

Dr. med. Ralph Schürer  
An der Pirschheide 28  
14471 Potsdam

## **8. Prognose / Verlauf/ Ergebnisbeurteilung von Rückenschmerzen**

Stand 24.09.2023

### **8.0. Einleitung**

2013 publizierte Hazard in der Spine ein Editorial, das die Problematik der Wahl der richtigen Behandlungsziele und deren Erfassung und die Geschichte dieses Problems so gut beschreibt, dass es hier in einer freien Übersetzung des Verfassers wiedergegeben werden soll.

1987 publizierte Spine, das offizielle Journal der International Society for the Study of the Lumbar Spine (ISSLS) zwei grundlegende Artikel, die sich mit Behandlungsergebnissen bei unteren Rückenschmerzen beschäftigten. In dem einen forderten die Herausgeber Alf Nachemson und Henry LaRocca, dass „alte Feststellungen in Frage gestellt und neue Methoden angemessen untersucht werden müssen“ (Nachemson & LaRocca, 1987). Sie sprachen sich für randomisierte, kontrollierte Studien aus, die „wissenschaftlich validierte Formen der Einstufung (grading) von Rückenschmerz und Behinderung“ nutzen, wie die Million, MMPI, Oswestry und Waddell – Fragebögen. In dem zweiten wandte Gordon Waddell das biopsychosoziale Modell von George Engel auf Rückenschmerzen an (Engel, 1977, Waddell, 1987). Er hatte beobachtet, dass sich die Behinderung durch Rückenschmerzen ungeachtet der Fortschritte der modernen biomedizinischen Ansätze in der Diagnostik und Therapie schmerzhafter Pathologien zu einer Epidemie in westlichen Kulturen entwickelt hatte. Er erkannte an, dass die meisten Menschen zu irgend einem Zeitpunkt Rückenschmerzen hatten und sich die meisten schnell davon erholten oder zumindest ohne medizinische Behandlung weiter machten. Bei denen, die das nicht taten, schwächte die Chronizität die Beziehung zwischen beobachtbarer Pathologie, Schmerz, physischer Beeinträchtigung und Beschäftigungsstatus. Mehr als durch die Pathologie wurden die Entscheidungen des Arztes hinsichtlich Diagnostik, Therapie und der Frage, ob die Schwere des Schmerzes zu Arbeitsunfähigkeit und Entschädigungsleistungen führt, durch den Distress (d.h. nicht nur Stress, sondern auch Leiden, Kummer, Qual usw., d. Verf.) und das Krankheitsverhalten des Patienten bestimmt. Deshalb schlug Waddell ein neues, biopsychosoziales Modell vor, welches den biomedizinischen Fokus erweitern und alle Aspekte der physischen, psychologischen und sozialen Situation der Person berücksichtigen würde. Ebenfalls 1987 beschrieb Nordin Hadler die „Zwickmühle (predicament)“ bei Rückenschmerzen, in der eine Person mit diesem Symptom sich zu entscheiden hat, ob sie damit selbst fertig wird oder medizinische Hilfe sucht, wodurch sie zu diesem Zeitpunkt zu einem Patienten mit einer Krankheit wird (Hadler, 1987). Wenn die Krankheit die physische Kapazität einschränkt, kann der Patient zu einem Anspruchsteller auf den Gebieten von Behinderung und Entschädigung werden mit der Pflicht, die Schwere seiner Notlage zu beweisen.

Hazard führt dann aus, dass es Ziel dieses Editorials sei, kurz die Entwicklungen in der Epidemiologie und den entstehenden Konzepten zu Rückenschmerzen in den USA im letzten Vierteljahrhundert darzustellen.

### **25 Jahre biomedizinischer Praktik und Forschung**

Das Geschäft hinter dem biomedizinischen Modell brummte in jeder Hinsicht, aber wohin führte das? Unsere diagnostischen Werkzeuge und Techniken haben sich nicht wesentlich verbessert, und wir können für die meisten Menschen mit Rückenschmerzen immer noch keine definitive Diagnose

stellen. Die große Mehrheit der Menschen ohne red flags für Krebs oder Infektion haben keine klare, relevante und reversible Läsion im MRT und der klinische Nutzen eines zeitig in einer Rückenschmerzepisode angefertigten oder statt einer Röntgenaufnahme durchgeführten MRTs ist sehr begrenzt, teilweise sogar schädlich (Jarvik et al., 2003, Modic et al., 2005, Graves et al., 2012). Die meisten Menschen beenden den diagnostischen Prozess schließlich mit einem oder mehreren Etiketten, die bestenfalls „unspezifische Bezeichnungen ohne pathoanatomische Bestätigung“ darstellen (Deyo & Weinstein, 2001). 2005 wurde in einer elektronischen Umfrage unter ISSLS-Mitgliedern gefragt „welche einzelne Frage zu Diagnose oder Therapie von Rücken- oder Nackenschmerzen wäre in den nächsten 10 Jahren am wichtigsten zu beantworten?“. Von den 44 Antwortenden dachte die große Mehrheit, eine bessere Diagnose wäre entscheidend: die meisten bezogen sich auf eine anatomische oder biochemische Diagnose im Intervertebralsegment, der Rest empfahl eine klarere Identifikation von klinischen Subgruppen mit entsprechend angepasster Therapie. Bis heute wurde keine dieser Forderungen im klinischen Alltag zufriedenstellend gelöst. Es gibt wenig Übereinstimmung über angemessene diagnostische Strategien, im Ergebnis gibt es eine weite Vielfalt in der Praxis (Chou et al., 2007). Mittlerweile geben wir mehr und mehr Geld für Diagnostik und Therapie sowohl für Nacken- als auch für Rückenprobleme aus. In den USA stiegen die Gesamtausgaben des Gesundheitswesens für Erwachsene mit Rücken- und Nackenproblemen zwischen 1997 und 2005 um 65 % (Martin et al., 2008). Die schnell und stetig steigenden Ausgaben für MRTs, Facetten- und epidurale Steroidinjektionen, Opioidverordnungen und chirurgische Verfahren sind gut dokumentiert (Deyo et al., 2009), für 2005 werden die jährlichen Kosten in den USA auf 86 Milliarden Dollar geschätzt (Martin et al., 2008).

Was hat dieses massive Investment gebracht? 1993 beschäftigte sich John Frymoyer mit dieser Frage und empfahl der ISSLS die Gleichung: Wert = (Patientenzufriedenheit + Behandlungsergebnis)/Kosten (Frymoyer, 1993). Weil die Kosten immer mehr steigen, ist es umso trauriger festzustellen, dass die Ergebnisse bei Amerikanern mit Wirbelsäulenproblemen nach Angaben des durch die Agency for Healthcare Research and Quality durchgeführten Medical Expenditure Panel Survey Jahr für Jahr schlechter werden. Unter den jährlich durchschnittlich 1774 Personen mit Wirbelsäulenproblemen, die die SF-12 Physical and Mental Health Summaries ausfüllten, stieg der Anteil derjenigen, die Einschränkungen in physischen, sozialen und Arbeits-/Schul- oder häuslichen Aktivitäten angaben, von 1997 bis 2006 ständig und signifikant (Martin et al., 2009).

Die Zufriedenheit von Patienten ist ein Problem mit vielen Facetten, wozu die persönliche Wertung und Wahrnehmung von Symptomen, Lebensqualität und Behandlungserfahrungen gehören (Hudak & Wright, 2000). Eine bemerkenswerte Erscheinung bei Menschen mit chronischen Rückenschmerzen ist, dass die Patientenzufriedenheit nur schwach mit den üblichen, allgemein in wissenschaftlichen Studien benutzten Messgrößen der Behandlungsergebnisse korreliert: Schmerz und Funktion, gemessen mit den Fragebögen, die von Nachermon und LaRocca empfohlen wurden und den daraus folgenden Instrumenten zur Erfassung des generellen Gesundheitszustandes wie der 36-Item Short Form Health Survey (Hazard et al., 1994). Da sich die Beobachtung Waddells der Nichtübereinstimmung von Schmerz, Beeinträchtigung und Behinderung bei Patienten mit chronischen und noch mehr bei akuten Rückenschmerzen bestätigte (Kovacs et al., 2004) und die Standardmessgrößen der Behandlungsergebnisse die Patientenzufriedenheit nicht gut zu erklären scheinen, war die Bestimmung des Wertes einer Therapie tatsächlich schwierig.

Wie gut haben uns diese Messungen gedient? Idealerweise hätten sie uns geholfen, aus den Myriaden klinischer Studien, die den Empfehlungen von Nachermon und LaRocca folgten, nützliche Therapien zu identifizieren. Allerdings haben sogar die besten randomisierten kontrollierten Studien bei dem relativ klaren klinischen Szenario von Radikulärsyndromen bei Bandscheibenvorfällen widersprüchliche Ergebnisse erbracht (Weinstein et al., 2006, 2008). Das Gleiche trifft für neuere Studien zu Fusion und nichtoperativer Therapie bei nichtmalignem Rückenschmerz zu (Fritzell et al., 2001, Brox et al., 2003, Fairbank et al., 2005). Neben anderen Problemen hatte die Größe der Differenz zwischen den Behandlungsergebnissen operierter und nicht operierter Patienten trotz

Signifikanz in einigen Fällen nur marginale klinische Relevanz. Dies war ein allgemeines Problem klinischer Studien, die generische Fragebögen zur Erfassung von Behandlungsergebnissen nutzten, eine Strategie, die generell nur relativ triviale Ergebnisse brachte (Deyo, 2004), ungeachtet der Bemühungen, dieser Effektgröße mit Standards wie der minimal bedeutsamen Veränderung und der Änderungssensibilität (responsiveness) der Fragebögen eine Bedeutung zu geben (Copay et al., 2007, 2008, Frost et al., 2008). Dieses Unvermögen, Behandlungs“champions“ zu identifizieren, hat Kliniker frustriert, die Behandlungen verschreiben müssen, wohl wissend, dass Patienten, die für das gleiche klinische Problem die gleiche Behandlung bekommen, unerklärlicherweise sehr unterschiedliche Behandlungsergebnisse haben können. Davon ausgehend, dass wir üblicherweise für unsere Rückenschmerzpatienten keine einzelne beste Therapie empfehlen können, wurde vorgeschlagen, die Patienten aus einem vorurteilslosen Menü von Therapieoptionen ihre Therapie selbst auswählen zu lassen (Weinstein, 2005). Dieses Modell der geteilten Entscheidungsfindung muss aber noch zeigen, dass es die Behandlungsergebnisse verbessert. Eng verwandt ist die Choosing Wisely Initiative der American Board of Internal Medicine Foundation, die die Kosten im Gesundheitswesen dadurch reduzieren will, dass vor Untersuchungen und Therapien gewarnt wird, die keinen nachgewiesenen Nutzen haben, wozu auch das frühe MRT bei Rückenschmerzen gehört (Cassel & Guest, 2012). Diese Ziel entwickelte sich aus Howard Brodys Aufruf an die medizinischen Fachgesellschaften, eine Liste der „top five“ der unnötigen Tests und Prozeduren zu erstellen (Brody, 2012), eine Herausforderung, die von der ISSLS zukünftig aufgenommen werden könnte.

## 25 Jahre Behinderung durch Rückenschmerzen

Was hat das letzte Vierteljahrhundert biomedizinischen Wachstums getan, um aufzuhalten was Waddell als westliche Epidemie der Behinderung durch Rückenschmerzen bezeichnete? Nach den meisten Berichten hat sich die Prävalenz von Rückenschmerzen nicht wesentlich verändert (Deyo et al., 2006). Die Rate an Behinderung variiert nach Quelle und Definition, aber nach einem oft zitierten Bericht, der 2000 publiziert wurde, ist etwa 1 % der Erwachsenen in den Vereinigten Staaten anhaltend wegen Rückenschmerzen behindert und ein weiteres Prozent vorübergehend (Nachemson et al., 2000). Die maßgeblichsten Trendberichte sind alarmierender. In den USA war die Zahl der Anträge auf Social Security Disability Income (eine Art Erwerbsunfähigkeitsrente, d. Verf.) bei Erkrankungen der Atemwege und des Herz-Kreislaufsystems zwischen 1996 und 2005 stabil oder rückläufig. Im gleichen Zeitraum stieg der Anteil dieser Leistungen für Erkrankungen des muskuloskeletalen Systems, in erster Linie für Rückenschmerzen, kontinuierlich von 20,6 % in 1996 auf 25,4 % in 2005 (Deyo et al., 2009).

Was bedeutet diese Behinderung und führt uns die Definition dieser Behinderung zu realen Lösungen? Auf einer Bevölkerungsebene stammt die Definition einer Behinderung aus medizinischen, sozialen ökonomischen und kulturellen Modellen (Verbrugge & Jette, 1994, Waddell, 2002, Jette, 2005). Im medizinischen Modell, das Waddell als in den 1980ern vorherrschend beschreibt (Waddell, 1987), muss eine Person mit einer Krankheit oder Verletzung, die sie an einer einträglichen Beschäftigung oder dem Leben in einer gewöhnlich verfügbaren Wohnung hindert, medizinisch wiederhergestellt werden oder soweit entschädigt werden, dass ein Überleben möglich ist. Das ist es, was die Frustration der im Gesundheitswesen Praktizierenden am meisten verstärkt, weil Personen mit einer Behinderung durch Rückenschmerzen sich oft trotz medikamentöser oder chirurgischer Interventionen nicht funktionell erholen. Hierin liegt der Haken von Hadlers Rückenschmerzzwickmühle: der Patient mit Arbeitsunfähigkeit wegen Krankheit muss die Schwere seines Problems beweisen und der Beweis erfordert die Einhaltung ärztlicher Empfehlungen. Die Erfassung von Symptom und Schwere der Beeinträchtigung ist im Setting und in der Zeit für eine Konsultation zumindest schwierig. Im sozialen Modell ist Behinderung keine persönliche Eigenschaft, sondern ein Mangel angepasster physischer und sozialer Ausstattung (environment), der nach politischen Lösungen ruft. Nach dem ökonomischen Modell hält die Verfügbarkeit einer Entschädigung von der Arbeit ab, also ist die Reduzierung von Behinderung eine Sache der Begrenzung von Entschädigungsleistungen. Im kulturellen Modell bestimmen kollektive Haltungen

und Vorstellungen der Gesellschaft die Behinderungspolitik und die Behinderungsraten wechseln durch komplexe Trends dieser subjektiven Einflüsse.

Die kritische Frage für Gesundheitspolitiker und Forscher bleibt: wie können sie ihre Anstrengungen hinsichtlich Individuen mit Rückenschmerzen besser gegen die stetige Flut von Behinderungen richten und diese eindämmen?

