

## Ergebnisqualität in der medizinischen Rehabilitation der Rentenversicherung

- Workshop der Deutschen Rentenversicherung
- 25. November 2008
- München

## Hinweis:

Alle in dem vorliegenden Gesamtbericht verwendeten Berufs-, Personen und Gruppenbezeichnungen gelten sowohl für Frauen als auch für Männer.

Die Betreuung der Qualitätssicherung der medizinischen und beruflichen Rehabilitation erfolgt im Bereich 0430 „Reha-Qualitätssicherung, Epidemiologie und Statistik“ der Deutschen Rentenversicherung Bund, 10704 Berlin.

## Inhaltsverzeichnis

<b>Vorwort</b> <b>Koordinatoren der Reha-Qualitätssicherung</b>	<b>6</b>
<b>Haaf, Hans-Günter, Dr.</b> <b>Reha-Erfolg – Ist die Reha überhaupt wirksam?</b>	<b>10</b>
<b>Farin, Erik, PD Dr.</b> <b>Methodische und konzeptionelle Anforderungen an die Ergebnismessung in der medizinischen Rehabilitation.</b>	<b>48</b>
<b>Dörning, Hans, Dipl.-Soz.-Wiss.</b> <b>Anforderungen an ein adäquates System zur Bewertung des Reha-Erfolgs – das Beispiel „Qualitätsbarometer“.</b>	<b>84</b>
<b>Morfeld, Matthias, Prof. Dr.</b> <b>Die Operationalisierung von Rückkehr an den Arbeitsplatz als unterschätzte Herausforderung an die Abbildung der Ergebnisqualität.</b>	<b>100</b>
<b>Meyer, Thorsten, Dr.</b> <b>Die Bedeutung subjektiver Erfolgsmaße für die Ergebnisqualität.</b>	<b>114</b>
<b>Wirtz, Markus, Prof. Dr.</b> <b>Adaptives Testen: Perspektiven für die Ergebnisqualität.</b>	<b>132</b>
<b>Schulz, Holger, PD Dr.</b> <b>Möglichkeiten und Grenzen der Bestimmung der Ergebnisqualität in der Rehabilitation durch Therapeuteneinschätzungen am Beispiel des SF-8.</b>	<b>156</b>
<b>Widera, Teresia, Dr.</b> <b>Ausblick zur Ergebnisqualität – Indikatoren einer erfolgreichen Rehabilitation sowie Einflussfaktoren auf das Rehabilitationsergebnis.</b>	<b>168</b>
<b>Referenten und Ansprechpartner</b>	<b>196</b>

## Vorwort

Mehr als 15 Jahre nach Einführung des Qualitätssicherungsprogramms der Rentenversicherung ist Qualitätssicherung fester Bestandteil des Rehabilitationssystems in Deutschland. Rehabilitationseinrichtungen und Leistungsträger erhalten eine regelmäßige, aussagekräftige Berichterstattung zur Qualitätssicherung, die zur Steigerung der Qualität der Rehabilitation geführt hat. Seit 2001 bestehen gesetzliche Regelungen zur Qualitätssicherung nach § 20 Abs. 1 Sozialgesetzbuch IX.

Das Programm für die medizinische Rehabilitation wird ständig weiterentwickelt.

Im Bereich der Struktur- und Prozessqualität und bei der Einbeziehung der Rehabilitandinnen und Rehabilitanden in die Qualitätssicherung ist das Programm schon weit entwickelt und anerkannt.

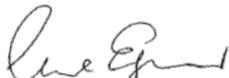
Im Bereich der Ergebnisqualität werden im Rahmen der Rehabilitandenbefragung die Zufriedenheit der Rehabilitanden mit der Rehabilitation sowie die subjektive Einschätzung des Behandlungserfolges berücksichtigt. Hier wird derzeit geprüft, inwieweit Analysen zum Eingliederungserfolg als wichtigen Parameter der Ergebnisqualität in das Programm einbezogen werden könnten. Und zwar sowohl in der Qualitätssicherung der medizinischen Rehabilitation als auch in der Qualitätssicherung der Leistungen zur Teilhabe am Arbeitsleben.

Einrichtungsbezogene Ergebnisse des Qualitätssicherungsprogramms sollen mittelfristig allen Interessierten und insbesondere den Versicherten zugänglich gemacht werden. Ziel ist es, aus den Ergebnissen des Qualitätssicherungsprogramms perspektivisch deutlichere Konsequenzen für die Belegung und Vergütung von Rehabilitationseinrichtungen zu ziehen. Hierbei ist die Ergebnisqualität natürlich von besonderem In-

teresse. Ergebnisse der Qualitätssicherung, und dabei insbesondere solche aus dem Bereich Ergebnisqualität, werden aber auch künftig wichtige Argumente in der sozialpolitischen Diskussion um die Weiterentwicklung des Gesundheitswesens sein.

Als ersten großen Schritt zur Weiterentwicklung des QS-Programms der Rentenversicherung im Bereich der Ergebnisqualität haben wir einen Workshop mit wissenschaftlichen Experten zu diesem Thema durchgeführt, der in diesem Band dokumentiert wird.

Den Experten des Workshops möchten wir für ihre bereitwillige Mitwirkung und für Ihre Vorträge danken, den Teilnehmern für die rege Diskussion, die wir als sehr befruchtend empfunden haben. Schließlich danken wir den MitarbeiterInnen des Bereichs Reha-Qualitätssicherung der Deutschen Rentenversicherung Bund für die Organisation und Dokumentation des Workshops.



Uwe Egner



Dr. Eberhard Grosch



Dr. Here Klosterhuis



Heinz-Gert Verhorst



Michael Zellner

Die Koordinatoren der Reha-Qualitätssicherung  
der Rentenversicherung

Die Zusammenarbeit zur Reha-Qualitätssicherung innerhalb der Rentenversicherung und deren rentenversicherungsweite Umsetzung erfolgt durch fünf Koordinatoren. Hierbei handelt es sich um Mitarbeiter der Regionalträger und der Deutschen Rentenversicherung Bund.

**Uwe Egner**

Fachbereichsleiter Zentrale und Grundsatzaufgaben  
Abteilung Rehabilitation  
Deutsche Rentenversicherung Bund  
10704 Berlin

**Dr. Eberhard Grosch**

Leitender Arzt  
Deutsche Rentenversicherung  
Braunschweig-Hannover  
30875 Laatzen

**Dr. Here Klosterhuis**

Leiter des Bereichs Reha-Qualitätssicherung,  
Epidemiologie und Statistik  
Deutsche Rentenversicherung Bund  
10704 Berlin

**Heinz-Gert Verhorst**

Abteilung für Rehabilitation  
Grundsatzreferat  
Deutsche Rentenversicherung Westfalen  
48125 Münster

**Michael Zellner**

Leiter der Abteilung Kliniken  
Deutsche Rentenversicherung Bayern Süd  
81729 München

# Reha-Erfolg – Ist die Reha überhaupt wirksam?

**Dr. Hans-Günter Haaf**

Deutsche Rentenversicherung Bund

Geschäftsbereich Sozialmedizin und Rehabilitation

Bereich Rehabilitationswissenschaften

## Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>11</b>
<b>2. Sozialmedizinischer Verlauf nach der medizinischen Rehabilitation</b>	<b>12</b>
<b>3. Ergebnisse zur Wirksamkeit aus dem Förderschwerpunkt „Rehabilitationswissenschaften“</b>	<b>14</b>
<b>4. Rehabilitationselemente mit besonderer Wirksamkeit</b>	<b>20</b>
<b>5. Weitere aktuelle Studien zur Wirksamkeit der medizinischen Rehabilitation</b>	<b>28</b>
<b>6. Methodische Aspekte der Wirksamkeitsstudien</b>	<b>33</b>
<b>7. Fazit</b>	<b>34</b>
<b>8. Literatur</b>	<b>38</b>

## 1. Einleitung

Die Ziele der Rehabilitation in der Rentenversicherung sind gesetzlich festgelegt. Die Rehabilitationsleistungen sollen negative Auswirkungen von Krankheit und Behinderung auf die Erwerbsfähigkeit beseitigen. Handlungsbedarf wird bereits dann gesehen, wenn eine erhebliche Gefährdung der Erwerbsfähigkeit vorliegt. Allgemeines Ziel der Rehabilitation in der Rentenversicherung ist die möglichst dauerhafte (Re-)Integration in das Erwerbsleben.

Die Bewertung der Wirksamkeit der Rehabilitation bezieht sich zunächst auf die berufliche Wiedereingliederung als primäres Ziel der Reha-Leistungen der Rentenversicherung. Die (Re-)Integration reicht aber als Erfolgskriterium alleine u. a. deshalb nicht aus, da sie wesentlich von externen Faktoren (u. a. Arbeitsmarkt) mitbestimmt wird (vgl. Schliehe & Haaf, 1996). Zudem ist auch die Rehabilitation der Rentenversicherung dem übergeordneten Ziel verpflichtet, den Rehabilitanden eine Teilhabe am Leben in der Gesellschaft zu ermöglichen. Aus diesem gesetzlichen Auftrag leitet sich nicht zuletzt die ganzheitliche Ausrichtung der Rehabilitation ab. In Evaluationsstudien werden demgemäß auch Indikatoren zur Leistungs- und Erwerbsfähigkeit der Rehabilitanden einbezogen. Die ganzheitliche Ausrichtung der Rehabilitation begründet zudem weiter gefasste Outcomekriterien, z. B. aus dem Bereich der gesundheitsbezogenen Lebensqualität. Dabei ist auch zu berücksichtigen, dass gerade die subjektive Einschätzung des eigenen Gesundheitszustands und der eigenen Leistungsfähigkeit wichtige Prädiktoren für den Reha-Erfolg und die berufliche Wiedereingliederung sind. Ferner sind indikationsspezifische Reha-Ziele, u. a. bezogen auf die psychische Befindlichkeit oder Risikofaktoren, zu berücksichtigen.

