(1) No(0): Data were collected from a register, routine data</p>	Bias
I8	<p>Was the same mode of data collection used for all patients? (Yes(2),Unclear(1),No(0))</p>	Bias
I9	<p>Was the number of non-responders/ dropouts unlikely to affect the results?</p> <p>Yes(2): <20% Unclear(1) No(0):>20%</p>	Bias
	<p><u>Risk that the mode of data collection and/ or patient flow introduced bias: low, unclear, high</u></p> <p>Low(0): Consider low if I7,I8 and I9 are answered as following: yes/yes/yes, yes/yes/unclear Unclear(1): Consider unclear if I7,I8 and I9 are answered as follows: unclear/unclear/unclear,yes/yes/no, unclear/unclear/no High(2): Consider high if I7,I8 or I9 are answered as follows: no/no/ no, no/no/yes, no/no/unclear</p>	Overall assessment 3
Domain C: Determination of the underlying aetiology of the symptom (refers only to review question "What are the underlying conditions and their respective frequencies (differential diagnosis)?")		
Item		
I10	<p>Was the aetiologic category clearly defined?</p> <p>Yes(2): Definition of the etiologic category was provided</p>	Information

	Unclear(1) No(0): Only name but no definition of the etiologic category was provided	
I11	Was the diagnostic work up likely to correctly classify the respective aetiology? Yes(2): Appropriate single reference test or composite reference standard OR delayed-type follow up on combination with independent expert panel OR Post-hoc diagnosis established by treating physician Unclear(1) No(0): preliminary diagnosis at the time of consultation OR single reference test or composite test that was not appropriate	Bias
I12	Did every patient receive the same diagnostic work up to detect the respective aetiology? (Yes(2),Unclear(1),No(0))	Bias
	<u>Risk that the diagnostic work up introduce bias: low, unclear, high</u> Low(0): Consider low if I11 and I12 are answered with yes Unclear(1): Consider unclear if I11 and I12 are answered as follows: yes/unclear, unclear/unclear, yes/no High(2): Consider high if I11 or I12 are answered as follows: no/no, no/unclear	Overall assessment 4
Domain D: Determination of the prognosis (refers only to review question “What is the prognosis of patients with the respective symptom presenting in the respective setting?”)		
Item		
I13	Was the prognostic outcome clearly defined? (Yes(2),Unclear(1),No(0))	Information
I14	Did the study design include a comparison group without the symptom?(Yes(2),Unclear(1),No(0))	Information
I15	Was the work up/ measurement likely to correctly classify the respective prognostic outcome? Yes(2): reliable and valid measurement AND length of follow up appropriate acute back pain: 4 weeks, chronic pain: 12 weeks Unclear(1) No(0):unreliable or not valid measurement OR length of follow-up not appropriate	Bias
I16	Did every patient receive the same work up/ mode of data collection to verify the respective prognostic outcome? (Yes(2),Unclear(1),No(0))	Bias
	<u>Risk that the prognostic work up introduced bias: low, unclear, high</u> Low(0): Consider low if I15 and I16 are answered with yes Unclear(1): Consider unclear if I15 and I16 are answered as follows: yes/unclear, unclear/unclear, yes/no High(2): Consider high if I15 or I16 are answered as follows: no/no, no/unclear	Overall assessment 5

8. Akademische Lehrer:

Marburg:

Albert, Arabin, Aumüller, Barth, Bartsch, Bauer, Baum, Baumann, Becker, Bösner, Cetin, Czybayko, Daut, Del Rey, Dodel, Donner-Banzhoff, Feuser, Fendrich, Fuchs-Winkelmann, Görg, Gress, Hegele, Hertl, Holst, Hofmann, Höglinger, Hoyer, Jerrentrup, Josephs, Kann, Kanngießler, Kill, Klose, Koolman, Kühne, Kühnert, Langer, Lill, Löffler, Lohoff, Maier, Maisch, Mandrek, Moll, Moosdorf, Mueller, Müller, Mutters, Neff, Neubauer, Neumüller, Oertel, Pagenstecher, Rausch, Renz, Röhm, Rothmund, Ruchholz, Schäfer, Schlosser, Schmidt, Schofer, Schütz, Seitz, Steiniger, Steinfeld, Tackenberg, Vogelmeier, Voigt, Wagner, Waldmann, Weihe, Werner, Wulf

Irland:

Bartels, Ikele, Mohammed, Ni Bhuinneain

9. Danksagung:

Ich möchte mich ganz herzlich bei meiner Doktormutter Prof. Becker für die tolle Zusammenarbeit bedanken, die mir wann immer nötig mit Rat und Tat zur Seite stand. In gleichem Maße möchte ich auch Herrn Haasenritter danken, der mir insbesondere bei technischen und methodischen Fragen und Problemen behilflich war.

Besonders hervorheben möchte ich auch die erfolgreiche Zusammenarbeit mit meiner Co-Doktorandin Frau Grevenrath und mich für ihren Einsatz in der gemeinsamen Sache bedanken.

Auch meinem Partner Andreas Gudeev, meiner Familie und meinen Freunden danke ich für ihre Geduld und ihr Interesse. Ein besonderer Dank gilt noch meinen Eltern ohne deren Unterstützung ich mir nicht die nötige Zeit hätte nehmen können, um diese Arbeit zu einem Ende zu bringen.

Ein großes Dankeschön noch zu guter Letzt an meine beiden Korrekturleser Klaus H. Schmidt und Maika Dänner, die für die nötige Genauigkeit in Sachen Rechtschreibung und Zeichensetzung Sorge getragen haben.