## 25 Jahre Lektionen, gelernt von Patienten in funktionellen Rehabilitationsprogrammen

1987 unterstrich Waddell (Waddell, 1987) die Notwendigkeit, Schmerz von Behinderung zu unterscheiden und empfahl, einen Wechsel von der Überbewertung von Schmerz allein in Richtung Rehabilitation und Wiederherstellung von Funktion. Gleichzeitig entwickelte Tom Mayer multidisziplinäre Rehabilitationsprogramme für Menschen mit behindernden Wirbelsäulenproblemen. Diese funktionellen Wiederherstellungsprogramme (Functional Restoration Programs, FRP) in Texas und später in Vermont mischten Verhaltensunterstützung mit fortschreitendem physischen Training in Richtung der Normen für Beweglichkeit, Kraft und Ausdauer vergleichbarer Gesunder ohne Rückenschmerzen (Mayer et al., 1987, Hazard et al., 1989). Obwohl erste klinische Studien hinsichtlich funktioneller Wiederherstellung und Arbeitsfähigkeit vielversprechend waren, erbrachten Langzeitnachbeobachtungen eine kritische Einsicht in die Natur von Rückenschmerzen und Behinderung. Fünf Jahre nach Behandlung im Vermont-FRP waren individuelle Patienten oft mit ihren Behandlungsergebnissen unzufrieden, obwohl die traditionellen medizinischen Behandlungsergebnisse auf Gruppenniveau sehr positiv gewesen waren (Hazard et al., 1994). Individuelle Interviews mit diesen Patienten und Diskussionen in kleinen Gruppen mit hunderten nachfolgender FRP-Teilnehmern offenbarten das Versagen typischer Fragebögen und von Messgrößen physischer Kapazitäten in der Frage, inwieweit die Funktion wiederhergestellt wurde oder nach dem Symptomspiegel, der für ihn oder sie von Bedeutung war. Beispielsweise können zwei Personen mit dem gleichen Oswestry- oder Schmerzscore sehr unterschiedliche Grade an Zufriedenheit aufweisen, wenn der eine nicht in der Lage war, die einzige verfügbare Arbeit auszuüben und der andere sich im Ruhestand befindet. Folgende Jahrzehnte dieser Kleingruppeninterviews in den Vermont- und anschließend New Hampshire-FRPs zeigten, dass durch Rückenschmerzen behinderte Menschen eine Vielfalt an Vorstellungen und Einstellungen zu ihrem Schmerz und ihrer Funktionskapazität haben, die schließlich ihre Erwartungen an die Therapie bestimmen, weit entfernt von ihren Symptomen und welche Pathologie sie auch immer entdeckt haben mögen (Hazard et al., 2009a). Noch wichtiger, die persönlichen Funktionsziele für Arbeit, Erholung und tägliche Aktivitäten reichen weit, vom aktiven Militärdienst über die Teilzeitarbeit als Kassierer bis zum einfachen Wunsch, mit den Enkeln spielen zu können.

Aus diesen Interviews und Diskussionen entwickelte sich das Zielerreichungsmodell (goal achievement model) für die Behandlung und die Erfassung von Behandlungsergebnissen bei Personen mit schmerzhaften, behindernden spinalen Problemen. 2001 fanden die Forscher des Vermont-FRP, dass das Ausmaß, in dem Patienten glaubten, ihre funktionellen Ziele erreicht zu haben, ein stärker bestimmender Faktor für deren generelle Zufriedenheit mit der Therapie war, als ihre Symptome und objektiven und subjektiven funktionellen Ergebnisse (Hazard et al., 2001). Aktueller wurden 2 Kohorten von Absolventen des New Hampshire-FRPs mit personalisierten Langzeitfragebögen hinsichtlich Schmerz und funktionellen Behandlungsergebnissen (NRS, der Subscore für physische Funktion des 36-Item Short Form Health Survey) befragt, außerdem schätzten sie selbst das Maß der Erreichung ihres vor der Behandlung selbst gesetzten Ziels ein (Hazard et al., 2009, 2012). In beiden Studien war wiederum die funktionelle Zielerreichung am engsten mit der Zufriedenheit mit „Ihrem Fortschritt mit Ihrem Schmerzproblem“ verbunden.

Das Zielerreichungsmodell für die Behandlung von Rückenschmerzen

1993 fasste John Ware, der Entwickler des 36-Item Short Form Health Survey, einen Trend der medizinischen Literatur des zwanzigsten Jahrhunderts so zusammen: ein medizinisches Behandlungsergebnis bedeutet jetzt das Ausmaß, in dem eine Veränderung in dem funktionellen Verhalten eines Patienten oder seines Wohlbefindens seinen Bedürfnissen oder Erwartungen entspricht („a medical outcome has come to mean the extent to which a change in a patient's behavioral functioning or well-being meets the patient's needs or expectations“) (Ware, 1993). 2001 empfahl das Institute for Medicine eine mehr patientenzentrierte Behandlung, die „respektvoll ist und auf die individuellen Wünsche, Bedürfnisse und Werte des Patienten bei allen klinischen Entscheidungen eingeht“ (Institute of Medicine, 2001). 2012 wurde eine zielorientierte Patientenbetreuung als Paradigma für die Behandlung von Menschen mit vielen chronischen Erkrankungen und schwerer Behinderung vorangebracht (Reuben & Tinetti, 2012). Das Zielerreichungsmodell ist gut positioniert, um dieses Bedürfnis in klinischen und wissenschaftlichen Bestrebungen zur Reduzierung von Behinderungen bei Rückenschmerzen zu erfüllen. In diesem Modell wird bei dem initialen Patientenkontakt zu traditioneller Anamneserhebung, körperlicher Untersuchung und Bildgebung die Erfassung funktioneller und symptombezogener Ziele hinzugefügt. Leistungsanbieter und Patient werten und wählen gemeinsam diagnostische und therapeutische Optionen, die zu den Zielen des Patienten passen. Dieser Prozess ist wichtig, um Personen zu helfen, die nicht nur behindernde Rückenschmerzen, sondern oft auch noch einen Komplex von Kombinationen medizinischer, emotionaler und sozialer (environmental) Probleme aufweisen. Dies sind die echten Herausforderungen, die Waddell so passend in seinem biopsychosozialen Modell beschrieb. Das Zielerreichungsmodell gibt auf verschiedene Weise das biopsychosoziale Modell wieder, da es dem Patienten gestattet, mit seinem Therapeuten die überzeugenden Vorstellungen, Haltungen und Werte innerhalb seiner biologischen, physischen und sozialen Situation herauszuarbeiten und in gemeinsam akzeptable Ziele und machbare Pläne umzusetzen. Obwohl das Symptom und die physische und soziale Funktion mit traditionelleren Ergebnisfragebögen und Test erfasst und gemessen werden können, verschiebt sich der Fokus der Erfassung von Behandlungsergebnissen in Richtung des Ausmaßes, in dem jeder Patient die Ergebnisse erreicht, die für ihn oder sie wichtig sind.

Die Erfassung der Zielerreichung wurde in der Vergangenheit auf mentale Gesundheitsprobleme angewandt, wobei die Zielsetzung primär durch den Therapeuten erfolgte (Cardillo & Smith, 1994). Das hier beschriebene Zielerreichungsmodell, das erst in den letzten Jahren eingeführt wurde, um das Problem von Rückenschmerz und Behinderung anzugehen, basiert auf den persönlichen Zielen des Patienten. Eingedenk Nachemsons und LaRoccas Ruf nach wissenschaftlicher Validierung müssen wir viel über die Vorstellungen und Haltungen (attitudes), welche die Ziele eines Individuums bestimmen, lernen, über die besten Methoden, die echten Ziele und Prioritäten des Patienten herauszufinden, die Messung der persönlichen Zielerreichung und deren Einfluss auf den Gesundheitszustand der Population. Ein besseres Verständnis dieser 4 Bereiche der Zielerreichung und die Untersuchung relevanter Interview- und Fragebogenstrategien zu deren Enthüllung und Messung kann unsere klinischen und wissenschaftlichen Anstrengungen, Menschen mit behindernden Rückenschmerzen zu helfen, nur unterstützen. Für die Gesundheitsindustrie gedeiht immer noch das biomedizinische Modell mit seiner Orientierung auf typische Behandlungsergebnisse mit seinem Honorar-für-Leistung – Paradigma. Hoffentlich können wir Patienten mit Rückenschmerzen besser helfen, weniger behindert zu sein, indem wir Waddells biopsychosoziales Modell in ein praktisches Modell für Forschung und Praxis überführen: das Zielerreichungsmodell (Hazard, 2013).

Auf ein anderes Problem weisen McIntosh et al. (2018) hin: So lange Forscher und Kliniker keine anspruchsvollen und rigorosen Methoden von statistischer bzw. externer Validität einsetzen, werden sie unkorrekte Annahmen vornehmen und falsche Schlüsse hinsichtlich Behandlungseffekten und Ergebnissen ziehen. Diese Aussage fußt auf eine Analyse der in den Zeitschriften Spine und Spine Journal zwischen Januar 2013 und Januar 2014 publizierten Arbeiten zur Prognose von Rückenschmerzen.

Garcia et al. (2019) suchten nach Prädiktoren, die die Behandlungsergebnisse bei chronischen Rückenschmerzen voraussagen können und fanden die Behinderung zu Behandlungsbeginn als universalen Prädiktor für die Behandlungsergebnisse nach 6 Monaten und den Schmerz zu Beginn für die Behandlungsergebnisse nach 12 Monaten.

Tagliaferri et al. (2019) empfehlen Klinikern, die multidimensionale Natur chronischer Rückenschmerzen zu beachten, zu der physische (Beeinträchtigung, Muskelkraft, Ausdauer, Leistungsfähigkeit für Aktivitäten des täglichen Lebens und Körperzusammensetzung), psychologische (Kinesiophobie, Angst-Vermeidung, Schmerzkatastrophisierung, Selbst-Effizienz, Depression, Angst, Schlafqualität), soziale (soziales Funktionieren, Arbeitsunfähigkeit) und auf die Gesundheit bezogene Lebensqualität gerichtete Messgrößen gehören, individuell auf das einzelne Individuum abgestimmt.

Hanke et al. (2021) fanden 33 Tests zur Einschätzung des Return-to-Activity – Status bei unspezifischen Rückenschmerzen, wobei die Gehfähigkeit, das Verhalten bei Heben und Tragen von Gegenständen, die motorische Kontrolle, die Muskelkraft und die Beweglichkeit eine übergeordnete Rolle spielen.

Ältere Arbeiten siehe Gesamtliteraturverzeichnis

- Brody,H. From an ethics of rationing to an ethics of waste avoidance  
N Engl J Med 366 (2012)1949 – 51
- Cassel,C.K. Guest,J.A.  
Choosing wisely: helping physicians and patients make smart decisions about their care  
JAMA 307 (2012)1801 -2
- Garcia,A.N. Costa,L.O.P., da Cunha Menezes Costa,L., Hancock,M., Cook,C.  
Do prognostic variables predict a set of outcomes for patients with chronic low back pain: a long-term follow-up secondary analysis of a randomized control trial  
J Man Manip Ther (2019) online 4.4.2019;  
[doi.org/10.1080/10669817.2019.1](https://doi.org/10.1080/10669817.2019.1)
- Graves,J.M. Fulton-Kehoe,D., Jarvik,J.G., Franklin,G.M.  
Early imaging for acute low back pain. One-year health and disability outcomes among Washington state workers  
Spine 37 (2012)1617 - 27
- Hanke,A. Schoch,W., Keller,M., Richter,R.  
Funktionelle Testungen zur Ermittlung des Return-to-Activity-Status bei Patienten mit unspezifischen Kreuzschmerzen

Sportverl Sportschad (2021) Mai doi.org/10.1055/a-1272-9781

Hazard,R.G. Goal achievement model for low back pain (editorial)

Spine 38 (2013)1431 – 5

McIntosh,G. Steenstra,I., Hogg-Johnson,S., Carter,T., Hall,H.

Lack of prognostic model validation in low back pain prediction studies. A systematic review

Clin J Pain 24 (2018)748 - 54

Reuben,D.B. Tinetti,M.F.

Goal-oriented patient care – an alternative health outcomes paradigm

N Engl J Med 366 (2012)777 – 9

Tagliaferri,S.D. Miller,C.T., Owen,P.J., Mitchell,U.H., Brisby,H., Fitzgibbon,B., Masse-Alarie,H., van Oosterwijck,J., Belavy,D.

Domains of chronic low back pain and assessing treatment effectiveness

Pain Pract 19 (2019) doi.org/10.1111/papr.12846 (Zugriff 18.11.2019)

### **8.1. Qualitätskriterien**

2006 musste die IMMPACT- Studiengruppe (Turk et al., 2006) feststellen, dass es an standardisierten und umfassenden Verlaufs-/Ergebnisparametern (outcome measures) für Studien zu Schmerzen mangelt, die adäquate vergleichbare Informationen für relevante Untersuchungsgruppen darstellen, und die es Forschern gestatten, Gruppen mit unterschiedlichen demographischen oder krankheitsspezifischen Charakteristika miteinander zu kombinieren und zu vergleichen.

In einem systematischen Review identifizierten Chapman et al. (2011) 75 verschiedene Ergebnisvariablen für chronische Rückenschmerzen. Davon konnten 29 gleich wieder ausgeschlossen werden, da sie jeweils nur in einer Publikation verwendet worden waren. Die in der englischsprachigen Literatur am häufigsten benutzen Scores für funktionelle Ergebnisse waren der Oswestry Disability Index, der Roland-Morris Disability Index und das Bewegungsausmaß (range of motion); für Schmerzen am häufigsten genutzt wurden die numerische Schmerzbewertungsskala (NRS), das Brief Pain Inventory, der Pain Disability Index, das McGill Pain Questionnaire und die visuelle Analogskala (VAS). Für psychologische Fragestellungen kamen am häufigsten das Fear Avoidance Beliefs Questionnaire, die Tampa Scale for Kinesiophobia und das Beck Depression Inventory zum Einsatz. Die allgemeine Lebensqualität wurde am häufigsten mit der Short Form 36, dem Nottingham Health Profile, der Short Form 12 und das Sickness Impact Profile eingeschätzt. Als objektive Messgrößen fungierten die Arbeitsfähigkeit/ Wiederaufnahme der Arbeit, Komplikationen bzw. Nebenwirkungen und die eingesetzten Medikamente. Um eine Kostenbewertung vorzunehmen, kamen Euro-Quol 5 und SF 6 am häufigsten zum Einsatz.