## 2. Sozialmedizinischer Verlauf nach der medizinischen Rehabilitation

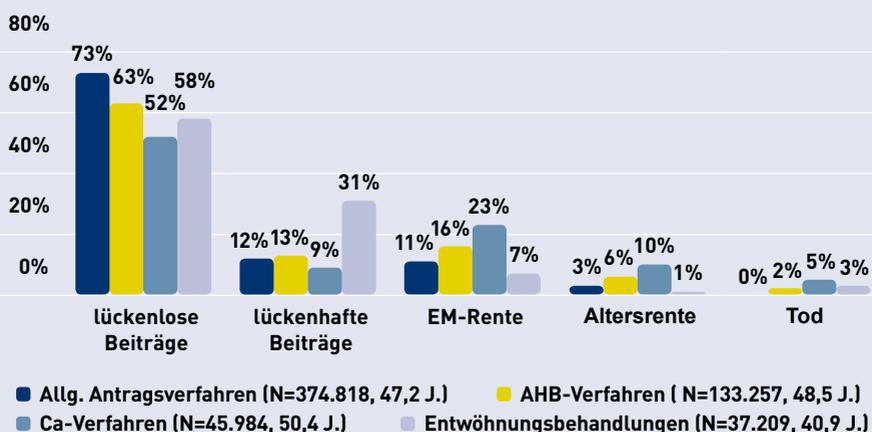
Seit den 1990er Jahren nutzt die Rentenversicherung systematisch ihre prozessproduzierten Daten, um Aussagen über den Erfolg der Leistungen zur medizinischen Rehabilitation zu treffen (Buschmann-Steinhage & Zollmann, 2009). Im Sozialmedizinischen Verlauf wird der Frage nachgegangen, zu welchem Anteil Rehabilitanden nach der Rehabilitationsleistung aus dem Erwerbsleben ausscheiden (in Erwerbsminderungs- oder Altersrente bzw. durch Versterben) und zu welchem Anteil Rehabilitanden im Erwerbsleben verbleiben (und ob dabei lückenlos oder lückenhaft Beiträge an die Rentenversicherung entrichtet werden).

Grundlage der deskriptiv-statistischen Analysen bildet die Reha-Statistik-Datenbasis (RSD), die Verlaufsauswertungen über einen Zeitraum von acht Jahren gestattet. In ihr sind personenbezogene Daten über Leistungen zur medizinischen Rehabilitation sowie Leistungen zur Teilhabe am Arbeitsleben, über Renten und Beitragszeiten sowie soziodemografische Angaben enthalten. Die aktuelle RSD umfasst den Zeitraum von 1999 bis 2006. Alle pflichtversicherten Rehabilitanden mit medizinischer Rehabilitation im Jahr 2004 wurden bis zwei Jahre nach Ende der Leistung beobachtet. Zur Berücksichtigung der Gender-Perspektive wurden zudem getrennte Analysen für Männer und Frauen durchgeführt.

Die Ergebnisse zum sozialmedizinischen Verlauf nach medizinischer Rehabilitation sind seit Jahren relativ stabil: Ca. 70 % aller Rehabilitanden weisen in den zwei Jahren nach Ende der Leistung lückenlose Beitragszahlungen auf und 14 % lückenhafte Beitragszahlungen. Aus dem Erwerbsleben ausgeschieden sind 12 % in Erwerbsminderungsrente, 4 % in Altersrente. Etwa 1 % der Rehabilitanden ist aus dem Erwerbsleben heraus verstorben. Die Unterschiede nach ausgewählten

Verfahrensarten (keine disjunkten Gruppen) sind in der Abbildung 1 wiedergegeben. Besonders hervorzuheben ist der Anteil von Rehabilitanden mit lückenhaften Beitragszahlungen (31 %) bei den Entwöhnungsbehandlungen. Allerdings gehen diese Rehabilitanden auch zu einem geringeren Anteil in Erwerbsminderungs- (7 %) oder Altersrente (1 %) als Rehabilitanden anderer Verfahrensarten. Sie sind mit 41 Jahren auch deutlich jünger. Am ältesten sind mit 50 Jahren die Rehabilitanden mit Krebserkrankungen (Ca-Verfahren), die entsprechend auch zu höheren Anteilen aus dem Erwerbsleben ausscheiden (23 % EM-Rente, 10 % Altersrente).

**Abbildung 1: Sozialmedizinischer Verlauf von pflichtversicherten Rehabilitanden zwei Jahre nach medizinischer Rehabilitation im Jahr 2004 nach ausgewählten Verfahrensarten**



Quelle: Reha-Statistik-Datenbasis (RSD), Erhebungsjahr 2007, Berichtszeitraum 1999-2006

Mit der hohen Rate von Versicherten, die nach einer Rehabilitationsleistung der Rentenversicherung im Erwerbsleben verbleiben, weist die Auswertung des sozialmedizinischen Verlaufs auf die Wirksamkeit der medizinischen Rehabilitation hin.

### 3. Ergebnisse zur Wirksamkeit aus dem Förderschwerpunkt „Rehabilitationswissenschaften“

Der Förderschwerpunkt „Rehabilitationswissenschaften“ stellt das bisher umfassendste Programm in der Reha-Forschung in Deutschland dar. Das Förderprogramm wurde gemeinsam vom Bundesministerium für Bildung und Forschung (BMBF) und der Deutschen Rentenversicherung initiiert. Über einen Zeitraum von insgesamt acht Jahren (1998-2005) wurden acht regionale Forschungsverbände gefördert. Konkrete Zielsetzung war es, Qualität und Umfang anwendungsorientierter Forschung auf dem Gebiet der Rehabilitation zu steigern und die strukturelle Verankerung der Rehabilitationswissenschaften an Universitäten und Hochschulen nachhaltig zu stärken. Mit der systematischen Förderung wurde die wissenschaftlich begründete Weiterentwicklung der Rehabilitation langfristig sichergestellt.

Demgemäß fokussierten die insgesamt etwa 150 Forschungsprojekte vor allem auf die Weiterentwicklung der Reha-Konzepte. Die vorliegende Übersicht bezieht Projekte aus dem Förderschwerpunkt ein, aus denen sich Aussagen zur Wirksamkeit der medizinischen Rehabilitation ableiten lassen. Die einbezogenen Projekte verfolgten zum Teil spezifische Fragestellungen. Die betrachteten Aussagen zur Wirksamkeit der Rehabilitation stellen oft nur einen Ausschnitt aus den vielfältigeren Ergebnissen dar. Für eine vertiefte Information zu den Projekten stehen fast alle Abschlussberichte im Internet zur Verfügung (vgl. Haaf, 2005).

Die betrachteten 38 Evaluationsstudien befassen sich mit verschiedenen Indikationen. Am häufigsten wurde mit 10 Studien die Rehabilitation bei chronischen Rückenschmerzen evaluiert. Die psychosomatische und die kardiologische Rehabilitation wurden jeweils in 9 Studien thematisiert. Die weiteren Projekte lassen sich der Rehabilitation bei Krebserkrankungen (4), bei neu-

rologischen Erkrankungen (3), bei Erkrankungen der Verdauungsorgane (1) sowie bei Kindern und Jugendlichen (2) zuordnen. Im Folgenden werden die wichtigsten Ergebnisse kursiv dargestellt.

### **3.1 Rehabilitation bei chronischen Rückenschmerzen**

Chronische Rückenschmerzen stellen ein schwerwiegendes Gesundheitsproblem dar, das zu bedeutsamen Einschränkungen der gesundheitsbezogenen Lebensqualität führt und die Teilhabe der Patienten an Beruf und Gesellschaft beeinträchtigen kann. Zudem verursachen chronische Rückenschmerzen hohe direkte und indirekte Krankheitskosten.

Die medizinische Rehabilitation bei chronischen Rückenschmerzen erfüllt eine wichtige Versorgungsaufgabe bei Patienten, deren Erkrankung durch Chronifizierung gekennzeichnet und von beruflichen sowie psychosozialen Problemlagen überlagert ist. Die mittlere Erkrankungsdauer lag in den vorliegenden Studien bei mehreren Jahren. Die Rehabilitation steht dabei oftmals am Ende der Behandlungskette und muss somit auch vorherige Versorgungsdefizite auffangen. Nach internationalen Übersichtsarbeiten kann die Wirksamkeit von rehabilitativen Interventionen für intensive multidisziplinäre Maßnahmen als gesichert gelten (Guzman et al. 2001). Die deutsche Rehabilitation ist grundsätzlich multidisziplinär ausgerichtet. Nach den vorliegenden Ergebnissen erzielt sie insbesondere kurzfristige Effekte in verschiedenen gesundheitsbezogenen Outcome-Dimensionen. Diese Effekte lassen jedoch nach der Literaturübersicht von Hüppe & Raspe (2003, 2005) mittelfristig meist wieder nach.

Bei dieser Metaanalyse ist sowohl die methodische Heterogenität der einbezogenen Studien als auch die Unterschiedlichkeit der Ergebnisse zu beachten. In einigen

der betrachteten Projekte erbrachte die Rehabilitation vergleichsweise nachhaltigere Erfolge, so dass diesen

Konzepten Vorbildcharakter zukommen kann.

Der Nachhaltigkeit der Rehabilitation stehen entgegen: hoher Chronifizierungsgrad, berufliche und soziale Problemlagen sowie fehlende Weiterführung der begonnenen Behandlungsansätze in der nachgelagerten Versorgung. Insbesondere das körperliche Training benötigt einen längeren Zeitraum, um seine Wirkung entfalten zu können.

Die Arbeitsgruppe um Greitemann konnte in einer kontrollierten Studie mit einer größeren Stichprobe zeigen, dass sich die Wirksamkeit der stationären Rehabilitation durchaus verstetigen lässt (Dibbelt, Büschel & Greitemann, 2005a+b; Greitemann, Dibbelt & Büschel, 2004). Der evaluierte Reha-Ansatz zeichnet sich u. a. durch eine assessmentgesteuerte, individuelle Zuweisung zu spezifischen Behandlungspfaden aus. Neben den psychologischen und bewegungstherapeutischen Elementen ist die klare berufliche Orientierung des Reha-Konzepts hervorzuheben. Besonders bemerkenswert ist, dass der innovative Reha-Ansatz mittlerweile in die Routineversorgung übernommen werden konnte. Auch eine Übertragung auf weitere Reha-Einrichtungen wurde erreicht (s.u.).