Auf theoretischem Gebiet wird derzeit die klassische Testtheorie von der Item-Response-Theorie (Hampelten, 2000) abgelöst, die zunehmend CAT (Computer Adaptive Testung) – Verfahren einsetzt (De Vet et al., 2003). (Im Nachhinein betrachtet scheint sich diese Aussage allerdings nicht bestätigt zu haben, d. Verf.)

Kamper et al. (2010) beschreiben das derzeitige Dilemma: in 82 Rückenschmerzstudien wurden 66 verschiedene Messgrößen für eine Heilung/Erholung von Rückenschmerzen eingesetzt, 59 von diesen Messgrößen wurden jeweils nur in einer Studie eingesetzt.

Van Tulder et al. (2009) fanden in einer Untersuchung, dass die Einhaltung der Anforderungen der aus 11 Items bestehende Checkliste der Cochrane Back Review Gruppe mit der Effektstärke randomisierter Studien zu Rückenschmerztherapien korreliert.

Der Wert von Ergebnismessungen für die Erfassung der Effektivität einer Therapie wird in der Literatur betont (Beattie & Maher, 1997, Bombardier, 2000, Koes et al., 1995, Wittink & Carr, 2008). Interessanterweise wird bei der Beschreibung der Effektivität einer Behandlung in der englischsprachigen Literatur zwischen „efficacy“ und „effectiveness“ unterschieden. Während „efficacy“ – Studien untersuchen, wie eine Behandlung unter optimalen Umständen, d.h. unter kontrollierten Bedingungen an sorgfältig ausgewählten Patienten wirkt, beschreibt „effectiveness“, wie die Behandlungsergebnisse in der täglichen Praxis sind (Turner & Turk, 2008).

Smuck (2016) weist in einem Editorial darauf hin, dass mittlerweile akzeptiert ist, dass eine statistisch signifikante Differenz bei den Behandlungsergebnissen hinsichtlich Schmerz oder Funktion bedeutungslos ist, wenn die Veränderungen nicht klinisch relevant sind.

Für die Beurteilung des Evidenzniveaus gelten die Leitlinien der Cochrane Back Review Group (van Tulder et al., 2003, van der Velde et al., 2007):

- Starke Evidenz: übereinstimmende Befunde mehrerer hochqualitativer RCT's (randomisierter klinischer Studien)
- Moderate Evidenz: übereinstimmende Befunde mehrerer minderqualitativer RCT's und/oder CCT's (kontrollierter klinischer Studien) und /oder eine hochqualitative RCT
- Begrenzte Evidenz: eine RCT oder CCT niedriger Qualität
- Keine Evidenz: keine klinischen Studien (weder RCT's oder CCT's)
- Widersprüchliche Evidenz: abweichende Befunde mehrerer verschiedener Studien (RCT's und/oder CCT's),

siehe auch die Einleitung zu Kapitel 5.

Die Beurteilung der internen Validität, d.h. der Qualität einer Studie erfolgt nach den Leitlinien der Cochrane Back Review Group (van Tulder et al., 2003, van der Velde et al., 2007) wie folgt, wobei die Fragen jeweils mit ja/nein/unklar bewertet werden:

- Wurde eine Methode genutzt, um eine Sequenz zufälliger Zuordnung zu schaffen (Gene Rating a Radom allocation sequence)?
- War die Zuordnung zur Behandlung verblindet?
- Waren die Behandler hinsichtlich der Intervention verblindet?

- Wurden Begleittherapien vermieden oder waren diese vergleichbar?
- War die Compliance in allen Gruppen vergleichbar?
- Waren die Patienten hinsichtlich der Intervention verblindet?
- War derjenige, der die Behandlungsergebnisse beurteilte, gegen die Art der jeweiligen Intervention verblindet?
- Waren die Messgrößen des Behandlungserfolges geeignet?
- Wurde die Rate an Studienabbrüchen/ die Ausfallrate beschrieben und war diese akzeptabel?
- War das Zeitschema der Erfolgsbeurteilung in beiden Gruppen vergleichbar?
- Enthielt die Analyse eine intention-to-treat Analyse?

Nur die Antwort „ja“ auf die jeweilige Frage wird mit einem Punkt bewertet, als hochqualitativ wird eine Studie eingeschätzt, wenn mindestens 6 von 11 Punkten erreicht werden (van der Velde et al., 2007).

Eine Möglichkeit zur Beurteilung der Validität randomisierter Studien (Bronfort, 2004) ist im Appendix 2 dargestellt. Appendix 3 stellt die von Jadad et al. (1996) aus Oxford beschriebene Möglichkeit dar, mit nur 3 Items die Qualität von RCT's einzuschätzen.

Es gibt Untersuchungen, die zeigen konnten, dass nichtrandomisierte Studien und unverblindete RCT's eher die Überlegenheit einer Innovation über die Standardbehandlung nachweisen (Colditz et al., 1989) und dass RCT's, in denen die Zuordnung zu Behandlungsarmen nicht geheim waren oder in denen die Verblindung nicht klar dargestellt wurde, zu besseren Therapieergebnissen im Studienarm führten (Schulz et al., 2008).

#### Statistische Größen

Eine interessante Methode zur Wertung der Ergebnisse klinischer Studien in Metaanalysen haben Cruccu & Taylor (2007) in Form eines Diagramms mit den beiden Achsen Reliability und Nettogewinn entwickelt. Die Reliability wird durch Gruppengröße und Qualität der Arbeit bestimmt. Alle einbezogenen Arbeiten werden unter Beachtung der statistischen Regeln der Gruppengröße 1 – 5 zugeordnet

Auf Probleme der Unterschiede zwischen statistischer Signifikanz und klinischer Relevanz gehen van Tulder et al. (2007) ein. In randomisierten, kontrollierten Untersuchungen (RCT's) werden Schlussfolgerungen oft aus statistischen Signifikanzen abgeleitet. Wenn  $P < 0,05$  ist, wird üblicherweise geschlussfolgert, dass die Intervention effektiver als der Vergleich / die Vergleichsgruppe ist, wenn  $P > 0,05$ , gilt die Vergleichsbehandlung als effektiver. Dieser  $P -$  Wert ist nach van Tulder et al. nicht sehr informativ, er zeigt nur die Wahrscheinlichkeit des beobachteten Effektes, nicht dessen Größe bzw. Stärke. Der  $P -$  Wert zeigt nicht, ob der Effekt von klinischer Bedeutung ist. Die klinische Bedeutung wird nach Beaton et al. (2002) durch die **minimale klinisch bedeutende Differenz (MCID)** bestimmt.

Hurst & Bolton (2004) beschäftigen sich mit der klinischen Signifikanz von Veränderungen bei Patienten mit Rücken- und Nackenbeschwerden und nutzen dafür den Reliable Change Index ( $> 1,96$ ), die Effektgröße (etwa 0,5) und verschiedene Patienten – Scores.

Gatchel et al. (2010c) setzen sich in einem Editorial im Spine mit der Frage der MCID ausführlicher auseinander. Die MCID wurde von Jaeschke et al. (1989) als der kleinste Unterschied definiert, der vom Patienten als positiv (beneficial) empfunden wird. Es existieren zwei prinzipielle Herangehensweisen, die MCID zu bestimmen: verteilungsbasiert oder ausgangssituations- bzw. Anker-basiert („anchor-based“).

In einem Literaturüberblick stellen Chung et al. (2017) für Rückenschmerzen fest, dass die MCID zunehmend zur Beschreibung eines Therapieerfolges genutzt wird. Dabei mussten die Autoren feststellen, dass nicht nur verschiedene Kalkulationsmethoden zu unterschiedlichen Werten führen, sondern auch ähnliche Methoden unterschiedliche Werte erzeugen. Außerdem wurden verschiedene MCIDs bei kleiner Studienpopulation für eng begrenzte Indikationen oder Methoden kalkuliert, was ein erhebliches Fehlerpotential (bias) bietet.

Yuksel et al. (2019) weisen darüber hinaus darauf hin, dass die MCID bei operativen und konservativen Therapien von Wirbelsäulendeformitäten unterschiedlich ist.

Asher et al. (2020) untersuchten die Wertigkeit des MCID in der Chirurgie der LWS und kommen zu der Einschätzung, dass eine Verbesserung von Schmerz und Behinderung 12 Monate nach Operation um 30 % eine valide MCID darstellt.

Das verteilungsbasierte Herangehen zielt auf die statistischen Charakteristika der Patientengruppe und vergleicht eine beobachtete Veränderung mit einem Index der Variabilität um zu bestimmen, ob die Veränderung substantiell ist. Als Effektgröße wird eine standardisierte statistische Größe bezeichnet, die die Veränderung eines Scores nach Therapie mit der Standardabweichung vor Therapie vergleicht. Eine andere Methode bestimmt die kleinste erkennbare Veränderung, die größer als der Messfehler auf einem bestimmten Konfidenzniveau ist. Nach Gatchel et al. ist die klinische Bedeutung dieser Werte allerdings unklar.

Das ausgangswertbasierte Herangehen untersucht die Veränderung eines durch den Patienten berichteten Wertes eines Therapieergebnisses in Beziehung auf ein externes Kriterium, den Anker. Diese externen Anker können objektive Ergebnisse wie beispielsweise den Medikamentenverbrauch oder subjektive Einschätzungen des Behandlungsergebnisses sein. Als schwerer methodischer Fehler gilt nach Gatchel et al. allerdings, wenn beide benutzten Ergebnisse subjektiv und miteinander verknüpft sind.

Jacobson et al. (1999) entwickelten den Reliable Change Index (Index zuverlässiger Veränderungen), der zwei eigenständige, aber miteinander verbundene Funktionen hat.(1) er ermittelt die Zuverlässigkeitseigenschaften einer Ergebnisgröße, er bestimmt, ob die Veränderung mit dem was als klinisch bedeutsame Veränderung angesehen wird, korrespondiert. Angenommen, die Scores der gesunden Normalbevölkerung und die Scores der unbehandelten Patientenpopulation sind annähernd normalverteilt und überlappen sich nicht wesentlich, dann kann angenommen werden, dass eine signifikante Veränderung aufgetreten ist, wenn das Behandlungsergebnis einer von 3 Kategorien angehört: es liegt (1) innerhalb von 2 Standardabweichungen vom Mittelwert Gesunder, (2) mehr als 2 Standardabweichungen des Mittelwert der Unbehandelten in Richtung Gesunder, und (3) auf der gesunden Seite des Überlappungsbereichs beider Gruppen. Es kann nach Gatchel et al. (2010c) allerdings insbesondere bei chronischen muskuloskeletalen Schmerzen schwierig bis unmöglich sein, eine gesunde Norm zu definieren.

Der IMMPACT – Konsensusbericht von 2005 (Dworkin et al., 2005) geht von einem positiven klinischen Ergebnis bei einer Reduktion des selbstberichteten Schmerzes um 30 % aus.

Van Tulder et al. (2007) gehen von einer MCID für Schmerz von 20 % und für Funktion von 10 % aus.

Im Ergebnis eines Workshops 2006 in Amsterdam wurde in dieser Hinsicht folgender Konsensus erreicht: minimal wichtige Veränderungen werden bei Veränderungen der Visuellen Analogskala von 15, der Numerischen Analogskala von 2, der Roland Disability Questionnaire von 5, des Oswestry Disability Index von 10 des Quebec Back Pain Disability Questionnaire von 20 angenommen (Ostelo et al., 2008). Fritz et al. (2009) schlagen als MCID eine Reduktion des Oswestry – Index um 50 % vor. In einem Kommentar zu dieser Arbeit weist Spratt (2009) auf die Problematik von 50 % des Ausgangswertes für verschiedene Skalen hin und schlägt vor, eher 50 % der möglichen maximalen Verbesserung zu nehmen.

Diaz-Ambas et al. (2017) stellen fest, dass unterschiedliche Methoden zur Berechnung eines MCID zu unterschiedlichen Werten führen. Die Autoren nutzen die Subskalen PCS (physische Komponenten) und MCS (mentale Komponenten) des SF-12 und fanden bei Patienten mit subakuten und chronischen Rückenschmerzen bei Werten von > 3,77 für PCS und > 3,29 für die MCS eine klinisch wichtige Veränderung der Lebensqualität.

Bei der Bewertung von Verlaufsparametern die auf Patientenangaben beruhen, ist die mögliche Veränderung im Antwortverhalten (Anderson et al., 2009, Schwartz & Finkelstein, 2009) zu beachten. So kann beispielsweise ein Patient nach der Schmerzstärke auf der VAS befragt, heute nach seinen bisherigen Erfahrungen einen Wert von 7 angeben, später, nach der Erfahrung einer ersten Nierenkolik, könnte der gleiche Schmerz möglicherweise nur noch mit 5 beschrieben werden. In einem Überblick zu dieser Fragestellung betonen Barclay-Goddard et al. (2009) die Notwendigkeit eines Konsensus in Terminologie und theoretischen Modellen, damit alle Forscher von gleichen Ausgangspunkt ausgehen. Die klinische Bedeutung und die besten Möglichkeiten der Bestimmung dieses Phänomens („response shift“) müssen bestimmt werden und es ist notwendig zu bestimmen, wie das veränderte Antwortverhalten am besten bestimmt werden kann, wenn es von klinischer Bedeutung ist. In Beziehung auf die gesundheitsbezogene Lebensqualität ist der patientengenerierte Index (PGI, Ahmed et al., 2005) interessant: Der Patient wird aufgefordert, 5 für ihn wichtige Lebensgebiete (z.B. Autofahren, Reisen, Anziehen) zu benennen und seine Fähigkeiten auf diesen Gebieten mit 0 – 10 Punkten zu bewerten. Anschließend kann er die Wertigkeit dieser Gebiete mit insgesamt 12 Marken bewerten. Über die Zeit zeigen sich Veränderungen in der Priorität in der veränderten Reihenfolge in der Priorität der gewählten Domänen (Barclay-Goddard et al., 2009).

Hirschfeld & Zernikow (2013) untersuchten die Differenzierung zwischen milden, mäßigen und starken Schmerzen anhand der NRS bei Kindern und fanden auf der 11-stufigen Skala (0 – 10) als Grenzwerte NRS 4 und 8. Bei Patienten mit chronischen muskuloskelettalen Schmerzen fanden Boonstra et al. (2014) Werte auf einer VAS bis 3,4 mit leichten funktionellen Einschränkungen verbunden, von 3,5 bis 6,4 mit mäßigen und darüber mit starken Einschränkungen assoziiert.

Birnie et al. (2012) weisen auf die Bedeutung individueller Faktoren hin. So ist es nicht das Ziel der Erfassung der klinischen Bedeutung von Schmerzen, ob das Individuum das erreicht, was für die meisten Patienten befriedigend wäre, sondern ob sie oder er erreicht, was für das Individuum persönlich von Bedeutung ist. Die Autoren schlagen deshalb vor, Therapieziele individuell zu bestimmen.

Ältere Arbeiten siehe Gesamtliteraturverzeichnis

Asher,A.M.

Oleisky,E.R., Pennings,J.S., Khan,I., Sivagenasan,A., Devin,C.J., Bydon,M., Asher,A.L., Archer,K.R.

Measuring clinically relevant improvement after lumbar spine surgery: is it time for something new?

Spine J 20 (2020)847 - 65

- Birnie,K.A. McGrath,P.J., Chambers,C.T.  
When does pain matter? Acknowledging the subjectivity of clinical significance  
*Pain* 153 (2012)2311 - 4
- Boonstra,A.M. Schiphorst Preuper,H.R., Balk,G.A., Steward,R.E.  
Cut-off points for mild, moderate, and severe pain on the visual analogue scale for pain in patients with chronic musculoskeletal pain  
*Pain* 155 (2014)2545 - 50
- Chung,A.S. Copay,A.G., Olmscheid,N., Campbell,D., Walker,J.B., Chutkan,N.  
Minimum clinical important difference. Current trends in the spine literature  
*Spine* 42 (2017)1096 - 1105
- Diaz-Ambas,M.J. Fernandez-Serrano,M., Royuela,A., Kovacs,F.M., Gallego-Izquierdo,T., Ramos-Sanchez,M., Llorca-Palomera,R., Pardo-Hervas,M., Martin-Pariente,O.S.  
Minimal clinically important difference in quality of life for patients with low back pain  
*Spine* 42 (2017)1908 - 16
- Hirschfeld,G. Zernikow,B.  
Variability of “optimal” cut points for mild, moderate and severe pain: Neglected problems when comparing groups  
*Pain* 154 (2013)154 – 9
- Hurst,H. Bolton,J.  
Assessing the clinical significance of change scores recorded on subjective outcome measures  
*J Manipul Physiol Ther* 27 (2004)26 – 35  
[doi.org/10.1016/j.jmpt.2003.11.03](https://doi.org/10.1016/j.jmpt.2003.11.03)
- Smuck,M. A slap on the back and a pat on the head (editorial)  
*Spine J* 16 (2016)1500 -2
- Yuksel,S. Ayhan,S., Nabihev,V., Domingo-Sabat,M., Vila-Casademunt,A., Obeid,I., Perez-Grueso,F.S., Acaroglu,E.,

Minimal clinical important difference of the health-related quality of life scales in adult spinal deformity calculated by latent class analysis: is it appropriate to use the same values for surgical and nonsurgical patients?

Spine J 19 (2019)71 - 8

## 8.2.Therapieziele/Zielgrößen

Um eine Ergebnisbewertung durchführen zu können, ist es primär nötig, Therapieziele zu definieren.

Über die zu erfassenden Zielgrößen gibt es bislang keine Einigkeit. So fanden Henschke et al. (2014) in 28 RCT's zur Therapie chronischer Rückenschmerzen 20 unterschiedliche Definitionen eines Therapieerfolgs.

Um die Ergebnisse einer Therapie einschätzen zu können, hat ein internationales multidisziplinäres Forum Messgrößen definiert, um die physische Funktionalität, die Schmerzintensität und die gesundheitsbezogene Lebensqualität erfassen zu können (Chiarotto et al., 2016, 2018).

Vorgeschlagen wurden der Oswestry Disability Index (ODI) Version 2.1a oder der 24-Item Roland-Morris Disability Questionnaire für die physische Funktionalität, die NRS für die Schmerzintensität und SF-12 oder ein 10-Item PROMIS Global Health Form für die schmerzbezogene Lebensqualität.

Eine europäische Arbeitsgruppe (Kaiser et al., 2018) hat zur Erfassung der Therapieergebnisse einer interdisziplinären multimodalen Schmerztherapie folgende 8 Domänen definiert: Schmerzintensität, Schmerzfrequenz, emotionales Wohlbefinden, Zufriedenheit mit der sozialen Rolle und sozialen Aktivitäten, Produktivität (bezahlt und unbezahlte, zu Hause und auf Arbeit, einschließlich Anwesenheit und Fehlzeiten), gesundheitsbezogene Lebensqualität und die Zufriedenheit des Patienten mit der Erreichung von Therapiezielen.

Die Cochrane Back Review Group definiert 5 für den Patienten bedeutsame Domänen, die als Ergebnisse evaluiert werden sollen: Schmerzen, rückenspezifische Funktionen, Einschränkungen der Arbeitsfähigkeit, genereller Gesundheitszustand und die Patientenzufriedenheit (Marin et al., 2013).

Pfingsten et al. (1997a) wählten als Kriterien des Behandlungserfolges die Parameter a) Wiederherstellung der Arbeitsfähigkeit, b) Schmerzreduktion und c) subjektive Patientenbeurteilung bezüglich des Behandlungserfolges.

Neben der Schmerzintensität (SI) und Beeinträchtigung (Disability Score, DS) gilt die Lebensqualität als eine weitere zentrale Therapieerfolgsvariable (Scholich et al., 2011).

Nach Niemier et al. (2012) ist die Wiederherstellung der Funktionsfähigkeit ein Grundprinzip der multimodalen Schmerztherapie und beinhaltet sowohl die körperlichen bzw. biologischen wie auch die psychosozialen Funktionen.

Die Notwendigkeit zu wissen, welche Behandlungen Schmerzen reduzieren und welche die gesundheitsbezogene Lebensqualität verbessern und für welche Behandlungen eine Kostenerstattung sinnvoll ist, hat zur Entwicklung einer Vielzahl von Ergebnis- / Verlaufsparametern geführt, allein für muskuloskelettalen Schmerz wurden durch die WHO 147 unterschiedliche Instrumente gezählt (Bowling, 2000). In der ICF (WHO's International Classification of Functioning, Disability and Health; WHO, 2001) wurde ein Rahmen geschaffen der die Charakteristika eines

Gesundheitszustandes als Gesundheitsstörung oder Erkrankung in den Säulen „gestörte Körperfunktionen und –strukturen“, „Einschränkung von Aktivitäten“ und „Eingeschränkter Teilnahme“ beschreibt, wobei jeder Komplex wiederum von externen und persönlichen Faktoren beeinflusst wird (s. Appendix 6).

In einem systematischen Review der Literatur zu objektiven Messgrößen für funktionelle Einschränkungen bei degenerativen Erkrankungen der LWS fanden Stienen et al. (2019) 21 verschiedene objektive Messgrößen. Dabei war der häufigste Test der Test der Zeit für Aufstehen und Losgehen (timed up- and go – test), gefolgt von einem Laufbandtest und der Zeit für 5 x Aufstehen aus dem Sitzen. Die Autoren plädieren für eine internationale Abstimmung auf vergleichbare Tests.

Messwerte der Lebensqualität sind subjektiv und variabel, aber die Verbesserung der Lebensqualität ist das nach der Verlängerung der Lebensdauer wichtigste Ziel jeder medizinischen Behandlung. Schmerz kann an der Reduktion der Lebensqualität maßgeblich beteiligt sein und schwerer Schmerz kann alles andere überlagern, andererseits gehört Schmerz aber auch zum täglichen Leben (Farrar, 2010). Nach Scholich et al. (2011) wird die Lebensqualität bei Rückenschmerzpatienten nicht von organpathologischen Parametern sondern wie die Beeinträchtigung durch die Schmerzintensität und zahlreiche psychologische Faktoren beeinflusst, wobei ein Zusammenhang mit Depression, Angst und Krankheitsverarbeitung besteht.

Im deutschsprachigen Raum fällt auf, dass bei Rückenschmerzen völlige Uneinigkeit über Therapieziele herrscht (baldige Schmerzlinderung versus alltagstaugliche Rückenbelastbarkeit und Autonomie des Patienten (Kügelgen 2003). Nach Kügelgen (2003) fördern alle Anreize des deutschen Gesundheitssystems die Chronifizierung, bei Betroffenen wie bei Leistungsanbietern, darüber hinaus verdrängt sektorales Denken bei den Kostenträgern das Fallmanagement. Die DGSS definiert als Behandlungsziel bei Patienten mit chronischen Schmerzen die Wiederherstellung der objektiven und subjektiven Funktionsfähigkeit mit Steigerung der Kontrollfähigkeit und des Kompetenzgefühls der Betroffenen (Arnold et al., 2009). Heinrich et al. (2011) fanden, dass unter Praxisbedingungen die über den Funktionsfragebogen Hannover – Rücken bestimmbare Funktionskapazität als Maß für den Therapieerfolg brauchbar ist.

Imagama et al. (2011) fanden, dass die Lebensqualität japanischer Männer in mittlerem und höheren Alter durch die sagittale Balance, den lumbalen Lordosewinkel, die spinale Beweglichkeit und die Kraft der Rückenstrecker bestimmt wird.

Haugen et al. (2011) weisen auf ein Problem der Ergebniseinschätzung hin: in einer Multicenterstudie an Patienten mit lumbalen Radikulärsyndromen wurden die Patienten nach 12 Monaten nach den Behandlungsergebnissen befragt. Je nachdem, welches von 5 Erfassungsinstrumenten genutzt wurden, wurden sehr gute Ergebnisse im Bereich zwischen 49 % und 58 % ermittelt. Dass bedeutet, dass auch für Studien zur Ergebnisqualität einheitliche Erfassungsinstrumente eingesetzt werden müssten.

Beale et al. (2011) untersuchten 60 RCT's zur psychologischen Behandlung von chronischen Schmerzen hinsichtlich definierter Zielparameter und fanden nur die Parameter „physische Aktivitäten“ und „emotionales Wohlbefinden“ sowohl für Studienteilnehmer als auch für die klinischen Forscher als wichtige Ergebnisvariablen.

In einer 2008 veröffentlichten Studie zu wichtigen Zielgrößen bei der Behandlung von Patienten mit chronischen Schmerzen zeigte sich nach Befragung von 959 Patienten neben der Schmerzlinderung als wichtigste Aspekte Freude am Leben, emotionales Wohlbefinden, Ermüdung, Schwäche und schlafbezogene Probleme (Turk et al., 2008).

Passik & Weinreb (2000) definierten 4 Zielkriterien für eine Schmerztherapie, die sie mit den 4 A's mnestisch charakterisierten: Analgesie, Aktivitäten des täglichen Lebens, Nebenwirkungen (adverse events) und abweichendes Verhalten bei der Medikamenteneinnahme.

Mehling et al. (2011) kommen zu der Schlussfolgerung, dass die beste Genauigkeit in der Erfassung von Behandlungsergebnissen erreicht wird, wenn eine Wertung der selbstempfundenen Genesung mit Messgrößen von Schmerz und Behinderung kombiniert werden.

Kirschneck et al. (2009) schlagen vor, die ICF-Interventionskategorien zu nutzen, da in der manuellen Medizin kein allgemein anerkanntes Vorgehen für das Assessment existiert (Mein et al., 2001, van de Veen et al., 2005, Niemier et al., 2007).

Die **Zielerreichungsskala** (Kiersuk & Sherman, 1968) ist eine interessante Möglichkeit, Behandlungserfolge anhand der Erreichung von durch den Patienten selbst definierten Zielen (2 – 5) einzuschätzen. Mannion et al. (2010) zeigten bei Patienten mit chronischen Rückenschmerzen, dass etwa 10 % der von Patienten als wichtig eingeschätzten Therapieziele im Roland-Morris-Questionnaire gar nicht erwähnt werden.

Bei der Erfassung und Beurteilung von Therapieergebnissen spielen **Zielvorgaben** eine wichtige Rolle. Leboeuf-Yde et al. (2005) schlagen vor, das Therapieergebnis nach 3 Monaten und nicht nach 12 Monaten zu beurteilen, da eine Verbesserung des Gesundheitszustandes schnell eintritt und andererseits häufig Rezidive auftreten. Farrar (2010) weist darauf hin, dass bei der Messung von Schmerz zwei Zielstellungen existieren, die Erfassung der Stärke des aktuell empfundenen Schmerzes und die Erfassung von **Veränderungen** des Schmerzstatus im Einfluss verschiedener Interventionen über die Zeit; wobei die Messgrößen übereinstimmen können, aber nicht müssen.

Jensen et al. (2012) untersuchten die Sensitivität von Ergebnismessungen mit einer einzelnen Domäne gegen eine Multidomänenerfassung bei chronischen Rückenschmerzen und fanden keine Überlegenheit der aufwändigeren Multidomänenerfassung gegen die Erfassung der Schmerzstärke oder, noch besser, von **PGART** (patient global assessment of response to therapy). Letzteres ist eine einfache Frage: Wie würden Sie Ihr Ansprechen auf die (zu testende) Therapie Ihrer Rückenschmerzen werten: 0 – exzellent (ideales Ergebnis, schmerzfrei), 1 – gut (befriedigendes Ergebnis mit gelegentlichen Schmerzepisoden), 2 – mäßig (akzeptabler Effekt, könnte aber besser sein), 3 schlecht (geringe Wirkung, unbefriedigend) 4 – gar nicht (keine Verbesserung, unwirksam).

Bereits 1997 betonen Flores et al. (1997) die Notwendigkeit, objektivere, stärker quantifizierbare Instrumente für die Erfassung von Behandlungsergebnissen zu schaffen, die die physikalischen, psychologischen und sozioökonomischen Aspekte von Rückenschmerzen beschreiben. In der Schmerzforschung geht heute der Trend dazu, patientenzentrierte Verlaufsparameter zu erfassen, d.h. Verlaufsparameter, die aus Sicht des Patienten wichtig sind (Wittink & Carr, 2008). Hazard et al (2009) entwickelten ein Verfahren, in dem der Patient seine persönlichen Ziele und deren Wichtigkeit definiert, woraus ein Score entwickelt wird, mit dem individuell das Therapieergebnis beurteilt werden kann. Siehe dazu auch Kap. 7.3..