Eine ebenfalls nachhaltigere Wirkung der Rehabilitation konnten die Arbeitsgruppen um Fikentscher sowie Leidl erreichen. Den evaluierten Reha-Ansätzen ist die Intensivierung der psychologischen Betreuung gemeinsam (Fikentscher et al., 2005; Leidl, Schweikert & Seitz, 2002). Bei der Studie von Leidl und Mitarbeitern sind die bedeutsamen Einsparungen von 1.670 € p. P. für die Interventionsgruppe hervorzuheben. Das Konzept von Fikentscher und Mitarbeitern war auf eine spezielle Zielgruppe (sensomotorisches Dekonditionierungssyndrom und suppressives Schmerzverhalten) zugeschnitten.

### **3.2 Rehabilitation bei psychosomatischen Erkrankungen**

Die psychosomatische Rehabilitation ist der am inten-

sivsten evaluierte Indikationsbereich. Auf dieser Grundlage konnte eine umfassende und aussagekräftige Meta-Analyse mit einer großen Zahl von einbezogenen Studien durchgeführt werden, die einen Überblick zu den Ergebnissen der psychosomatischen Rehabilitation bei mehreren Tausend Rehabilitanden gibt (Huber et al., 2004; Löschmann et al., 2005). Die Resultate der Studien sind weitgehend homogen und belegen den Erfolg der Rehabilitation eindrucksvoll. Unmittelbar nach der Maßnahme lassen sich große Effekte für das körperliche und psychische Befinden bei den meist schon langjährig erkrankten Rehabilitanden nachweisen. Nach einem Jahr dokumentiert sich die Wirkung noch in mittleren Effektstärken. Gesundheitsökonomisch von Bedeutung ist der deutliche Rückgang der AU-Zeiten im Jahr nach der Rehabilitation. Den Erfolg der psychosomatischen Rehabilitation belegen zudem ausnahmslos alle weiteren im Förderschwerpunkt „Rehabilitationswissenschaften“ zur psychosomatischen Rehabilitation durchgeführten Studien.

In mehreren Projekten des Förderschwerpunkts wurden Ansätze für eine weitere Optimierung evaluiert. Die von Beutel et al. (2002) untersuchte Belastungserprobung erwies sich als generell wirksam und kann zu einer stärkeren Berufsorientierung in der psychosomatischen Rehabilitation beitragen. Es zeigten sich aber auch Akzeptanzprobleme. Zudem kamen weniger als 20% der Patienten für die Belastungserprobung in Frage. Über eine hohe Akzeptanz von prä- und poststationären ambulanten Reha-Modulen berichten Bischoff et al. (2003). Die poststationäre Maßnahme konnte mittlerweile in der Klinik in ein Nachsorgekonzept eingebunden werden. Ergebnisse von Mussgay et al. (2004) sowie Mussgay & Rüddel (2005) weisen auf die Bedeutung von körperlichem Training auch in der Psychosomatik hin. Schmitz et al. (2002) zeigten in einer randomisierten Studie mit einer großen Stichprobe auf, dass ein hoher Anteil suchtkranker Patienten von Persönlichkeitsstörungen betroffen ist. Mit dem evalu-

ierten Gruppenprogramm konnte die Wirksamkeit der Rehabilitation für diese besonders fordernde Patientengruppe erhöht werden.

### **3.3 Rehabilitation bei kardiologischen Patienten**

Für koronare Herzerkrankungen ist die Wirksamkeit von multimodalen, rehabilitativen Interventionen sehr gut abgesichert. Die Reduktion von Risikofaktoren u. a. durch körperliches Training, Nikotinabstinenz, Gewichtsreduktion und Cholesterinsenkung kann nicht nur den Gesundheitszustand verbessern und die Lebensqualität erhöhen, sondern auch die Mortalität senken. Eine wirksame Verringerung der Risikofaktoren setzt allerdings eine längerfristige Compliance der Patienten und eine Lebensstiländerung voraus. Eine dreiwöchige Rehabilitationsmaßnahme kann die Voraussetzungen dafür schaffen und diesen Prozess einleiten, der dann verstetigt werden muss.

Nach Brenner & Rothenbacher (2002) erreicht die kardiologische Rehabilitation auch im Vergleich mit internationalen Resultaten beachtliche kurzfristige Effekte. Die mittel- bis langfristigen Ergebnisse sind jedoch uneinheitlich. In vier Studien konnte nach einem Jahr eine Reduktion von Risikofaktoren festgestellt werden. Weniger stabil scheinen die Reha-Erfolge bei der psychischen Befindlichkeit zu sein. Die Rehabilitation kann den Erhalt der beruflichen Eingliederung als ein primäres Reha-Ziel bei der überwiegenden Mehrheit der Rehabilitanden erreichen. Nur ein sehr geringer Anteil der Versicherten wird im Jahr nach der Rehabilitation wegen verminderter Erwerbsfähigkeit berentet.

### **3.4 Rehabilitation bei Krebserkrankungen**

Delbrück & Witte (2004) haben im Rahmen des Förderschwerpunkts eine internationale Vergleichsstudie zur Situation der onkologischen Rehabilitation in Ländern der EU durchgeführt. Die Autoren weisen auf die besonders günstige Lage in Deutschland hin und betonen, dass die ganzheitliche Rehabilitation den Forderungen

von Selbsthilfegruppen und Ärzten entspricht, wie sie in den letzten Jahren vermehrt erhoben werden und teilweise auch bereits als gesundheitspolitische Ziele von Regierungen formuliert sind. Weis, Moser & Bartsch (2002) betonen zudem, dass dem intensiven, beratenden und therapeutischen Angebot der Reha-Kliniken in Deutschland größtenteils schlecht ausgebaute Strukturen bei der ambulanten Krebsnachsorge gegenüber stehen.

Nach den Studien aus dem Förderschwerpunkt erfüllt die Rehabilitation bei Krebserkrankungen ihren spezifischen Versorgungsauftrag. Die körperliche und psychische Befindlichkeit der Rehabilitanden wird deutlich gebessert und ihre Krankheitsverarbeitung günstig beeinflusst.

In zwei onkologischen Projekten wurden auch Patienten einbezogen, die keine Rehabilitation in Anspruch genommen haben (Weis, Moser & Bartsch, 2002; Welk et al., 2005). Charakteristisch waren die deutlich besseren Ausgangswerte dieser Patienten gegenüber den schwerer eingeschränkten Rehabilitanden. In beiden Studien konnte die Rehabilitation die schlechteren Ausgangswerte ausgleichen und zudem spezifische Effekte im psychosozialen Bereich und bei der Krankheitsverarbeitung erzielen.

Die aufgezeigte höhere Bedürftigkeit der Rehabilitanden lässt auf eine gute individualspezifische Indikationsstellung seitens der zuweisenden Vorbehandler schließen. Patienten mit erhöhtem Bedarf werden den Reha-Angeboten angemessen zugewiesen. Weis, Moser & Bartsch (2002) heben hervor, dass Patienten, die sich nicht für die Rehabilitation entscheiden, auch keine ambulanten, psychosozialen Angebote in Anspruch nehmen. Positive Ergebnisse erbrachten zwei Studien zur Teilnahme von Angehörigen an der Rehabilitation (Schönberger & v. Kardorff, 2002; Welk et al., 2005). Auch die begleitenden Partner profitieren vom Aufenthalt in der Reha-

Klinik. Sie verbesserten sich signifikant im Hinblick auf die psychische Lebensqualität, während dies für die zu Hause gebliebenen Partner nicht gilt. Zudem ist die Zufriedenheit der Rehabilitanden sowie der Angehörigen hoch.

## 4. Rehabilitationselemente mit besonderer Wirksamkeit

Die Ergebnisse des Förderschwerpunkts „Rehabilitationswissenschaften“ weisen darauf hin, dass Patientenschulungen und beruflich orientierte Ansätze die Wirksamkeit der Rehabilitation erhöhen können. Diese Behandlungselemente sind auch Schwerpunkt der Umsetzung der Erkenntnisse aus dem Förderschwerpunkt in die Praxis (vgl. Zwingmann et al., 2004). Deshalb werden die vorliegenden wissenschaftlichen Ergebnisse zu diesen rehabilitativen Elementen vertiefend dargestellt. Als weiteres aussichtsreiches Konzept hat sich in einer Untersuchung des Förderschwerpunkts die Telefonnachsorge erwiesen. Diese Studie hat mittlerweile einige weitere Forschungsarbeiten angeregt, die ebenfalls referiert werden.

### 4.1 Patientenschulungen

Eine gezielte Schulung der chronisch kranken Rehabilitanden hat sich u. a. in mehreren Studien des Förderschwerpunkts „Rehabilitationswissenschaften“ als besonders wirksam erwiesen. Eine umfassende Übersicht zu den Ergebnissen geben Faller et al. (2005). Unter Einbeziehung der vorliegenden nationalen Studien und internationaler Übersichtsarbeiten bzw. Metaanalysen kommen die Autoren zu dem Schluss, dass Patientenschulung als Teil der medizinischen Rehabilitation in unterschiedlichem Ausmaß medizinisch, psychosozial und sozio-ökonomisch effektiv ist. Positive Resultate liegen für die Indikationsgebiete Rheumatologie, Orthopädie, Kardiologie, Asthma,

Diabetes mellitus und Onkologie vor. Die Stärkung der Patientenschulung in der Rehabilitation ist deshalb ein wesentliches Ziel in der Umsetzung der Ergebnisse. Im Folgenden sollen insbesondere die ökonomischen Effekte exemplarisch näher beleuchtet werden.

Bönisch & Ehlebracht-König (2001) untersuchten ein Schulungsprogramm für Patienten mit Spondylitis ankylosans und verwandten Spondylarthropathien während einer stationären Rehabilitation. In der quasi-experimentellen Studie wurden 167 Rehabilitanden der Interventionsgruppe und 156 Teilnehmer der Kontrollgruppe zugeordnet. Die Interventionsgruppe erhielt ein interdisziplinäres Schulungsprogramm der Deutschen Gesellschaft für Rheumatologie, die Kontrollgruppe zusätzliche unspezifische Behandlungen. Für beide Gruppen zeigten sich nach der Rehabilitation signifikante Verbesserungen u. a. für Schmerzen, krankheitsbezogenes Wissen, Selbstwirksamkeit, Funktionskapazität und Lebensqualität. Im Wissenszuwachs sowie bei den Schmerzen zeigte sich die Interventionsgruppe signifikant überlegen. Ein Jahr nach der Rehabilitation war die Interventionsgruppe in verschiedenen Merkmalen weiterhin stärker gebessert und stellte seltener einen Rentenanspruch. Eine gesundheitsökonomische Analyse ergab deutliche Einsparungen bei den Folgekosten, die überwiegend durch geringe Arbeitsunfähigkeitstage realisiert wurden. Der return on investment lag zwischen 17:1 und 38:1 (Krauth et al., 2003).

Die langfristigen Effekte eines Schulungsprogramms für Patienten mit chronischer Polyarthrit (cP) untersuchten Mau & Bräuer (2003). Grundlage der randomisierten Studie mit 186 Patienten war ebenfalls ein standardisiertes Schulungsprogramm der Deutschen Gesellschaft für Rheumatologie. Nach diesen Ergebnissen verlängerte sich die Zeitdauer zwischen Ende der Rehabilitationsmaßnahme und Eintritt einer Erwerbsminderungsrente von 34 Monaten in der Kontrollgruppe auf durchschnittlich 49 Monate in der Interventions-

gruppe. Aus gesellschaftlicher Perspektive lässt sich daraus eine Einsparung indirekter Kosten in Höhe von ca. 24.000 € über den gesamten Beobachtungszeitraum errechnen. Die Schulung selbst erfordert einen Mehraufwand von lediglich ca. 137 € pro Teilnehmer.

Positive ökonomische Effekte konnten Wittmann, Spohn und Petro (2001) in einer randomisierten Studie für eine standardisierte Patientenschulung bei chronisch obstruktiver Bronchitis an 212 Patienten zeigen. Unter Berücksichtigung der direkten Kosten konnten 455 €, bei zusätzlicher Berücksichtigung der indirekten Kosten 1.869 € pro geschultem Patienten im Jahr eingespart werden. Als ebenfalls wirksam erwies sich ein neu entwickeltes Schulungsprogramm bei nierenkranken Rehabilitanden (Fritschka, 2002, n = 281). Hinsichtlich der indirekten Folgekosten zeigte sich bei einer Substichprobe von Erwerbstätigen (n = 176) ein signifikant geringerer Anteil an Arbeitsunfähigkeitstagen und Berentungen nach der Schulung.

In einer jüngeren Studie evaluierten Meng et al. (2009) das Curriculum Rückenschule des Gesundheitstrainingsprogramms der Deutschen Rentenversicherung Bund. Die Studie bezog 360 orthopädische Rehabilitanden in einem randomisierten Kontrollgruppendesign ein. Zum Ende der Rehabilitation zeigten sich signifikante größere Gruppenunterschiede vor allem in schulungsbezogenen Variablen. Die Interventionsgruppe wies ein höheres Krankheits- und Behandlungswissen auf. Signifikante, kleine bis mittlere Differenzen ergaben sich für die Schmerzbelastung und Aspekte der Schmerzverarbeitung. Keine bedeutsamen Gruppenunterschiede zeigten sich im Funktions- und Gesundheitsstatus. Nach sechs Monaten blieben die erreichten Effekte in beiden Gruppen für den Gesundheitszustand und die Funktionsfähigkeit insgesamt erhalten. Die Interventionsgruppe zeigte sich weiterhin vor allem in schulungsbezogenen Aspekten und in der Schmerzbewältigung überlegen.

## 4.2 Berufliche Orientierung

Das Rehabilitationsziel der Deutschen Rentenversicherung ist beruflich orientiert: Die Erhaltung oder Wiederherstellung der Erwerbsfähigkeit, die Integration oder Reintegration in das Erwerbsleben. Erfahrungsgemäß reicht eine medizinische Orientierung allein dazu oft nicht aus. Forschungsprojekte belegen die hohe Bedeutung subjektiver Einstellungen zur Erwerbsfähigkeit für die berufliche Wiedereingliederung und begründen damit eine gezieltere Arbeit an diesen Einflussgrößen. Die Deutsche Rentenversicherung unterstützt deshalb die aktuelle Entwicklung und Evaluierung von berufsorientierten Angeboten.

Nach Müller-Fahrnow et al. (2005) versteht sich die medizinisch beruflich orientierte Rehabilitation (MBOR) „...als Ansatz mit schärferer Fokussierung auf die Folgen gesundheitlicher Beeinträchtigungen in der Berufs- und Arbeitsrealität. Sie leitet Reha-Diagnostik und Reha-Behandlung in der medizinischen Rehabilitation von gesundheitlich bedingten spezifischen beruflichen Problemlagen, von der gestörten Teilhabe am Erwerbsleben ab.“ (S. 289). Mehrere Studien zeigten mittlerweile auf, dass durch berufsbezogene Behandlungselemente die Wirksamkeit der Rehabilitation erhöht werden kann. Indikative Schwerpunkte sind dabei Kardiologie, Orthopädie und Psychosomatik.

Im Rahmen des Förderschwerpunkts belegen die Ergebnisse von Keck et al. (2000) den Erfolg einer eintägigen berufsorientierten Maßnahme für berufliche Problempatienten in der kardiologischen Rehabilitation. In die Basiserhebung der prospektiven Studie wurden 915 AHB-Patienten aufgenommen, die jünger als fünfzig Jahre waren. 60 Rehabilitanden, bei denen Probleme bei der beruflichen Wiedereingliederung zu erwarten waren, wurde die Möglichkeit zur Teilnahme an einer eintägigen berufsorientierten Maßnahme in einem Berufsförderungswerk eingeräumt. Nach 6 Monaten waren diese Rehabilitanden häufiger beruflich integriert als

Patienten mit vergleichbaren beruflichen Risiken. Die Autoren bewerten die Maßnahme damit als erfolgreich.

In einer randomisierten Studie untersuchten Kittel (2007) sowie Kittel, Liebich & Karoff (2008) ein berufsbezogenes Programm, das neben psychologischen Einzelinterventionen aus realitätsnahen arbeitsplatzbezogenen Übungen besteht, die der Evaluation der funktionalen Leistungsfähigkeit (EFL) dienen und zur Entwicklung einer realistischen Selbsteinschätzung beitragen sollen. Die Rehabilitanden in der Interventionsgruppe waren ein Jahr nach der Rehabilitation u. a. signifikant häufiger beruflich integriert. Demgegenüber konnten Zwerenz et al. (2008) für ein berufsbezogenes Schulungsprogramm zwar Verbesserungen bei berufsbezogenen Einstellungen und in der Behandlungszufriedenheit erreichen, aber keine höheren Eingliederungsraten nach 6 Monaten erzielen.

In einem assoziierten Projekt des nordrhein-westfälischen Forschungsverbunds evaluierten Greitemann et al. (2004) verschiedene Maßnahmen zur Verbesserung der beruflichen Integration bei Rückenschmerzpatienten, u. a. Informationsveranstaltungen zu Rentenrecht und zur beruflichen Rehabilitation, intensive Eingangsdagnostik und berufliche Beratung sowie Arbeitsplatzanalyse und arbeitsplatzbezogenes Bewegungstraining.

Auf der Basis einer multiprofessionellen Diagnostik wurde zudem die Bildung homogener Patientengruppen in drei Behandlungspfaden ermöglicht. Die Effektivität dieser Maßnahmen wurde mit einem Standard-Reha-Programm verglichen. In zeitlichen Blöcken nahmen insgesamt 482 Patienten an dem Projekt teil. Bei Reha-Ende wies die Interventionsgruppe größere Verbesserungen auf, die auch nach 3 und 10 Monaten erhalten blieben. Die Effektstärke für den Reha-Status lag zum Reha-Ende bei 1,2, drei Monate später bei 0,7 und hatte auch nach 10 Monaten eine mittlere Ausprägung. Die Effektstärken in der Vergleichsgruppe waren

geringer (zwischen 0,2 und 0,5). Nach 10 Monaten nahmen die AU-Tage in der Kontrollgruppe um 20 % in der IG hingegen um 38 % ab. Zudem fühlte sich die Interventionsgruppe weniger stark beruflich belastet und war beruflich zufriedener. Die Rehabilitanden sahen ferner ihren Beratungsbedarf zu sozialmedizinischen und sozialrechtlichen Problemen vergleichsweise stärker erfüllt.

Bethge et al. (2009) evaluierten ein berufsbezogenes kognitiv-behaviorales Gruppenkonzept als Teil einer integrierten medizinisch-berufsorientierten orthopädischen Rehabilitation. 236 Rehabilitanden wurden zufällig der beruflich orientierten oder der üblichen Reha zugeordnet. Unmittelbar nach der Maßnahme zeigten sich in der Interventionsgruppe (IG) eine höhere Selbstwirksamkeitserwartung und eine bessere Vorbereitung auf die Zeit nach der Rehabilitation. Sechs Monaten später waren die Teilnehmer des berufsbezogenen Konzepts häufiger beruflich integriert und berichteten über geringere psychosoziale Beschwerden. Die Autoren weisen darauf hin, dass die Effekte in der IG deutlich höher liegen als in der Meta-Analyse von Hüppe & Raspe (2003, 2005). Beispielsweise wird für Schmerzen eine Effektstärke von 1,08 und für die Funktionsfähigkeit von 0,52 erreicht. Die Effizienz einer weiteren berufsbezogenen Maßnahme für die Orthopädie konnten Blume et al. (2007) in einer ebenfalls randomisierten Studie zeigen. Durch die höhere Quote beruflich reintegrierter Rehabilitanden erzielte die Rentenversicherung u. a. höhere Beitrags-einnahmen in den 18 Monaten nach der Rehabilitation. In einem Kooperationsprojekt wurden berufsbezogene Gruppenkonzepte vergleichend in der verhaltenstherapeutischen und der tiefenpsychologischen psychosomatischen Rehabilitation evaluiert (Koch & Hillert, 2006; Zwerenz et al., 2007). Das kognitiv-verhaltenstherapeutische Gruppentherapiekonzept „Stressbewältigung am Arbeitsplatz“ umfasst 8 Sitzungen à 90 Minuten. 289 beruflich belastete Rehabilitanden erhielten entweder eine Rehabilitation mit dem Gruppenkonzept

oder eine Standardrehabilitation. Zwölf Monate nach Entlassung waren die Rehabilitanden der berufsbezogenen Intervention häufiger wieder erwerbstätig (80 %). Zudem nahmen sie häufiger berufliche Nachsorgemaßnahmen in Anspruch, stellten seltener einen Rentenantrag und waren mit der Behandlung zufriedener.