Verbunt et al. (2009) betonen, dass bei Patienten mit muskuloskeletalen Schmerzen die Erfassung des **Niveaus der täglichen Aktivität** wichtiger ist als die Erfassung des Niveaus der schmerzbezogenen Einschränkungen („disability level“). Die WHO definiert dabei physische Aktivität als „jede körperliche Bewegung, die durch Aktivität der Skelettmuskulatur entsteht und einen substantiellen Anstieg über den Ruheenergiebedarf darstellt“ (Caspersen et al., 1985, WHO, 2007). In der ICF – Klassifikation der WHO ist „Aktivität“ als „Ausführung einer Aufgabe oder Aktion durch ein Individuum“ beschrieben. Nach Verbunt et al. (2009) muss dabei zwischen physischer Aktivität (physical activity) und physischer Übung (physical exercise) unterschieden werden: Eine physische Übung ist eine spezielle Form physischer Aktivität, die nicht zufällig entsteht, sondern geplant und

strukturiert mit dem Ziel der Verbesserung oder Erhaltung verschiedener Aspekte der körperlichen Fitness durchgeführt wird (WHO, 2007).

In den meisten Rückenschmerzstudien werden multiple Zielgrößen (outcomes) genutzt. Zu den häufig genutzten Messgrößen gehören Schmerz, Funktion, genereller Gesundheitszustand und Arbeitsfähigkeit (Deyo et al., 1998). Nach van Tulder et al. (2007) gibt es aber nur wenige Studien, die den **MCID (minimale klinisch bedeutende Differenz)** dieser Messwerte bestimmten. Jaeschke et al. (1989) definieren den MCID als die kleinste Differenz in der interessierenden Domäne eines Scores, die vom Patienten als positiv /Verbesserung empfunden wird und die, ohne problematische Nebenwirkungen und exzessiven Kosten aufzuweisen, die Empfehlung einer Therapieänderung nahe legt. In einem Editorial bemängelt Carragee, (2010), dass in der Rückenschmerzforschung das von Jäschke in seiner Definition verlangte Gleichgewicht völlig vernachlässigt wird, indem man bei der Bestimmung des MCID einseitig auf den retrospektiv vom Patienten eingeschätzten Nutzen fokussiert und den Aspekt der Kosten und Risiken völlig außer Acht lässt.

Nach Carragee (2010) sind statistisch bestimmte MCIDs häufig lächerlich klein im Vergleich zu der in Erwähnung gezogenen Maßnahme. Ein aus dem Standardfehler der Messung berechneter MCID von z.B. 1,12 kann zur Empfehlung führen, bei einer großen Zahl von Patienten eine mehrtägige anterior-posteriore spinale Fusion durchzuführen. Eine andere Strategie mit zweifelhafter Validität ist es, den MCID mit Hilfe zeitlich – bis zu 2 Jahre- verschobener Befragungen des Patienten über generelle Zufriedenheit oder „Verbesserung“ zu definieren. Dies ist dann eine sogenannte post – hoc – Klammer – Methode, in der die Klammer keine externe Messgröße ist, sondern eine weitere subjektive und retrospektive Einschätzung durch den Patienten. Es gibt Studien, die mit dieser Methode große Mengen an Patienten dokumentieren, die „zufrieden“ sind, obwohl sie nur geringe oder gar keine mit Standardmethoden nachweisbaren Verbesserungen aufweisen. Wir sind also mit der bekannten menschlichen Natur konfrontiert – eine schlechte Wahl wird selten zugegeben. Carragee (2010) führt das für die USA typische Beispiel eines Patienten an, der die Kostenübernahme von etwa 100 000 \$ für eine große spinale Fusion und die Nachsorge einklagt und jedermann davon erzählt, und dem es nach der Operation trotzdem nicht besser geht. Sollten wir annehmen, dass diese Vorgeschichte keinen Einfluss hat, wenn er gefragt wird, ob dies eine gute Entscheidung war? Slozar et al. (2000) berichten genau dies: Obwohl nur 10 % der von ihnen untersuchten Patienten nach spinaler Fusion eine objektivierbare Verbesserung zeigten, sagten aber 50 % , dass sie diese Operation mit den gleichen Ergebnissen wieder durchführen ließen. Copay et al. (2010) konnten zeigen, dass weniger als die Hälfte ihrer Patienten auf wiederholte Fragen zu Behandlungsergebnissen konsistent antworteten.

Gatchel & Mayer (2010) überprüften eine auf dem Konsensus von Experten berufende Methode der Bestimmung des MCID, wonach eine subjektiv empfundene Verbesserung um 30 % dem MCID darstellt. Der so ermittelte MCID korrelierte wenig mit objektiven Messgrößen wie den weiteren Behandlungsbedarf oder der Arbeitsfähigkeit. Andererseits weisen Glassman & Carron darauf hin, dass bestimmte Ergebnisse, wie z.B. die Arbeitsfähigkeit, von Patienten unter Umständen gar nicht gewollt werden.

Nur wenige Studien bestimmten den MCID für Funktion (Beurskens et al., 1996, Riddle et al., 1998, Stratford et al., 1998, Fritz & Irrgang, 2001, Davidson & Keating., 2002, van der Roer et al., 2006) und Schmerz (Hägg et al., 2003, Hayden et al., 2005) bei Rückenschmerzpatienten.

Spratt (2009) untersuchte anhand der Skala „physische Funktion“ der SF-36 Ergebnisse der SPORT – Studie und fand, dass die MCID – Werte einer 30- prozentigen Verbesserung bei Patienten mit einem Score unter 50 substantiell weniger veränderungssensitiv waren , bei Scores über 50 aber den Werten der Standardfehlerbestimmung nach der klassischen Testtheorie entsprechen.

Zur Einschätzung der Behinderung durch Rückenschmerzen werden vor allem in englischsprachigen Publikationen das RMDQ (Roland and Morris Disability Questionnaire (Roland & Morris, 1983)) und das ODI (Oswestry Low Back Disability Questionnaire (Fairbank et al., 1980)) eingesetzt, s. Kapitel 4. Als MCID definierten Bombardier et al. (2001) für das RMDQ einen Wert von 2,5 und mehr, Ostelo & de Vet (2005) für ODI 10 Punkte.

Carragee & Cheng (2010) führen ein „**minimal akzeptables Behandlungsergebnis**“ ein, indem intern – subjektive und extern-objektive Schwellwerte für die Feststellung eines erfolgreichen Behandlungsergebnisses gefordert werden: für die Schmerzreduktion 4 oder mehr auf einer 10 – Punktskala, für funktionelle Einschränkungen 20 oder mehr Punkte auf dem Oswestry Disability Index. Außerdem wurden bei 90 % der untersuchten Patienten eine weitere Arbeitsunfähigkeit und eine weiter erforderliche Opiattherapie als ungenügendes Behandlungsergebnis gewertet. In dieser großen Patientenpopulation sahen nur wenige Patienten den mit üblichen Methoden bestimmten MCID für Schmerz und Funktion als akzeptables Behandlungsziel an, eine Operation mit dem Ziel, nur diesen MCID zu erreichen wurde abgelehnt.

In Holland definieren Jellema et al. (2006) eine klinisch bedeutsame Verbesserung als Reduktion der funktionellen Beeinträchtigung (gemessen durch das Roland-Morris Disability Questionnaire) um mindesten 30 % und eine subjektive Verbesserung auf einer siebenstufigen Likert-Skala (sehr viel / viel / etwas besser, unverändert, etwas / viel / sehr viel schlechter).

Mesrian et al. (2007) aus Heidelberg empfehlen eine dreistufige globale Selbsteinschätzungsskala: schlechter / unverändert / besser und verweisen darauf, dass bei Nutzung der VAS zur Erfolgskontrolle eine signifikante Veränderung nicht automatisch klinisch relevant ist (Deyo & Patrick, 1995, Beaton et al., 2001, Bombardier et al., 2001); bei genügend großer Stichprobe kann jede kleine unbedeutende Veränderung statistische Relevanz gewinnen. Bei Patienten mit chronischen Rückenschmerzen wird bei Nutzung der VAS eine Verminderung der Schmerzstärke von 18/100 bzw. 2/10 als klinisch relevant beschrieben (Farrar et al., 2000, 2001, Hägg et al., 2001), Mesrian et al. (2007) fanden 25/100 als relevant, für eine Verschlimmerung der Schmerzen genügten allerdings schon 9/100. In einem Kommentar zu dieser Arbeit rügt Zenz (2007), dass nicht definiert wurde, ob dem Patienten das Extrem als „stärkste Schmerzen“ oder als „stärkste vorstellbare Schmerzen“ vorgestellt wurde, was von zentraler Bedeutung sei

Lauridsen et al. (2007) bemängeln die verschiedenen Varianten, das Ergebnis einer Rückenschmerztherapie darzustellen, häufig sind Forscher und Kliniker nicht mit den im Einzelfall gebrauchten Verfahren vertraut, was die Interpretation von Ergebnissen erschwert. Als Lösung dieses Sullivan & Ballantyne (2016) warnen, dass die einseitige Orientierung auf eine reine Schmerzlinderung durch Opioide das falsche Ziel für eine Behandlung ist, zu einer Auswahl der falschen Patienten für eine hochpotente Therapie führt und unser Verständnis von chronischem Schmerz verzögert. Unser Ziel in der Behandlung von Patienten muss es dagegen sein, nicht die Schmerzstärke zu reduzieren, sondern die Lebensqualität zu steigern.

Dilemmas schlagen die Autoren vor,

1. dem Patienten eine klare einführende Frage zu stellen, die sich auf Veränderungen in dem interessierenden Bereich konzentriert und ein klares Zeitfenster definiert,
2. eine siebenstufige Veränderungsskala zu benutzen mit Antwortoptionen, die kurz, klar und miteinander logisch verknüpft sind (Wie würden Sie ihre Rückenschmerzen heute im Vergleich zum Zeitpunkt des Beginns der Therapie beschreiben: 1. viel besser, 2. besser, 3. etwas besser, 4. unverändert, 5. etwas schlechter, 6. schlechter, 7. viel schlechter)

3. eine klare Definition festzulegen, wann sich die Beschwerden eines Patienten gebessert haben und wann sie als unverändert gelten.

Außerdem schlagen Lauridsen et al. (2007) eine globale Einschätzung der Entwicklung durch den Patienten vor, z.B. durch eine numerische Skala von 0 bis 10, auf der der Patient markiert, für wie bedeutend er die erreichten Veränderungen hält, 0 = unbedeutend, 10 = sehr bedeutend.

Bombardier (2000) identifizierte 5 Bereiche, die zur Erfassung der Behandlungsergebnisse essenziell sind: Spezifische Rückenfunktionen, allgemeiner Gesundheitszustand, Schmerz, Arbeitsunfähigkeit und Patientenzufriedenheit (Liddle et al., 2004). Nach Liddle et al. (2004) wird diese Komplexität der Ergebniserfassung in Studien jedoch kaum angewandt, am häufigsten werden rückenspezifische Funktion und Arbeitsfähigkeit zur Beurteilung des Behandlungsergebnisses genutzt. Dionne et al. (1999) weisen auf die Bedeutung der Erfassung von Schmerz, Funktion und Arbeitsfähigkeit zur Beurteilung des Behandlungsergebnisses hin.

In einem Konsensuspapier legten Turk et al. (2003) folgende Kerndomainen für die Erfassung von Behandlungsergebnissen fest: 1. Schmerz, 2. körperliche Funktionen (physical functioning), 3. emotionale Funktionen, 4. Zufriedenheit mit der Therapie und Grad der subjektiven Verbesserung, 5. Nebenwirkungen und Symptome und 6. Teilnehmerverhalten, d.h. Studienabbrüche und deren Gründe.

Sigl et al. (2006) untersuchten, inwieweit in verschiedenen Instrumenten zur Ergebnismessungen Elemente der ICF (Internationale Klassifikation von Funktion, Behinderung und Gesundheit) abgebildet werden und fanden die ICF in guter Übereinstimmung mit Items der NASS, Oswestry (ODI) und Roland-Morris – Skalen.

Bei der Interpretation von Metaanalysen zu Therapieergebnissen bei chronischen Rückenschmerzen ist Vorsicht geboten, da unterschiedliche Ergebnisparameter nach Untersuchungen von Morris et al (2015) zwar jeder für sich valide ist, es aber keine Übereinstimmung zwischen den Parametern gibt. Eine Vermischung der Parameter der schmerzbedingten Behinderung wird ausdrücklich nicht empfohlen.

Eine Möglichkeit des Klassifizierung der Behandlungsergebnisse stellt der **CPG (chronic pain grade**, von Korff et al., 1992) dar: CPG I: wenig Schmerz und Behinderung, II: wenig Schmerz aber starke Behinderung, III: stark behindernder, aber nur mäßig einschränkender Rückenschmerzen und IV: Rückenschmerzen stark behindernd und einschränkend.

Für operative Verfahren gilt, dass die Behandlung als erfolgreich gewertet werden kann, wenn 2 Jahre postoperativ der LBP – VAS niedriger als 3 ist, des Oswestry – Index unter 16 liegt, der Patient wieder voll arbeitsfähig ist und keine Schmerzmittel gebraucht werden. Als „minimal akzeptables Ergebnis“ (Carragee & Alamin, 2003) wird gewertet, wenn der Patient der Operation nochmals zustimmen würde (Carragee et al., 2006b).

In einem Kommentar zu einer Arbeit zur Neurostimulation des M. multifidus (Gilligan et al., 2021) wirft Atlas (2021) einige Fragen zur Bewertung der Methode auf. Sind die Einschlusskriterien für die Therapie mit dieser Methode in der Praxis nachvollziehbar? Hat die Methode Einfluss auf die rückenschmerzbedingte Behinderung und die Wiedererlangung der Arbeitsfähigkeit? Kann der Opioidverbrauch reduziert werden? Wie lange muss mit der Methode therapiert werden? Gibt es limitierende Nebenwirkungen?

Niederer et al. (2021) untersuchten 42 verschiedene Kriterien für Behandlungsergebnisse auf ihre Relevanz und fanden 25 potentiell bedeutsame Faktoren im Kontext von selbstberichteten Ergebnissen (Arbeitsstunden pro Woche, gesundheitsbezogene Lebensqualität, chronische Schmerzstärke, chronische Behinderung durch Schmerzen, Arbeitsunfähigkeitstage und

Trainingstage/Woche), objektive funktionelle Messergebnisse (Bewegungsausmaß Extension/Flexion, dynamische Balance, verschiedene Sprung-Test-Parameter, Bewegungsausmaß in Flexion, maximale Extensionskraft, Kraftausdauerwerte) und Körpercharakteristika (BMI, Herzfrequenz).