An dem parallel evaluierten tiefenpsychologisch fundierten Gruppenangebot nahmen 144 beruflich belastete Rehabilitanden teil. 135 Patienten, die ebenfalls berufliche Belastungen aufwiesen, bildeten die (nicht randomisierte) Kontrollgruppe. Der Katamnesezeitraum umfasste drei Jahre. Die Intervention erreichte vor allem eine höhere Behandlungszufriedenheit und eine häufigere Inanspruchnahme von beruflichen Reha-Leistungen. Unterschiede bezogen auf die Erwerbstätigkeit, u. a. Berentung und Arbeitslosigkeit, konnten nicht nachgewiesen werden.

#### **4.3 Telefonische Nachsorge**

Ein vergleichsweise einfacher und kostengünstiger Ansatz zur Verstetigung der Rehabilitationseffekte stellt die telefonische Nachsorge dar. Positive Ergebnisse zu Telefonnachsorgeprogrammen liegen für die Kardiologie aus den USA vor. China et al. (2005) haben den Ansatz erstmalig in Deutschland evaluiert. Dabei wurden die Rehabilitanden von speziell fortgebildeten Studienschwestern über ein Jahr regelmäßig telefonisch nachbetreut. Von den angesprochenen Rehabilitanden erklärte sich etwa die Hälfte zur Teilnahme bereit. Diese 344 Patienten wurden zufällig auf die telefonische Nachsorgegruppe und eine Kontrollgruppe, die lediglich schriftliche Informationen (Aufmerksamkeitsplacebo) erhielt, aufgeteilt.

Die Studienschwestern berichten über eine hohe Akzeptanz auf Seiten der Patienten. Krankheitsbezogene Ängste waren ein häufiges Gesprächsthema. Der Zeitaufwand für die Gespräche wird als hoch eingeschätzt. Insgesamt 297 Patienten beteiligten sich über 12 Monate

an der Studie. Die Ergebnisse weisen bei männlichen Patienten der IG auf bessere Werte bei einem Risikofaktorenscore hin. Für die deutlich kleinere Gruppe der erkrankten Frauen ließ sich dies jedoch nicht zeigen.

In einer weiteren randomisierten Studie konnten Kohlmeier et al. (2008) die Effektivität der telefonischen Nachsorge bestätigen. Nach einer kardiologischen AHB wurden die Rehabilitanden über 18 Monate telefonisch betreut und nach sechs Monaten einen Tag lang nachgeschult. In der Interventionsgruppe wurden signifikant weniger Rehabilitanden wegen verminderter Erwerbsfähigkeit berentet. Zudem zeigte sich eine Überlegenheit bei der Lebensqualität und somatischen Maßen.

Keck (2009) verglich ebenfalls für die kardiologische Anschlussrehabilitation ein strukturiertes Telefonnachsorgekonzept bzw. die Kombination mit einem individuellen Nachsorgetermin in der Klinik mit einer Standardtherapie. In der umfangreichen randomisierten Studie wurden insgesamt 198 Rehabilitanden mit telefonischer Nachsorge und 400 AHB-Patienten mit Standardreha behandelt. In der Telefon-Gruppe wurden die Rehabilitanden im Jahr nach der AHB drei Mal angerufen, bei einer Gesprächsdauer von etwa 20 Minuten. In dieser Gruppe ergaben sich eine tendenziell bessere berufliche Reintegration nach 12 Monaten und ein geringeres Berentungsrisiko. Zudem beschrieben die Rehabilitanden in der IG ihren Gesundheitszustand positiver und sie nahmen signifikant häufiger an einer Herzgruppe teil.

Eine telefonische Nachsorge bei Herzinsuffizienz untersuchten Mische et al. (2008) an einer relativ kleinen Stichprobe von 39 Patienten. Eine entsprechend geschulte Schwester rief die Patienten im Jahr nach der Rehabilitation insgesamt 16 Mal an. Die Rehabilitanden zeigten vor allem trainingsabhängige Verbesserungen der körperlichen Leistungsdaten und eine geringe Wiedereinweisung ins Krankenhaus. Zudem konnte die

Compliance mit der Medikation unterstützt werden. Keinen zusätzlichen Nutzen konnten Mangels et al. (2008) mit einer Ergänzung ihres verhaltensmedizinischen Reha-Konzepts durch eine telefonische Nachsorge in der Orthopädie erreichen. Die Autoren betonen zwar die Praktikabilität und Akzeptanz der Telefonate, es finden sich aber bei dem randomisierten Vergleich keine Gruppenunterschiede.

## 5. Weitere aktuelle Studien zur Wirksamkeit der medizinischen Rehabilitation

Nach Abschluss des Förderschwerpunkts „Rehabilitationswissenschaften“ wurden mittlerweile zahlreiche weitere Evaluationsstudien durchgeführt, die hier in einem kursorischen Überblick dargestellt werden. Grundlage für die Zusammenstellung sind die Tagungsbände des Rehabilitationswissenschaftlichen Kolloquiums, das als größter deutscher Reha-Kongress eine gute Übersicht zur Reha-Forschung zulässt. Dabei wurden vor allem Studien mit kontrolliertem Design zu den häufigeren Indikationen einbezogen.

### 5.1 Rehabilitation bei chronischen Rückenschmerzen

Ergebnisse aus dem Förderschwerpunkt „Rehabilitationswissenschaften“ und weitere Forschungsergebnisse weisen darauf hin, dass die Rehabilitation bei chronischen Rückenschmerzen durch die Intensivierung von psychologischen Behandlungselementen effektiver gestaltet werden kann. In einer randomisierten Studie untersuchten Mangels et al. (2006) einen verhaltensmedizinischen Rehabilitationsansatz der u. a. Gruppen zur Schmerzbewältigung, Einzelpsychotherapie sowie ggf. Stressbewältigungs- und soziales Kompetenztraining umfasste. 300 RehabilitandInnen, die überwiegend an Rückenschmerzen litten, wurden zufällig auf eine übliche Reha und den verhaltensmedizinischen Ansatz aufgeteilt. In beiden Gruppen konnten nach

12 Monaten noch deutliche Erfolge bei der gesundheitsbezogenen Lebensqualität festgestellt werden (Schwarz et al., 2007). Das verhaltensmedizinische Konzept hatte Vorteile bezogen auf die psychische Belastung und die Schmerzbewältigung.

Inwieweit durch die medizinische Rehabilitation die Depressivität bei Patienten mit chronischen Rückenschmerzen positiv beeinflusst werden kann, untersuchten Hampel et al. (2009). In das Vergleichsgruppendesign wurden Rehabilitanden mit niedriger und mittlerer bzw. hoher Depressivität einbezogen. Ein Teil der depressiveren Patienten erhielt zusätzlich fünf einstündige Gruppensitzungen mit kognitiv-behavioralem Training zur Depressionsbewältigung. Bei Ende der Rehabilitation zeigten sich alle Rehabilitanden gebessert. Aber lediglich für die Interventionsgruppe zeigten sich auch 6 Monate nach Reha noch Effekte für Depressivität, Ängstlichkeit und gesundheitsbezogene Lebensqualität. Nach einem Jahr war die Interventionsgruppe weiterhin weniger depressiv als zu Beginn der Rehabilitation.

Bethge & Müller-Fahrnow (2008) haben eine Meta-Analyse zur Wirksamkeit einer intensivierten stationären Rehabilitation bei muskuloskelettalen Erkrankungen vorgelegt. Dabei berücksichtigten sie sowohl beruflich orientierte als auch verhaltensmedizinische Ansätze. Die Autoren konnten fünf Studien aus den Jahren 1994 bis 2006 mit überwiegend hoher methodischer Qualität (vier randomisierte Studien) einbeziehen, darunter die hier schon erwähnten Studien von Leidl, Schweikert & Seitz (2002) sowie Schwarz et al. (2007). Die Forschungsergebnisse belegen, dass eine intensiverte Rehabilitation im Vergleich zur herkömmlichen Rehabilitation mittel- und langfristig den subjektiven Gesundheitszustand verbessert. Nach Einschätzung der Autoren findet sich moderate Evidenz dafür, dass eine intensiverte Rehabilitation, die eine psychologische Gruppenintervention umfasst, die

Schmerzenintensität vergleichsweise stärker reduziert. Aus den Studien lässt sich jedoch nicht ableiten, dass die Programme die Arbeitsunfähigkeitszeiten und die Funktionsfähigkeit stärker positiv beeinflussen als die übliche Rehabilitation.

Bei den Ergebnissen aus dem Förderschwerpunkt „Rehabilitationswissenschaften“ waren die nachhaltigen Effekte des umfassenden Konzepts der Arbeitsgruppe um Greitemann besonders hervorzuheben. Dieser Ansatz wurde nun in drei orthopädischen Rehakliniken erfolgreich implementiert (Fröhlich, Greitemann & Niemeyer, 2009). Die Studie wurde in einem Vergleichsgruppendesign durchgeführt (Blockrandomisierung). Insgesamt 437 Patienten wurden mit dem untersuchten Reha-Programm behandelt, 375 Rehabilitanden gehörten zur Kontrollgruppe. Die Ergebnisse weisen 10 Monate nach Reha-Ende einen verbesserten Reha-Status bei fast allen Rehabilitanden aus. Die IG zeigt dabei eine stärkere Verbesserung als die Kontrollgruppe auf. Die Überlegenheit der Studiengruppe betrifft vor allem die Bereiche Gesundheitsverhalten, Schmerzwahrnehmung und AU-Zeiten. Die Effektstärken liegen in der Interventionsgruppe verglichen mit den Ausgangswerten beispielsweise für Funktionsfähigkeit und Krankheitskompetenz bei 1,0 sowie für AU-Tage bei 0,5.

In einer Forschungsarbeit von Huber, Wiskemann & Heilmeyer (2007) wurden die Effekte eines intensivierten Krafttrainings bei chronischen Rückenschmerzen untersucht. In der randomisierten Studie mit insgesamt 233 Patienten trainierte die Interventionsgruppe im Anschluss an die stationäre Rehabilitation über einen Zeitraum von sechs bis neun Wochen in einem ambulanten Reha-Zentrum weiter. Die Kontrollgruppe (n = 118) konnte nach der Reha an einem rückenbezogenen Kursangebot der gesetzlichen Krankenkasse in Wohnortnähe teilnehmen. Während der stationären Phase zeigten sich in beiden Gruppen ein Kraftzuwachs und eine Schmerzreduktion. Nach sechs Monaten war in beiden Gruppen

eine weitere Verbesserung zu verzeichnen, die sich auch in der Funktionsfähigkeit (FFbH) ausdrückte. Die Interventionsgruppe erreichte insgesamt bessere Werte als die Kontrollgruppe. Die Autoren geben die Effektstärken bezogen auf die Ausgangswerte in der Interventionsgruppe mit 0,7 für den Kraftzuwachs, 0,6 für die Funktionsfähigkeit und 0,5 für die Schmerzen an.

### **5.2 Rehabilitation bei kardiologischen Patienten**

Einen besonderen Beleg der Wirksamkeit der kardiologischen Rehabilitation erbrachte eine Studie, in der Rehabilitanden nach einer akuten Herz-Kreislauf-Erkrankung mit Patienten verglichen wurden, die nach ihrem Krankenhausaufenthalt keine Reha in Anspruch nahmen (Waldmann et al., 2008). Obwohl die Rehabilitanden älter und kränker waren als die Vergleichsgruppe, traten nach der Rehabilitation signifikant seltener klinisch relevante Ereignisse (u. a. Versterben, akuter Myokardinfarkt, Rehospitalisation) auf. Für die Forschergruppe weisen diese Ergebnisse „...auf eine möglicherweise bislang unterschätzte Wirksamkeit der kardialen Rehabilitation hin“.

Zur Initiierung regelmäßiger körperlich-sportlicher Aktivitäten im Alltag bei kardiologischen Rehabilitanden setzten Höner et al. (2008) ein Interventionsmodell ein, das auf einem handlungspsychologischen Modell beruht. Es geht um motivationale und volitionale Determinanten, die regelmäßigen Sport begünstigen. In die kontrollierte Evaluationsstudie wurden 354 kardiologische Rehabilitanden einbezogen. In beiden Gruppen konnten die Motivations- und Volitionsstärke sowie handlungspsychologische Messgrößen bezogen auf sportliche Aktivitäten gesteigert werden. Für die körperliche Funktionsfähigkeit erzielten beide Gruppen positive Effekte. Bei den genannten Verbesserungen gab es insgesamt Vorteile für die Interventionsgruppe. Auf der Verhaltensebene konnte jedoch kein Effekt der Intervention nachgewiesen werden, in beiden Gruppen erhöhte sich das Sportvolumen nach der Rehabilitation.

Eine stärkere Steigerung der körperlichen Aktivität bei 87 Patienten mit diagnostizierter peripherer arterieller Verschlusskrankheit (pAVK) konnte Pochstein (2007) über einen Ansatz erreichen, der auf Handlungskontrollprozesse zielt. Bei der Interpretation der Ergebnisse müssen sowohl die relativ kleine Stichprobe der randomisierten Studie als auch die Frage der Übertragbarkeit auf das Reha-Setting beachtet werden.

### **5.3 Rehabilitation bei Krebserkrankungen**

Das Forschungsprojekt von Mehnert et al. (2007) zielte auf die Evaluation spezifischer psychoonkologischer Rehabilitationsangebote für psychisch besonders belastete Krebspatienten hinsichtlich psychosozialer und sozialmedizinischer Erfolgsparameter. Zu Beginn der Rehabilitation wiesen 60 % der 1.281 Patienten auffällige bis hohe Angstwerte und 32 % auffällige bis hohe Depressionswerte auf. In beiden Merkmalen verbesserten sich die Patienten vom Beginn der Rehabilitation bis zum Katamnesezeitpunkt nach einem Jahr signifikant. 161 besonders stark belastete Patienten erhielten eine verhaltensmedizinisch orientierte integrative psychoonkologische Intervention. Diese Gruppe konnte noch stärker bezogen auf ihre psychische Befindlichkeit profitieren. In der Einschätzung der Autoren bestätigen die Ergebnisse sowohl die Wirksamkeit der onkologischen Rehabilitation als auch die Effektivität des evaluierten Konzepts für psychisch besonders belastete Patienten. In einer aktuellen Studie berichten Bergelt et al. (2009) über hohe Zufriedenheitswerte bei 843 onkologischen Rehabilitanden. Auch der Reha-Erfolg wurde insgesamt positiv bewertet. Die Autoren betonen, dass sich keine Unterschiede zwischen ambulanter und stationärer Rehabilitation gezeigt haben. Die Ergebnisse sprechen nach ihrer Ansicht für einen weiteren Ausbau ambulanter Rehabilitationsangebote für Krebspatienten.

## 6. Methodische Aspekte der Wirksamkeitsstudien

Die für einen Wirksamkeitsnachweis aus methodischer Sicht als Goldstandard anzusehenden randomisierten Vergleichsstudien mit einer nicht rehabilitierten Kontrollgruppe konnten in den vorliegenden Studien nicht realisiert werden. Da die Versicherten einen Rechtsanspruch auf Rehabilitationsleistungen haben, würde die Realisierung eines solchen Designs zunächst rechtliche Fragen aufwerfen. Genauso schwer wiegt, dass nach den vorliegenden Erfahrungen und Erkenntnissen nur sehr wenige Rehabilitanden ihre Zustimmung zu einer Randomisierung (Reha vs. keine Reha) geben würden. Dabei ist zu bedenken, dass sich die Rehabilitanden nach vorliegenden Ergebnissen bewusst für eine Rehabilitation entscheiden: Die Versicherten sehen sich dann als reha-bedürftig an, wenn ihr subjektiver Gesundheitszustand schlecht ist und sie sich von einer Rehabilitation eine Besserung erwarten. In dem Projekt von Schweikert et al. (2007) waren nur einige wenige Rehabilitanden bereit, sich zufällig einer stationären oder einer ambulanten Rehabilitation zuweisen zu lassen. Dies weist darauf hin, dass die Rehabilitanden konkrete Erwartungen an die Rehabilitationsleistung haben. Neben den klaren Präferenzen der Versicherten spricht zudem gegen eine randomisierte Studie (Reha vs. keine Reha), dass mittlerweile eine große Zahl von Hinweisen auf die Wirksamkeit der medizinischen Rehabilitation vorliegt, die es als weiteres Argument ethisch bedenklich erscheinen lassen, den Versicherten eine Rehabilitationsleistung vorzuenthalten. In den vorliegenden rehabilitationswissenschaftlichen Evaluationsstudien werden immer häufiger bewährte Reha-Ansätze mit innovativen Konzepten randomisiert verglichen. Dieser Ansatz ist ein praktikabler Weg, komplexe Interventionen systematisch zu evaluieren und weiterzuentwickeln. Die vermehrte Verwendung dieses Designs in der deutschen Reha-Forschung ist ein wichtiger methodischer Fortschritt.

Unter der Voraussetzung, dass das übliche Reha-Verfahren in der Vergleichsbedingung als Quasi-Placebo angesehen wird, kann das skizzierte Design einen Wirksamkeitsnachweis erbringen. Einschränkend muss allerdings angemerkt werden, dass der beschriebene Ansatz nicht die Effizienz der Rehabilitation belegen kann, da nicht auszuschließen ist, dass andere Behandlungsverfahren ein gleiches Ergebnis mit geringeren Kosten erreichen könnten.

Bezogen auf Wirksamkeitsnachweise mit dem beschriebenen Untersuchungsansatz wurden Bedenken dahingehend geäußert, dass die übliche Rehabilitation nicht als Vergleichsbedingung angesehen werden kann, da von ihr nicht nur keine positiven, sondern vielmehr negative Effekte ausgehen könnten. Wäre nun der weiterentwickelte Reha-Ansatz wirksamer als das übliche Reha-Konzept dann sei dies kein Beleg für die Wirksamkeit, da er möglicherweise nur die negativen Effekte der üblichen Rehabilitation vermeidet und damit der „natürliche Verlauf“ ein gleiches Ergebnis bringen würde. Es liegen jedoch keine empirischen Belege für wesentliche Negativeffekte der Rehabilitation vor. Vor dem Hintergrund, dass viele Rehabilitanden bereits über mehrere Jahre im akutmedizinischen Bereich behandelt wurden, wäre eine ähnliche Argumentation auch bei einem Vergleich mit der „Usual-Care“ in der Akutversorgung denkbar.

## 7. Fazit

Nach den vorliegenden Ergebnissen des Förderschwerpunkts „Rehabilitationswissenschaften“ und den betrachteten jüngeren Studien seit 2005 verbessert die medizinische Rehabilitation den Gesundheitszustand der Rehabilitanden unmittelbar nach der Maßnahme klinisch bedeutsam. Die Wirksamkeit bleibt für viele Rehabilitanden auch mittel- bzw. längerfristig erhalten. Diese Ergebnisse bestätigen somit eine erste Litera-

turübersicht zu 39 Evaluationsstudien aus den Jahren 1990 bis 2000 (Haaf & Schliehe, 2001a+b). Mit den Übersichtsarbeiten von Haaf & Schliehe (2001a+b), Haaf (2005) und Faller (2005) sowie den hier betrachteten aktuelleren Studien liegt nun für den Zeitraum seit 1990 ein Überblick zu mehr als 100 Forschungsarbeiten vor, die insgesamt überzeugende Belege für die Wirksamkeit der medizinischen Rehabilitation liefern.