Ältere Arbeiten siehe Gesamtliteraturverzeichnis

- Atlas,S.J.  
Medical interventions have not solved the epidemic of chronic low back pain: will neurostimulation of the multifidus muscle be different? (Commentary)  
*Pain* 162 (2021)2464 – 5  
[doi.org/10.1097/j.pain.0000000000002259](https://doi.org/10.1097/j.pain.0000000000002259)
- Chiarotto,A.  
Boers,M., Deyo,R.A., Buchbinder,R., Corbin,T.P., Costa,L.O.P., Foster,N.E., Grotle,M., Koes,B.E., Kovacs,F.M., Lin,C.W.C., Maher,C.G., Pearson,A.M., Peul,W.C., Schoene,M.L., Turk,D.C., van Tulder,M.W., Terwee,C.B., Ostelo,R.W.  
Core outcome measurement instruments for clinical trials in nonspecific low back pain  
*Pain* 159 (2018)481 - 95
- Chiarotto,A.  
Terwee,C.B., Ostelo,R.W.  
Choosing the right outcome measurement instruments for patients with low back pain  
*Best Pract Res Clin Rheum* 30 (2016)1003 – 20  
[doi.org/10.1016/j.berh.2017.07.001](https://doi.org/10.1016/j.berh.2017.07.001)
- Gilligan,C.  
Volschenk,W., Russo,M., Green,M., Gilmore,C., Mehta,V., Deckers,K., de Smedt,K und weitere 25 Autoren  
An implantable restorative-neurostimulator for refractory mechanical chronic low back pain: a randomized sham-controlled clinical trial  
*Pain* 162 (2021)2486 – 98  
[doi.org/10.1097/j.pain.2021.0000000000002258](https://doi.org/10.1097/j.pain.2021.0000000000002258)
- Henschke.N.  
van Ernst,A., Froud,R., Ostelo,R.W.G.  
Responder analysis in randomized controlled trials for chronic low back pain: an overview of currently used methods  
*Eur Spine J* 23 (2014)772 – 8 doi.org/10.1007/s00586-013-3155-0
- Jensen,M.P.  
Schnitzer,T.J., Wang,H., Smugar,S.S., Peloso,P.M., Gammaitoni,A.

Sensitivity of single-domain versus multiple-domain outcome measures to identify responders in chronic low back pain.  
Pooled analysis of 2 placebo-controlled trials of etoricoxib

Clin J Pain 28 (2012)1 – 7

Kaiser,U.

Kopkow,C., Deckert,S., Neustadt,K., Jacobi,L., Cameron,P., de Angelis,V., Apfelbacher,C. und weitere 25 Autoren

Developing a core outcome domain set to assessing effectiveness of interdisciplinary multimodal pain therapy: the VAPAIN consensus statement on core outcome domains

Pain 159 (2018)673 - 83

Marin,T.J.

Furlan,A.D., Bombardier,C., van Tulder,M.

Fifteen years of the Cochrane Back Review Group

Spine 38 (2013)2057 - 63

Morris,T.

Hee,S.W., Stallard,N., Underwood,M., Patel,S.

Can we convert between outcome measures of disability for chronic low back pain?

Spine 40 (2015)734 - 9

Niederer,D.

Engel,T., Pfeifer,A.C., Arampatzis,A., Beck,H., Wippert,P.M., Schiltenwolf,M., Mayer,F.

Which functional outcomes can be measured in low back pain trials and therapies? A prospective 2-year factor-, cluster-, and reliability-multicenter analysis on 42 variables in 1049 individuals

Spine 46 (2021)1495 – 1508  
[doi.org/10.1097/BRS.0000000000004028](https://doi.org/10.1097/BRS.0000000000004028)

Niemier,K.

Noll-Hussong,M., Pfingsten,M.

Schmerz oder Funktion, was steht im Mittelpunkt der Schmerztherapie?

Schmerz 26 (2012)Suppl.1: 9 - 10

Stienen,M.N.

Ho,A.L., Staartjes,V.E., Maldaner,N., Veeravagu,A., Desai,A., Gautschi,O., Bellut,D., Regli,L., Ratliff,J.K., Park,J.

Objective measures of functional impairment for degenerative diseases of the lumbar spine: a systematic review of the literature

Spine J 19 (2019)1276 - 93

Sullivan,M.D.

Ballantyne,J.C.

Must we reduce pain intensity to treat chronic pain?

Pain 157 (2016)65 - 9

### **8.3 Patientenzufriedenheit**

Die Erfassung der Patientenzufriedenheit mit der Behandlung ist sehr wichtig (Liebenson & Yoemans, 1997), aber noch wenig untersucht (Liddle et al., 2004). Auf dem 9. internationalen Forum zur Forschung über den Rückenschmerz in der Grundversorgung 2007 in Spanien wurde vorgeschlagen, von Patienten definierte Behandlungsziele als zu erfassende Behandlungserfolgsparameter aufzunehmen (Cherkin et al. 2009).

In den USA zeichnet sich zu Beginn der 10'er Jahre des 21en Jahrhunderts der Trend ab, dass Kostenträger die Erstattung von Leistungen in der Rückenschmerzbehandlung von deren Qualität abhängig machen. Als allgemeine, billige und gut verfügbare Daten dafür gelten nach Truumees et al. (2013) Prozessgrößen der Behandlung und Scores der Patientenzufriedenheit. Zu Ersteren gehören beispielsweise die Dauer stationärer Therapie oder die Rate an Wundinfektionen bei Wirbelsäulenoperationen. Die Patientenzufriedenheit wird definiert als die individuelle Einschätzung, inwieweit die Behandlung den individuellen Erwartungen und Vorstellungen entsprochen hat (Brennan, 1995) oder als der Grad, in dem der Patient fühlt, eine hochqualitative Behandlung erhalten zu haben (Chow et al., 2009).

Turk & Wilson (2009) beschäftigen sich mit der Beziehung zwischen Schmerz und Leiden und belegen anhand der Literatur, dass der Begriff „Leiden“ unscharf definiert ist und oft synonym verwendet wird. Schmerz an sich scheint nicht auszureichen, um Leiden zu erzeugen, es scheint eher, dass die Interpretation der Symptome, des Einflusses auf das physische und emotionelle Funktionieren, die Ziele („meaning“) und die Coping-Ressourcen dafür von Bedeutung sind. Eine Behandlung kann das Leiden reduzieren, ohne den Schmerz zu lindern und umgekehrt (Turk & Wilson, 2009)

Hazard et al. (2012) verglichen die Schmerzreduktion und Erreichung persönlicher Ziele in der Rückenschmerztherapie mit der Patientenzufriedenheit und fanden die Erreichung der persönlichen Zielstellung wesentlich wichtiger dafür.

Die Zufriedenheit des Patienten mit der Behandlung steigt, wenn der Therapeut die Sichtweise des Patienten berücksichtigt (Hall et al., 1988) genauso wie die Compliance des Patienten mit dem Behandlungsprogramm und seinem Behandler (Becker, 1985, Kaplan et al., 1989).

Trentman et al. (2013) untersuchten, welche Faktoren die Patientenzufriedenheit an einer akademischen Schmerzklinik bestimmen und fanden, dass die Zufriedenheit der Patienten mit der Behandlungsqualität durch Gründlichkeit, Zuhören und die mit dem Behandler verbrachte Zeit bestimmt wird, während die generelle Einschätzung des Therapeuten mit sehr gut bzw. exzellent durch die Faktoren Gründlichkeit, Zuhören, Pünktlichkeit und klare Instruktionen bestimmt wurden.

Bible et al. (2015) befragten in den USA 200 Patienten innerhalb von 3 Wochen nach dem Erstkontakt mit einem Wirbelsäulentherapeuten und fanden eine verminderte Patientenzufriedenheit verbunden mit jüngerem Alter, geringerer Bildung und Rauchen. Negativ mit der Patientenzufriedenheit waren

auch männliches Geschlecht und eine Behandlung als Arbeitsunfall (treated under worker's compensation claim), während Familienstatus, Beschäftigungsstatus, mentaler Zustand, räumliche Entfernung zum Therapeuten, Schmerzcharakteristika, vorangegangene Therapien und eine aktuelle Opiateinnahme keinen signifikanten Einfluss auf die Patientenzufriedenheit hatten.

Hopayan & Notley untersuchten dieses Thema in einem systematischen Review und fanden, dass es darauf ankommt, wie die Behandlung abläuft: wichtig ist, dem Patienten das Gefühl einer gründlichen, vollständigen Untersuchung zu vermitteln, die individuellen Verhältnisse des Patienten zu berücksichtigen, ein gutes Verhältnis zum Patienten aufzubauen, Empathie zu zeigen und Informationen mit dem Patienten zu teilen. Kosten und Wartezeiten spielen dagegen eine geringere Rolle.

Eine niedrigere Zufriedenheit ist nach Truumees (2013) verbunden mit Depression und psychologischem Distress, weiblichem Geschlecht, niedrigerem Bildungsniveau, Ledigkeit, geringer sozialer Unterstützung, einem Haushaltseinkommen von weniger als 20000 US-\$, Arbeitslosigkeit, fehlendem Versicherungsschutz, anderen ernsthafteren Gesundheitsproblemen, einem schlechteren Gesundheitsverhalten (Rauchen und Alkoholkonsum), Erschöpfung und Schlafstörungen, geringer Selbstwirksamkeit (low self-efficacy), schlechten Schmerzcopingstrategien und Somatisierung.

In Australien untersuchten Henschke et al. (2013), welche Faktoren die Zufriedenheit von Patienten mit akuten Rückenschmerzen ein Jahr nach Diagnosestellung bestimmen und fanden, dass die Zufriedenheit mit den Behandlungsergebnissen durch die aktuelle Schmerzstärke stark bestimmt wird. Bei vergleichbarer Schmerzintensität waren diejenigen weniger zufrieden, die außerhalb Australiens geboren wurden, älter waren und eine stärkere Depression aufwiesen.

Die Patientenzufriedenheit selbst ist multidimensional und stellt eine Reaktion auf die Kombination von Behandlungsumfeld, Behandlungsprozess und Behandlungsergebnis dar (Pascoe, 1983). Eine Studie in der Allgemeinmedizin von Williams et al. (1995) ergab, dass die Proportion der erfüllten Erwartungen signifikant mit dem berichteten Grad der Zufriedenheit korrelierte. Die Patientenerwartungen und ihre Erfüllung haben einen therapeutischen Effekt (McPherson & Britton, 2001). Über die Beziehung von Patientenzufriedenheit und anderen Ergebnisparametern ist gerade bei LBP noch wenig bekannt (Breen & Breen, 2003). Hazard et al. (1994) untersuchten 65 Patienten mit chronischen Rückenschmerzen nach einem Rehabilitationsprogramm und fanden nach 5 Jahren nur schwache Korrelationen zwischen Patientenzufriedenheit und Schmerz, Behinderung und funktionellen Parametern. Kovacs et al. (2004, 2005) fanden, dass die Lebensqualität durch die schmerzbedingten Einschränkungen bestimmt wird, diese wiederum von der Dauer der Schmerzen, wobei die Stärke der Schmerzen unerheblich waren.

Thorne & Morley (2009) untersuchten, wie groß die Veränderungen des Ausgangszustandes sein müssen, damit Patienten mit Schmerzen das Behandlungsergebnis als akzeptabel bezeichnen undstellten dabei fest, dass die Erwartungen an die Schmerzlinderung mit 44 % – 75 % größer sind, als die in der Forschung als klinisch signifikant betrachteten Veränderungen von 30 % – 50 %. In einem Editorial zu dieser Arbeit kommen Vowles & Robinson (2009) zu der Schlussfolgerung, dass für den Patienten weniger die allgemeine Schmerzreduktion von Bedeutung ist, als der Einfluss, den der Schmerz auf den funktionellen Status des Patienten hat. Da Thorne & Morley weder einen Absolutwert der zur Erreichung der Akzeptanz erforderlichen Schmerzlinderung ermitteln konnten noch eine zur aktuellen Stärke des Schmerzes proportionale Schmerzlinderung, schlagen Vowles & Robinson vor, Behandlungsziele zu individualisieren und Funktionsbereiche zu definieren, in denen Therapieziele festgelegt werden können.

Eine Aufmerksamkeitsrichtung auf den Schmerz kann helfen, den Schmerz besser zu verstehen, eine Fokussierung auf den Schmerz kann aber auch negative Effekte in der Schmerzverarbeitung haben,

Kongstedt et al. (2021) fanden keine negativen Auswirkungen, wenn Patienten mit Rückenschmerzen wöchentlich nach Schmerzfrequenz und Schmerzintensität befragt wurden.

Ein Ausdruck der Patientenzufriedenheit ist die gefühlte Lebensqualität. Die Lebensqualität wird durch die WHO als die Wahrnehmung seiner Position im Leben durch ein Individuum, im Kontext der Kultur und des Wertesystems, in dem es lebt, und in Beziehung zu seinen Zielen, Erwartungen, Standards und Interessen (WHOQOL, 1998). Mason et al. (2008) fanden bei Rückenschmerzpatienten, dass die Lebensqualität von der aktuellen Schmerzstärke und der Akzeptanz der Schmerzen abhängt.

Bei Patienten nach Wirbelsäulenoperationen stellten Godil et al. (2013) allerdings fest, dass die Lebensqualität und das Maß der körperlichen Beeinträchtigung die Patientenzufriedenheit nicht gut vorhersagen.

Hush et al. (2009) untersuchten, was Patienten mit Rückenschmerzen unter dem Begriff „wieder gesund“ verstehen und fanden eine Reihe von Faktoren, die in den Domänen „Symptomminderung“, „Fähigkeit, selbst definierte Aktivitäten durchführen zu können“ und „Erreichen einer akzeptablen Lebensqualität“ zusammengefasst werden können. In einem Kommentar zu dieser Untersuchung betont Ammer (2009), dass zur Patientenzufriedenheit auch das Einbezogensein des Patienten in relevante Entscheidungen gehört.

Es ist wichtig, zwischen der Zufriedenheit des Patienten mit der Behandlung selbst und der Zufriedenheit mit den Behandlungsergebnissen zu unterscheiden, da bei Patienten mit chronischen Schmerzen die Zufriedenheit mit der Behandlung nicht eng mit der Schmerzlinderung verbunden ist, sondern von interpersonellen Beziehungen zwischen Patient und Therapeut abhängt (Hirsh et al., 2005). Petrie et al. (2005) zeigten, dass für Patienten, die einen Schmerztherapeuten aufsuchen, die Erklärung ihres Schmerzproblems genauso wichtig ist, wie die Behandlung oder die Schmerzlinderung. Bei Patienten mit chronische Rückenschmerzen war die Zufriedenheit mit der letzten Konsultation in einer Untersuchung aus den USA (Wallace et al., 2009) unter anderem davon abhängig, ob den Patienten Opiate verordnet wurden (positive Korrelation), was in einem Kommentar für Ballantyne (2009) aber eine Menge ungeklärter Fragen aufwirft.