Die Rehabilitation schafft dadurch die Voraussetzungen für eine weitere Erwerbstätigkeit. Dies bestätigt auch die statistische Analyse des sozialmedizinischen Verlaufs mittels der Routinedaten der Rentenversicherung. Die rehabilitierten Patienten sind meist seit Jahren chronisch krank und weisen neben körperlichen Symptomen und Beeinträchtigungen auch psychische Beschwerden auf. Die Rehabilitation bietet diesen Patienten die Möglichkeit einer intensiven und ganzheitlichen Auseinandersetzung mit ihrer chronischen Erkrankung, die in der Akutversorgung nicht gegeben ist.

Die Projekte des Förderschwerpunkts zeigen konkrete Perspektiven für die wissenschaftlich fundierte Weiterentwicklung der medizinischen Rehabilitation auf. Dies gilt vor allem für die positiven Erfahrungen mit berufsbezogenen Behandlungsmodulen und mit der Patientenschulung. Teils werden die gewonnenen Erkenntnisse bereits in der Praxis umgesetzt. Besonders intensive Bestrebungen zur Verbesserung der Wirksamkeit sind für die Rehabilitation bei chronischen Rückenschmerzen in den letzten Jahren zu beobachten. Hier erwiesen sich nun psychologisch fundierte und beruflich orientierte Ansätze auch als nachhaltig wirksam. Hervorzuheben ist das umfassende Konzept der Arbeitsgruppe um Greitemann, das mittlerweile auch in weiteren Reha-Einrichtungen erfolgreich implementiert werden konnte. Die Effektstärken sind bei den untersuchten neueren Ansätzen häufig auch mittel- und längerfristig deutlich höher als die von Hüppe & Raspe (2003, 2005) für ältere Studien berichteten. Wegen der ausgesprochen positiven Ergebnisse zur

Effektivität und Effizienz von Patientenschulungen in der medizinischen Rehabilitation in verschiedenen Indikationen wurde als Umsetzungsprojekt des Förderschwerpunkts die Entwicklung eines Zentrums Patientenschulung gefördert. Übergreifende Ziele des Projektes waren die Ermittlung von Ansatzpunkten zur Optimierung von Patientenschulungen und die Dissemination von Forschungsergebnissen in die Praxis (Ströbl et al., 2007). Eine zentrale Aufgabe des Zentrums Patientenschulung ist zudem die Entwicklung und Pflege einer Datenbank, in der Informationen über Schulungsprogramme enthalten sind und recherchiert werden können. Diese Datenbank ist mittlerweile online ([www.zentrum-patientenschulung.de](http://www.zentrum-patientenschulung.de)) verfügbar.

Aufgrund der wichtigen Zielsetzung und der positiven Forschungsergebnisse erscheint eine weitere Intensivierung der medizinisch-beruflichen Orientierung in der Rehabilitation besonders sinnvoll. Diese Entwicklung wird durch eine Forschungsarbeit unterstützt, in der Interventionen bzw. Interventionsbausteine zur gezielten Bearbeitung beruflicher Problemlagen systematisch gesammelt und wissenschaftlich bewertet werden (Gerlich et al., 2007). Dieses durch die Deutsche Rentenversicherung als Umsetzungsprojekt geförderte Forschungsvorhaben wird es den Reha-Praktikern erleichtern, geeignete Konzepte zur beruflich orientierten Rehabilitation auszuwählen und einzusetzen.

Nicht in allen Studien zur telefonischen Nachsorge konnten gleichermaßen positive Effekte aufgezeigt werden. Dies kann auch an den Unterschieden zwischen den Ansätzen beispielsweise bei der Häufigkeit und Dauer der Telefonate liegen. In der Tendenz sprechen die Ergebnisse jedoch für eine Wirksamkeit des Nachsorgeansatzes. Die Vorteile der telefonischen Nachsorge liegen vor allem in der Flexibilität und Anpassungsfähigkeit an die Bedürfnisse der Patienten und der betreibenden Institution, der Unabhängigkeit von Entfernungen sowie den relativ geringen Kosten.

Die methodische Qualität der rehabilitationswissenschaftlichen Studien hat sich in den letzten Jahren kontinuierlich verbessert. Die Forschungsarbeiten aus dem Förderschwerpunkt „Rehabilitationswissenschaften“ weisen ausreichend große Fallzahlen, ein zumindest mittelfristiges Follow-up und eine detaillierte Dokumentation in den Abschlussberichten auf (Haaf, 2005). Bei zwei der 38 betrachteten Projekte handelte es sich um systematische Literaturreviews. In 17 Projekten wurden verschiedene Reha-Konzepte bzw. Settings miteinander verglichen. Dabei wurde in 12 Untersuchungen randomisiert. Nur in einem dieser Projekte wurden dabei die Rehabilitanden entsprechend ihrer Anmeldung zur Rehabilitation den Studiengruppen zugewiesen (Pseudo-Randomisierung). Hinzu kommen 9 Forschungsarbeiten, über die Faller (2005) berichtet, von denen 8 mit kontrolliertem Design bzw. mit einer Vergleichsgruppe durchgeführt wurden. In zwei Studien erfolgte eine randomisierte Zuordnung.

In zwei onkologischen Studien wurde eine nicht rehabilitierte Vergleichsgruppe einbezogen (vgl. Haaf, 2005). Es zeigte sich allerdings, dass die Ausgangswerte der beiden Gruppen nicht vergleichbar waren, da überwiegend stärker beeinträchtigte und belastete Patienten eine Rehabilitation in Anspruch nehmen. Ähnliche Ergebnisse erzielten Waldmann et al. (2008) in einer jüngeren Untersuchung für die kardiologische Rehabilitation. Hier verbesserten sich die stärker belasteten und älteren Rehabilitanden stärker als Patienten, die keine AHB in Anspruch nahmen.

Aus den Jahren seit 2005 wurden 21 Arbeiten betrachtet. Insgesamt 19 Forschungsprojekte wurden in einem Kontrollgruppendesign durchgeführt, davon 11 mit randomisierter Zuordnung zu den Gruppen. Die Meta-Analyse von Bethge & Müller-Fahrnow (2008) bewertet die methodische Qualität von vier der fünf einbezogenen Studien als gut.

In den Studien geht es mittlerweile verstärkt um die wissenschaftlich fundierte Weiterentwicklung der Rehabilitation. Die Effektivität der Standardrehabilitation im Sinne einer Legitimationsforschung zu untersuchen, erscheint vor dem Hintergrund der vorhandenen Wirksamkeitsnachweise nicht mehr im Mittelpunkt zu stehen. Nach den vorliegenden Ergebnissen kann eine systematisch weiterentwickelte Rehabilitation als wirksam gelten.

## 8. Literatur

**Bergelt, C., Lehmann, C., Beierlein, V., Hagen-Aukamp, C., Kerschgens, C., Rhee, M., Frühauf, S., Meise, U., Otto, J., Berger, D. & Koch, U.** (2009): Prozess- und Ergebnisbewertung von InanspruchnehmerInnen ambulanter und stationärer onkologischer Rehabilitationsmaßnahmen. DRV-Schriften, Band 83, 389-391.

**Bethge, M., Herbold, D., Trowitzsch, L. & Jacobi, C.** (2009): Berufliche Leistungsfähigkeit nach einer medizinisch-beruflich orientierten orthopädischen Rehabilitation: Eine cluster-randomisierte Studie zur Wirksamkeit eines multimodalen Programms mit kognitiv-behavioralem Ansatz. DRV-Schriften, Band 83, 188-190.

**Bethge, M. & Müller-Fahrnow, W.** (2008): Wirksamkeit einer intensivierten stationären Rehabilitation bei muskuloskelettalen Erkrankungen: Systematischer Review und Meta-Analyse. Die Rehabilitation, 47/4. 200-209.

**Beutel, M. E., Hillert, A., Kayser, E., Cuntz, U. & Schatzenburg, L.** (2002): Berufliche Belastungserprobung als integrierter Bestandteil der psychosomatischen Rehabilitation - Evaluation im Rahmen eines randomisierten Kontrollgruppendesigns: Indikationen, Prädiktoren, Wirksamkeit. Abschlussbericht.

**Bischoff, C., Gönner, S., Ehrhardt, M., Limbacher, K. & Jäger, R. S.** (2003): Ambulante prä- und poststationäre Maßnahmen - ein Beitrag zur Flexibilisierung der stationären psychosomatischen Versorgung. Abschlussbericht.

**Blume, C., Streibelt, M., Thren, K. & Müller-Fahrnow, W.** (2007): Effizienz berufsbezogener Maßnahmen: Ökonomische Evaluation eines Medizinisch Beruflich Orientierten Modells in der orthopädischen Rehabilitation. DRV-Schriften, Band 72, 250-254.

**Bönisch, A. & Ehlebracht-König, I.** (2001): Evaluation eines Schulungsprogramms für Patienten mit Spondylitis ankylosans. Abschlussbericht.

**Brenner, H. & Rothenbacher, D.** (2002): Bedeutung kurzfristiger Reha-Erfolge für die Langzeitprognose. Abschlussbericht.

**Buschmann-Steinhage, R. & Zollmann, P.** (2009): Beschäftigung, Arbeitslosigkeit und Arbeitsunfähigkeit im 2-Jahresverlauf nach medizinischer Rehabilitation 2004: Neue Ergebnisse aus der Reha-Statistik-Datenbasis. DRV-Schriften, Band 83, 176-179.

**China, C., Mittag, O., Hoberg, E., Jürs, E., Kolenda, K. D., Richardt, G. & Raspe, H. H.** (2004): Telefonische Nachsorge bei KoronarpatientInnen. DRV-Schriften, Band 52, 151-153.