Butler & Johnson (2008) fanden, dass Arbeiter in den USA mit der Behandlung durch Chirurgen, Chiropraktoren und Physiotherapeuten zufriedener waren als mit der Behandlung durch Hausärzte (MDs), wobei die Zuordnung zur Gruppe der MDs recht diffus erscheint.

Haldeman (2012) kritisiert den Einsatz der Patientenzufriedenheit als Zielparameter, da diese zu unspezifisch sei und die Zufriedenheit mit der Erreichung von Zielen allgemein gültig sei. Außerdem gäbe es Studien, die die Zufriedenheit der Patienten mit allen möglichen Therapieformen Nachweise: Medikamente, Krankengymnastik, Manipulation, Akupunktur, die meisten Injektionsverfahren, Bandscheibenresektion, Fusionen und Arthrodesen (Choi et al., 2011, Marquardt et al., 2011), Rückenmarkstimulation (Sears et al., 2011), Hausübungs-programme wie auch für Yoga, Tai Chi, spirituelles Heilen und frei verkäufliche Nahrungsergänzungsmittel (Ho et al., 2009). Außerdem sei das funktionelle Ergebnis nicht der einzige Faktor, der die Zufriedenheit des Patienten bestimmt. Dazu gehören auch das Arzt – Patienten – Verhältnis (Farin et al., 2012), die Kommunikationsweise (Vangoonsveld & Linton, 2012) und die Erwartungen und Vorlieben des Patienten (Kalauokalani et al., 2001). Aus diesen Gründen fordert Haldeman (2012) objektive, reproduzierbare Verlaufspараметer.

Smuck (2012) hält in einem Kommentar zur gleichen Arbeit dagegen. Letztlich sei die MCID auch nur eine auf subjektiven Empfindungen und Vorstellungen des Patienten beruhende Größe. Der Charme in dem Ansatz von Hazard et al. (2012) sei eben, dass es sich dabei nicht um eine rein subjektive

Erfassung der allgemeinen Patientenzufriedenheit am Ende einer Behandlung handelt, sondern um die Erfassung, inwieweit vorher (!) festgelegte Erwartungen erfüllt wurden.

Ältere Arbeiten siehe Gesamtliteraturverzeichnis

- |              |  |
|--------------|--|
| Bible,J.E.   | Kay,H.E., Shau,D.N., O'Neill,K.R., Segebarth,P.B., Devin,C.J.  |
|              | What patient characteristics could potentially affect patient satisfaction scores during spinae clinic?  |
|              | Spine 40 (2015)1039 - 44   |
| Farin,E.     | Gramm,L., Schmidt,E.   |
|              | The patient-physician relationship in patients with chronic low back pain as a predictor of outcomes after rehabilitation  |
|              | J Behav Med (2012) epub ahaed of print (so zit. Haldeman, 2012)  |
| Godil,S.S.   | Parker,S.L., Zuckerman,S.L., Mendenhall,S.K., Devin,C.J., Asher,A.L., McGirt,M.J.  |
|              | Determining the quality and effectiveness of surgical spine care: patient satisfaction is not a valid proxy  |
|              | Spine J 13 (2013)1006 - 12   |
| Hazard,R.G.  | Spratt,K.F., McDonough,C.M., Olson,C.M., Ossen,E.S., Hartmann,E.M., Carlson,R.J., LaVoie,J.  |
|              | Patient centered evaluation of outcomes from rehabilitation for chronic disabling spinal disorders: the impact of personal goal achievements on patient satisfaction |
|              | Spine J 12 (2012)1132 - 7  |
| Henschke,N.  | Wouda,L., Maher,C.G., Hush,J.M., van Tulder,M.W.   |
|              | Determinants of patient satisfaction 1 year after presenting to primary care with acute low back pain  |
|              | Clin J Pain 29 (2013)512 - 7   |
| Hopayan,K.   | Notley,C.  |
|              | A systematic review of low back pain and sciatica patients' expectations and experiences of health care  |
|              | Spine J 14 (2014)1769 - 80   |
| Kongstedt,A. | Secher Jensen,T., Doktor,K., Hestbeak,L.   |
|              | Effects of weekly pain monitoring on back pain outcomes: a nonrandomized controlled study  |

Chiropr Man Ther 29 (2021)37 doi.org/10.1186/s12998-021-00393-2

Smuck,M.	Commentary: More or less satisfied?  Spine J 12 (2012)1140 - 1
Trentman,T.L.	Cornidez,E.G., Wilshusen,L.L., Chang,Y.H.H., Seamans,D.P., Rosenfeld,D.M., Freeman,J.A., Chien,J.J.  Patient satisfaction in an academic pain clinic  Pain Pract 13 (2013)372 – 9
Truumees,E.	Commentary: Appropriate use of satisfaction scores in spine care  Spine J 13 (2013)1013 - 6
Vangroenveld,K.L.	Linton,S.J.  The effect of validating and invalidating communication on satisfaction, pain and affect in nurses suffering from low back pain during a semistructured interview  Eur Pain J 16 (2012)239 – 46

#### 8.4. Prognose

Der Verlauf von Rückenschmerzen wird durch das Zusammenspiel biologischer, psychologischer und sozialer Faktoren bestimmt (Waddel, 1998), wobei rechtliche, kulturelle und soziodemographische Unterschiede den Verlauf beeinflussen können und beachtet werden sollten (Kovac et al., 2005, 2007). Nach Enthoven et al. (2006) wurden in den letzten Jahren eine Reihe von Faktoren identifiziert, die die Prognose bzw. den Verlauf von Rückenschmerzen bei Patienten in allgemeinmedizinischer Behandlung beeinflussen (Huskisson, 1974, Kleenerman et al., 1995, Skargreen & Öberg, 1998, Thomas et al., 1999, Enthoven et al., 2003). Diese Faktoren können nach der International Classification of Functioning, Disability and Health (ICF) der WHO (WHO, 2001, Cieza et al., 2004) in Körperfunktion/Struktur (B/F), Aktivität und Teilnahme (participation; A/P), Umwelt (environmental, E) und persönliche Faktoren (personal, P) unterteilt werden. Andere Studien haben die für die Prognose akuter Rückenschmerzen maßgeblichen Faktoren in demographische Faktoren, klinische (einschließlich Stärke und Ausdehnung der Schmerzen und der Beeinträchtigung durch den Schmerz) und psychosoziale Faktoren (einschließlich Copingstil und Arbeitsfaktoren) unterteilt (Thomas et al., 1999, Coste et al., 2004, Jones et al., 2006).

Turner et al. (2013) kommen zu der Schlussfolgerung, dass für die Praktikabilität unter Praxisbedingungen ein kurzes Screeninginstrument nötig ist, das Daten enthalten sollte zu (1) Schmerzlokalisation (diffuseness), (2) Funktionseinschränkungen (Schmerzintensität und Aktivitätsbeeinträchtigung) und (3) zeitlichen Eigenschaften (pain persistence).

Pfingsten et al. (1997a) fanden in einer Studie zur multimodalen Therapie, dass die vor der Behandlung abgegebene Selbsteinschätzung der Patienten über ihre vermutliche Arbeitsfähigkeit, die

Rentensituation sowie die Dauer der Arbeitsunfähigkeit die Rückkehr an den Arbeitsplatz sowie die subjektive Schmerzreduktion am Ende der Behandlung am besten voraussagen konnten.

Auf dem 10. Internationalen Forum für Forschung über Rückenschmerzen in der Primärversorgung 2010 in Boston wurde nach Pransky et al. (2011) festgestellt, dass in einem sozialen Kontext bestehende soziale Faktoren wie die Arbeitsumgebung, Entschädigungssysteme und Interessen von Vereinigungen oder Gewerkschaften („union concerns“) die Wiederaufnahme der Arbeit stärker beeinflussen als die häufig untersuchten psychosozialen Faktoren. Aylward & Drake (zit. bei Pransky et al., 2011) schlugen auf diesem Meeting vor, aus den Erfahrungen mit der Reintegration von Patienten mit schweren chronischen mentalen Problemen zu lernen und eher auf Programme zu verbessertem Coping und verbesserter Adaptation als auf medizinische Behandlungen und klinikbasiertem Training zu setzen. Das bedeutet auch, die Kontrolle des Prozesses von Anbietern medizinischer und rehabilitativer Leistungen auf die Patienten und deren Arbeitgeber zu verlagern, die Bedingungen zu entmedikalisieren und eine schnelle Wiederaufnahme einer Form von Arbeit anzustreben (Drake & Bond, 2008).

In einer Analyse dreier randomisierter klinischer Studien kommen Helmhout et al. (2010) zu dem Ergebnis, dass die Genesung bzw. eine funktionelle Verbesserung von Rückenschmerzen negativ von einer starken initialen Beeinträchtigung und einer stark ausgeprägten Fear-Avoidance – Einstellung beeinflusst wird. Von sozialen Faktoren wird die Heilung kurzfristig durch die Unterstützung Vorgesetzter und langfristig von der Unterstützung durch Kollegen positiv beeinflusst.

Nach Mäntyselkä (2012) wird das Risiko, wegen einer musculoskelettalen Erkrankung erwerbsunfähig zu werden durch die Symptomlast (symptom burden), die Summe der Gesundheitsprobleme, Komorbidität, niedrigen sozioökonomischen Status und niedriges Bildungsniveau gesteigert.

Interessanterweise ergaben Untersuchungen von Mlekusch et al. (2013), dass der Verlauf von Rücken- und auch Nackenschmerzen nicht von einer zentralen Hypersensitivität oder der Aktivität der konditionierten Schmerzmodulation (CPM, entspricht der DNIC,,s. Kap. 2.2.4.2) beeinflusst wird.

Grovle et al. (2013) fanden in Norwegen, dass die Wiedererreichung der Arbeitsfähigkeit bei Patienten mit lumbalen Radikulärsyndromen mit einer Reihe von Variablen verbunden ist. Dazu gehören Alter, allgemeiner Gesundheitszustand, Radikulärsyndrome in der Vorgeschichte, die Dauer der aktuellen Episode, der anfängliche Leidensdruck, Fear-Avoidance, Rückenschmerzen und Ergebnisse des Lasegue-Tests. Interessanterweise hatte eine operative Therapie keine signifikante Verbindung mit der Arbeitsfähigkeit.

Vavrek et al. (2015) fanden in einem RCT mit 400 Patienten mit chronischen Rückenschmerzen, dass sich ein Behandlungserfolg von spinalen Manipulationen nicht aus vor der Behandlung erhobenen Daten abschätzen lässt, zu denen unter anderen auch die Schmerzintensität, der psychische Zustand oder soziodemographische Daten gehörten. Am besten ließ sich der weitere Krankheitsverlauf aus der Stärke der Schmerzen nach 6 Wochen Behandlung abschätzen.

Der Einfluss des **Versicherungs-/Entschädigungsstatus** wird beispielhaft in einer Metaanalyse von Harris et al. (2005) deutlich: Die Ergebnisse üblicher Wirbelsäulen-chirurgischer Eingriffe waren bei arbeits-/berufsbedingten Schäden, für die in den USA der Arbeitgeber haftet, viermal schlechter als bei Eingriffen zu Lasten des Versicherungssystems. Unter diesem Aspekt ist die Tendenz der wissenschaftlichen Literatur, Rückenerkrankungen nicht mehr als Folge berufsbedingter Beanspruchung durch wiederholtes Beugen oder Verdrehen des Rumpfes (Wie et al., 2010) oder durch Arbeiten in Zwangshaltungen (Roffey et al., 2010) zumindest für Arbeitnehmer in den USA von erheblicher Bedeutung (Talmage, 2010).

Cancelliere et al. (2016) führten eine „best evidence synthesis of systemic reviews“ zu den Faktoren durch, die die Wiederaufnahme der Erwerbstätigkeit nach Verletzung oder Erkrankung beeinflussen, wobei sich die Mehrzahl der Studien mit muskuloskelettalen Erkrankungen, davon wiederum am häufigsten mit Wirbelsäulenproblem beschäftigten. Generell waren höhere Bildung, höherer sozioökonomischer Status, höhere Selbsteffizienz, optimistische Erwartungen hinsichtlich Genesung und Arbeitsfähigkeit, geringere Schwere von Krankheit/Verletzung, eine Koordination der Wiederaufnahme der Berufstätigkeit und multidisziplinäre Interventionen mit kürzeren Arbeitsunfähigkeitsdauern verbunden.

Burns et al. (2018) untersuchten, welche Patienten mit Rückenschmerzen als primärem Problem am meisten von einer physikalischen Therapie profitieren und fanden die besten Behandlungsergebnisse hinsichtlich Schmerz und Funktionseinschränkung bei Patienten mit einem gleichzeitigen Hüftproblem, mit einem BMI < 25,4 und einem NRS von 4 und weniger.

### Risikofaktoren

In einer prospektiven Studie über 2 Jahre fanden Dionne et al. (1999), dass zwischen Schmerzstärke, funktionellen Beeinträchtigungen als Risikofaktoren nur eine mäßige Übereinstimmung besteht, beide Messgrößen sind nicht gegeneinander austauschbar. Beide Faktoren sind mäßig mit der Wiederaufnahme der Berufstätigkeit assoziiert, interessanterweise aber nicht mit einer Verschlechterung der Arbeitsfähigkeit (Dionne et al., 1999).

Dunn & Croft (2006b) geben folgende Faktoren als Indikatoren einer schlechten Prognose an: starke Behinderung, hohe Schmerzintensität, Beinschmerz, Schmerzen im Oberkörper, quälender Schmerzcharakter, Katastrophisieren oder fear-avoidance-Verhalten. Gleichzeitig weisen sie darauf hin, dass entsprechend der Theorie der Regression zur Mitte (Morton & Torgerson, 2003) eine erneute Befunderhebung nach einem Monat die Qualität der Prognose wesentlich verbessern kann, persistierende Risikofaktoren verschlechtern die Prognose, beispielsweise steigt das Risiko einer schlechten Prognose bei persistierendem starken Schmerz auf das fünfzehnfache. Zur Theorie der Regression zur Mitte passen auch die Beobachtungen von Leboeuf-Yde et al. (2005), wo Patienten mit schweren Rückenschmerzen zu einer Beschwerdelinderung und Patienten mit leichten Beschwerden zu einer Verschlechterung der Beschwerden tendierten. Auch Carey et al. (2000) und Enthoven et al. (2003) fanden, dass die Erfassung der Befunde einen Monat nach Erstkonsultation die Einschätzung der Prognose sicherer macht.

In einer Untersuchung aus Chile fanden Diaz-Ledezma et al. (2009) bei Patienten mit akuten Rückenschmerzen überdurchschnittliche Arbeitsunfähigkeitszeiten bei Patienten, die bereits im Vorjahr wegen Rückenschmerzen arbeitsunfähig gewesen waren (+14 %), körperlich Tätige (+35 %) und Patienten mit Konsultationen bei orthopädischen Chirurgen (+ 43 %).

Das Risiko, dass Rückenschmerzen bei Erstkonsultation beim Allgemeinarzt dann über mehr als 6 Monate anhalten, wurde von Thomas et al. (1999) als deutlich erhöht nachgewiesen, wenn eine Kombination aus folgenden sechs Faktoren vorlag: weibliches Geschlecht, Rezidiv, Schmerzen im Bein, Schmerz an verschiedenen Körperstellen („widespread pain“), Einschränkung der Wirbelsäulenbeweglichkeit und unbefriedigende Tätigkeit.

Du Bois et al. (2009) führten eine logistische Regression über verschiedene Instrumente zur Erfassung eines erhöhten Risiko für eine Arbeitsunfähigkeit über mehr als 3 Monate durch und fand folgende 5 Fragend zielführend:

- Erwarten Sie, innerhalb von 6 Monaten wieder an Ihren Arbeitsplatz zurückkehren zu können?
- In welchem Ausmaß beeinträchtigt der Schmerz Ihre täglichen Aktivitäten?

- Ist es nicht empfehlenswert, physisch aktiv zu sein?
- Fühlen Sie sich allgemein nervös?
- Fühlen Sie sich allgemein beunruhigt/ängstlich?

Faber et al. (2006) fanden, dass die Wiederaufnahme der Berufstätigkeit (return to work) primär von den empfundenen funktionellen Beeinträchtigungen abhängt; rezidivierender LBP und männliches Geschlecht waren dagegen positiv mit der Arbeitsfähigkeit korreliert. In einer Studie von Campello et al. (2006) hatten 25 % der Patienten, die nach durchschnittlich 9 Monaten Arbeitsunfähigkeit nach einer multimodalen Therapie die Arbeit wieder aufnahmen, wieder schmerzbedingte Arbeitsunfähigkeiten. Die Unterbrechung der Arbeitstätigkeit war dabei mit einer Einschränkung der Beweglichkeit der Wirbelsäule in Flexion und Extension, mit Somatisierung und mit erhöhten Werten auf der obsessive-compulsiveness-Skala (entspricht Zwanghaftigkeit, fixe Ideen) der Symptom-Checklist 90 korreliert.

Patienten, die bis zur Bandscheibenoperation trotz Rückenschmerzen arbeiten, waren in einer Untersuchung von Anderson et al. (2006) zu 90 % später wieder arbeitsfähig, während dies bei denjenigen, die bereits vor der Operation wegen ihrer Schmerzen nicht mehr arbeiteten, nur 43 % waren.

Chorti et al. (2009) führten eine Literaturrecherche zum prognostischen Wert von Symptomveränderungen in der konservativen Therapie von Rückenschmerzen durch und fanden nur für die Veränderung von Schmerzlokalisation – und/oder –stärke in Zusammenhang mit wiederholten Wirbelsäulenbewegungen oder im Ergebnis von bestimmten Behandlungen Hinweise für die Behandlungsplanung. Für die Aussagen des Mobilisationstests in Bauchlage besteht in Verbindung zu einem Stabilisationsprogramm eine begrenzte Evidenz (Chorti et al., 2009).

Bei akuten Rückenschmerzen sind die Stärke des Schmerzes und die Anamnese von Rückenschmerzen für eine schlechte Prognose von Bedeutung (Sieben et al., 2005). Bei akutem LBP spielen psychologische Faktoren wie schmerzbezogene Furcht und Katastrophisieren eine wesentliche Rolle in der Chronifizierung (Swinkels-Meewisse et al., 2006b). Dies deckt sich mit den Ergebnissen von Pincus et al. (2002), die eine starke Evidenz für schlechte Resultate bei psychologischer Belastung (distress) und depressiver Stimmung, eine moderate Evidenz bei Somatisierung und nur eine schwache Evidenz für die Verbindung von Copingstil und Verlauf fanden. Macfarlane et al. (2006) stellen zu recht die Frage, wie dieses Wissen aus epidemiologischen Studien in Behandlungsregimen umgesetzt wurde.

Diese Befunde stehen in Übereinstimmung mit der Ergebnissen einer Studie von Myers et al (2008), die ergab, dass funktionelle Verbesserungen bei akuten Rückenschmerzen umso besser sind, je höher die Erwartungen des Patienten an den weiteren Verlauf sind.

Cruz et al. (2020) fanden maladaptive psychologische Faktoren und Arbeitslosigkeit als Risikofaktoren für unbefriedigende Behandlungsergebnisse bei unspezifischen Rückenschmerzen.

## Scores

In einer holländischen Studie (Jellema et al., 2007) zur Vorhersage des Verlaufs von Rückenschmerzen stellten die Autoren fest, dass die Risikoeinschätzung des Hausarztes mit der Frage „wie hoch schätzen Sie das Risiko, dass dieser Patient in seiner Funktionsfähigkeit in 3 Monaten noch eingeschränkt sein wird?“ auf einer numerischen Ratingskala (0 – 100 %) das Risiko einer Chronifizierung mindestens so gut erfasst, wie die Fragebögen des Örebro Schmerzscreenings für muskuloskeletalen Erkrankungen oder die Low Back Pain Perception Scala, die die Vorstellungen des Patienten über verschiedene Aspekte der Rückenschmerzen erfasst.

Neubauer et al. (2006) fanden bei Nutzung des Heidelberger Kurzfragebogens, dass die Items „wie stark war der geringste, am besten erträgliche Schmerz in der vergangenen Woche?“ und „wie viel Schmerz würden Sie ertragen, wenn Sie die Behandlung als erfolgreich einschätzen würden?“ die Entwicklung einer Chronifizierung am besten vorhersagen. Weitere Chronifizierungsrisiken sind eine Schmerzdauer über mehr als 8 Tage, ein niedriges Bildungsniveau und Schmerzen an anderen Körperstellen

Enthoven et al. (2006) fanden, dass bei Frauen das Maß der Behinderung 5 Jahre nach der Erstbehandlung signifikant durch eine längere Dauer der aktuellen Schmerzepisode, durch ein geringeres Niveau sportlicher Betätigung (exercise level) und durch eine höhere Schmerzfrequenz erhöht wird (wobei die Schmerzfrequenz auf einer 5-Punkt-Skala zwischen 0 = nie und 5 = Tag und Nacht ununterbrochen erfasst wurde).

Für die klinische Praxis und Forschung sowie für die Erfassung von Kosten – Nutzen – Relationen verschiedener Schmerztherapieformen ist es nötig, die Schmerzstärke zu erfassen. Nach Zelman et al. (2003) gibt es für einige Krankheiten den Marker „episodenfreier Tag“, ein Konzept, das für den chronischen Schmerz in dieser Form nicht tauglich ist. Als Korrelat wurde ein „Tag mit akzeptabler Schmerzkontrolle (manageable pain control)“ vorgeschlagen, bei dem die Schmerzen als mild im Gegensatz zu moderat und schwer eingeschätzt werden. Bei Patienten mit metastasierendem Karzinom wurden auf einer Skala von 0 bis 10 die Bereiche 0 – 4 als mild, 5 – 6 als moderat und 7 – 10 als schwer definiert (Serlin et al., 1995), für Rückenschmerzen fanden Jensen et al. (2001) die jeweiligen Obergrenzen 5 und 7 auf einer Skala des schlimmsten Schmerzes. Für eine Skala des durchschnittlichen Rückenschmerzes fanden Zelman et al. (2003) die Grenzwerte 5 und 8.

Dunn & Croft (2006a) nutzten als Marker eine Reduktion des Behinderungsscores nach dem Roland-Morris Disability Questionnaire (Roland & Morris, 1983) und fanden, dass sich ab einer Symptombelastung von mehr als drei Jahren die Behandlungsdauer bis zu einer Reduktion der schmerzbedingten Behinderung signifikant verlängert.

Das Risiko, nach einem Jahr noch an Rückenschmerzen zu leiden, kann nach von Korff & Miglioretti (2005) und Dunn et al. (2008) kalkuliert werden:

#### Variablen der Schmerzstärke

Durchschnittliche Schmerzstärke: 0-3 = 0; 4-6 = 1; 7-10 = 2,

stärkster Schmerz 0-4 = 0; 5-7 = 1; 8-10 = 2,

aktueller Schmerz 0-2 = 0; 3-4 = 1; 5-10 = 2,

Beeinträchtigung üblicher Aktivitäten 0-2 = 0; 3-4 = 1; 5-10 = 2

Beeinträchtigung v. Arbeit/Hausarbeit 0-2 = 0; 3-4 = 1; 5-10 = 2,

Beeinträchtigung v. familiären/sozialen Aktivitäten 0-2 = 0; 3-4 = 1; 5-10 = 2,

Tage, an denen in den letzten 6 Monaten die Aktivitäten eingeschränkt werden mussten

0-6 = 1; 7-14 = 2; 15-30 = 3; 31+ = 4,

andere prognostische Variablen

die bisherige Schmerzdauer wurde folgendermaßen gescort: weniger als 3 Monate = 0, 3-6 Monate = 1, 7-12 Monate = 2, 1-2 Jahre = 3, 3 und mehr Jahre = 4.

Die Anzahl anderer Schmerzlokalisationen (Schultergürtel, Arme, Nacken, Kopf) ergab folgenden Score: 0 = 0, 1 = 1, 2 = 2, 3 = 3 und 4 = 4.

Dunn et al. (2008) nahmen zur Ermittlung eines korrekten Depressionsscores die in der Arbeit von von Korff & Miglioretti (2005) für die SCL-90-R genutzten Kategorien und verwendeten diese auf eine Stichprobe der amerikanischen Bevölkerung, um die Proportionen dieser Kategorien in der Allgemeinbevölkerung zu erfassen. Diese Proportionen wurden dann auf den HADS – Score der britischen Allgemeinbevölkerung angewandt, um entsprechende Kategorien für den HADS zu ermitteln. Diese lauten: 0-3 = 0, 4-7 = 1, 8-10 = 2, 11-12 = 3 und 13-21 = 4.

Aus diesen Scores wurde die Wahrscheinlichkeit chronischer Rückenschmerzen ermittelt:

Niedriges Risiko                    0 – 7

Mittleres Risiko                    0 – 15

Möglicherweise chronisch      16 - 21

Wahrscheinlich chronisch      22+

Axen et al. (2014) untersuchten den Einfluss des Zeitpunkts der Erfassung von Behandlungsergebnissen auf die Sicherheit einer Prognose für zukünftige Rückenschmerzen und fanden, dass die aktuell in einem Zeitraum von 26 Wochen mehrfach gemessenen Behandlungsergebnisse stark schwanken. Dies erklärt die unterschiedliche Qualität von Prognosen, um halbwegs sichere Aussagen machen zu können, müssten mehrfache Messungen erfolgen.

Da Silva et al. (2017) bieten ein interessantes Modell zur Einschätzung der Wiederherstellung von Patienten mit akuten Rückenschmerzen, das auf der Dauer der akuten Episode, der Anzahl vorangegangener Episoden, Depressionssymptomen, der Schmerzintensität und der Veränderung der Schmerzintensität in der ersten Woche nach Therapiebeginn beruht.

In einer kleinen belgischen Studie mit 49 Patienten mit chronischen Rückenschmerzen (Denteneer et al., 2016) konnten 24 von einem Rücken-Rehabilitationsprogramm profitieren, 25 nicht. Als einziger signifikanter Risikofaktor für ein Mislingen der Therapie identifizierten die Autoren einen höheren Wert auf der Subskala „physische Funktion“ des SF-36 (PF-SF36).

Ältere Autoren siehe Gesamtliteraturverzeichnis

Axen,I.

Bergström,G., Bodin,L.

Using few and scattered time points for analysis of a variable course of pain can be misleading: an example using weekly text message data

Spine J 14 (2014)1454 – 9

- Burns,S.A. Cleland,J.A., Cook,C.E., Bade,M., Rivett,D.A., Snodgrass,S.  
Variables describing individuals with improved pain and function with a primary complaint of low back pain: a secondary analysis  
*J Manipulat Physiol Ther* 41 (2018)467 - 74
- Cancelliere,C. Donovan,J., Stochkendahl,M.J., Biscardi,M., Ammendolia,C., Myburgh,C., Cassidy,J.D.  
Factors affecting return to work after injury or illness: best evidence synthesis of systematic reviews  
*Chiropr Man Ther* 24 (2016)32: 1 - 64
- Cruz,E.B. Canhao,H., Fernandez,R., Caeiro,C., Branco,J.C., Rodriguez,A.M., Pimentel-Santos,F., Gomes,L.A. und weitere 4 Autoren  
Prognostic indicators for poor outcomes in low back pain patients consulted in primary care  
*PLoS ONE* 15 (2020)3: e0229265  
[doi.org/10.1371/journal.pone.0229265](https://doi.org/10.1371/journal.pone.0229265)
- da Silva,T. Macaskill,P., Mills,C., Maher,C., Lin,C., Hancock,M.J.  
Predicting recovery in Patients with acute low back pain. A clinical prediction model  
*Eur J Pain* 21 (2017)716 – 26 (Internetzugriff 5.8.2017)
- Denteneer,L. van Daele,U., de Hertogh,W., Truijen,S., Stassijns,G.  
Identification of preliminary prognostic indicators for back rehabilitation in patients with nonspecific chronic low back pain  
*Spine* 41 (2016)522 - 9
- Grovle,L. Haugen,A.J., Keller,A., Ntvig,B., Brox,J.I., Grotle,M.  
Prognostic factors for return to work in patients with sciatica  
*Spine J* 13 (2013)1849 - 57
- Mäntyselkä,P. Pain today – disability tomorrow. Commentary  
*Pain* 153 (2012)507 - 8
- Mlekusch,S. Schliessbach,J., Camara,R.J.A., Arendt-Nielsen,L., Jüni,P., Curatolo,M.  
Do central hypersensitivity and altered pain modulation predict the course of low back pain and neck pain?

Clin J Pain 29 (2013)673 - 80

Turner,J.A.

Shortreed,S.M., Saunders,K.W., LeResche,L., Berlin,J.A., von Korff,M.

Optimizing prediction of back pain outcomes

Pain 154 (2013)1391 - 1401

Vavrek,D.A.

Haas,M., Neradilek,M.B., Polissar,N.

Prediction of pain outcomes in a randomized controlled trial of dose-response of spinal manipulation for the care of chronic low back pain

BMC Musculoskel Disord 16 (2015)205