**China, C., Mittag, O., Hoberg, E., Jürs, E., Kolenda, K. D., Richardt, G. & Raspe, H. H.** (2005): Telefonische Nachsorge bei KoronarpatientInnen: Ergebnisse der Lübecker Nachsorgestudie. DRV-Schriften, Band 59, 328-330.

**Delbrück, H. & Witte, M.** (2004): Vergleich onkologischer Rehabilitationsmaßnahmen und -strukturen in Ländern der Europäischen Gemeinschaft. Abschlussbericht.

**Dibbelt, S., Büschel, C. & Greitemann, B.** (2005a):

Kontrollwahrnehmung, Krankheitsbewältigung und Wohlbefinden: Welche Ressourcen vermitteln positive Effekte des multimodalen Therapieprogramms „Rückenfit“ für Patienten mit chronischen Rückenschmerzen? DRV-Schriften, Band 59, 284-285.

**Dibbelt, S., Büschel, C. & Greitemann, B.** (2005b): Langfristige Effekte eines aktivierenden multimodalen Therapieprogramms für Patienten mit chronischen Rückenschmerzen. Unveröffentlichtes Manuskript.

**Faller, H., Reusch, A., Vogel, H., Ehlebracht-König, I. & Petermann, F.** (2005): Patientenschulung. Die Rehabilitation, 44, 21-31.

**Fikentscher, E., Schreiber, B., Müller, K., Bahrke, U. & Bandemer-Greulich, U.** (2005): Risikofaktorenspezifische Interventionsprogramme bei Rehabilitanden mit chronischen Rückenschmerzen. Abschlussbericht.

**Fritschka, E.** (2002): Evaluation der Schulung von nierenkranken Patienten im Rahmen der stationären Rehabilitation (Nierenschulung). Abschlussbericht.

**Fröhlich, S. M., Greitemann, B. & Niemeyer, R.** (2009): Langfristige Effekte eines multimodalen Behandlungsprogramms für Patienten mit chronischen Rückenschmerzen in der stationären orthopädischen Rehabilitation: Können die Effekte des Integrierten orthopädisch-psychosomatischen Konzepts repliziert werden? DRV-Schriften, Band 83, 320-321.

**Gerlich, C., Neuderth, S. & Vogel, H.** (2007): Systematisierung berufsbezogener Interventionen in der medizinischen Rehabilitation. DRV-Schriften, Band 72, 252-254.

**Greitemann, B., Dibbelt, S. & Büschel, C.** (2004): Berufliche Orientierung in der stationären orthopädischen Rehabilitation: Haben berufsbezogene Beratungs- und

Behandlungsangebote langfristige Effekte? DRV-Schriften, Band 52, 235-237.

**Guzman, J., Esmail, R., Karjalainen, K., Malmivaara, A., Irvin, E. & Bombardier, C.** (2001): Multidisciplinary rehabilitation for chronic low back pain: systematic review, *British Medical Journal*, 322/7301, 1511-1516.

**Haaf, H. G. & Schliehe, F.** (2001a): Wie wirksam ist die medizinische Rehabilitation? Übersicht zu den häufigsten Krankheitsgruppen (Teil 1). *Wirtschaftspsychologie*, 3/2, 111-124.

**Haaf, H. G. & Schliehe, F.** (2001b): Wie wirksam ist die medizinische Rehabilitation? Übersicht zu den häufigsten Krankheitsgruppen (Teil 2). *Wirtschaftspsychologie*, 3/3, 168-177.

**Haaf, H. G.** (2005): Ergebnisse zur Wirksamkeit der Rehabilitation. *Die Rehabilitation*, 44/5, 259-276.

**Hampel, P., Tlach, L., Gräf, T., Krohn-Grimberghe, B., Mantel, F. & Mohr, B.** (2009): Zur Wirksamkeit eines Trainings zur Depressionsbewältigung für Patienten mit chronisch unspezifischem Rückenschmerz in der stationären Rehabilitation – Eine 1-Jahres-Follow up-Studie. *DRV-Schriften*, Band 83, 322-325.

**Höner, O., Sudeck, G., Keck, M. & Lowis, H.** (2008): Auswirkungen einer motivationalen und volitionalen Sportförderung: Eine empirische Analyse mittel- und langfristiger Effekte auf der Verhaltens- und Gesundheitsebene. *DRV-Schriften*, Band 77, 199-201.

**Huber, C., Steffanowski, A., Löschmann, C., Nübling, R., Schmidt, J., Mayer, M. M. & Wittmann, W. W.** (2004): Metaanalyse der Effekte stationärer psychosomatischer Rehabilitation. *DRV-Schriften*, Band 52, 535-537.

**Huber, G., Wiskemann, J. & Heilmeyer, P.** (2007): Kraft-

training und die Rehabilitation des chronischen Rückenschmerzes - Eine kontrollierte und randomisierte Studie. *DRV-Schriften*, Band 72, 394-395.

**Hüppe, A. & Raspe, H.** (2003): Die Wirksamkeit stationärer medizinischer Rehabilitation in Deutschland bei chronischen Rückenschmerzen: Eine systematische Literaturübersicht 1980-2001. *Die Rehabilitation*, 42, 143-154.

**Hüppe A. & Raspe H.** (2005): Zur Wirksamkeit von stationärer medizinischer Rehabilitation in Deutschland bei chronischen Rückenschmerzen: Aktualisierung und methodenkritische Diskussion einer Literaturübersicht. *Die Rehabilitation*, 44, 24-33.

**Keck, M.** (2009): Intensivierte Nachsorge zur Verbesserung der kardiovaskulären Risikofaktoren sowie anderer relevanter Reha-Outcomes bei Patienten mit manifester koronarer Herzerkrankung mittels Telefondnachsorge. *DRV-Schriften*, Band 83, 357-358.

**Keck, M., Budde, H. G., Kallinke, D. & Behrens, J.** (2000): Berufsbezogene Orientierung und praktische Erprobung während der Phase II der kardiologischen Rehabilitation als Instrumente der beruflichen Früh-Rehabilitation. In: Bengel, J. & Jäcker, W. H. (Hrsg.). *Zielorientierung in der Rehabilitation*. Regensburg: Roderer-Verlag, 49-65.

**Kittel, J.** (2007): Effekte berufsorientierter Interventionen in der kardiologischen Rehabilitation ein Jahr nach der Rehabilitation – Ergebnisse einer randomisierten Kontrollgruppenstudie. *DRV-Schriften*, Band 72, 259-260.

**Kittel, J., Liebich E. & Karoff, M.** (2008): Gesundheitsökonomische Evaluation einer intensiven berufsbezogenen kardiologischen Rehabilitation. *DRV-Schriften*, Band 77, 421-423.

**Koch, S. & Hillert, A.** (2006): Was leistet eine berufsbezogene Therapiegruppe für die psychosomatische Rehabilitation? Ergebnisse von 12-Monatskatamnesen einer kontrollierten Evaluation. DRV-Schriften, Band 64, 52-54.

**Kohlmeyer, M., Redaelli, M., Büscher, G., Schwitala, B., Lauterbach, K. W. & Mayer-Berger, W.** (2008): Nachhaltigkeit in der Sekundärprävention. Sekundärprävention bei Patienten mit Koronarer Herzkrankheit durch Anschlussheilbehandlung und anschließende konzeptintegrierte Nachsorge (SeKoNa). DRV-Schriften, Band 77, 404-405.

**Krauth, C., Rieger, J., Mau, W., Bräuer, W. & Schwartz, F.W.** (2003): Gesundheitsökonomische Evaluation eines Patientenschulungsprogramms Spondylitis Ankylosans in der stationären Rehabilitation. In: Petermann, F. (Hrsg.): Prädiktion, Verfahrensoptimierung und Kosten in der medizinischen Rehabilitation. Regensburg: Roderer-Verlag, 293-307.

**Leidl, R., Schweikert, B. & Seitz, R.** (2002): Gesundheitsökonomische Analysen. Abschlussbericht.

**Löschmann, C., Steffanowski, A., Schmidt, J., Wittmann, W. W. & Nübling, R.** (2005): Evidenz stationärer psychosomatischer Rehabilitation – Ergebnisse der MESTA-Studie. DRV-Schriften, Band 59, 438-440.

**Mangels, M., Schwarz, S., Holme, M. & Rief, W.** (2006): Evaluation eines verhaltensmedizinischen Ansatzes in der orthopädischen Rehabilitation – Eine randomisierte, kontrollierte Studie. DRV-Schriften, Band 64, 337-339.

**Mangels, M., Schwarz, S., Holme, M., Worringer, U. & Rief, W.** (2008): Verbesserung der Nachhaltigkeit in der stationären orthopädischen Rehabilitation durch verhaltensmedizinische Therapiebausteine oder ein telefonisches Nachsorgeprogramm – Eine randomisierte, kontrollierte Studie. DRV-Schriften, Band 77, 330-332.

**Mau, W. & Bräuer, W.** (2003): Sozialmedizinische Langzeitprognose und Inanspruchnahme von Leistungen zur Rehabilitation von Patienten mit chronischer Polyarthrit. In: Petermann, F. (Hrsg): Prädiktion, Verfahrensoptimierung und Kosten in der medizinischen Rehabilitation. Regensburg: Roderer-Verlag, 245-292.

**Mehnert, A., Müller, D., Kegel, D., Gärtner, U., Friedrich, G., Bootsvelde, W., Leibbrand, B., Barth, J., Gaspar, M., Jacobi, C., Berger, D. & Koch, U.** (2007): Längerfristige Wirksamkeit eines integrativen Psychoonkologischen Rehabilitationskonzepts hinsichtlich psychosozialer und sozialmedizinischer Erfolgsparameter – Eine multizentrische Evaluationsstudie. DRV-Schriften, Band 72, 484-485.

**Meng, K., Seekatz, B., Roßband, H., Worringer, U., Falter, H. & Vogel, H.** (2009): Kurzfristige Effektivität des Curriculums Rückenschule des Gesundheitstrainingsprogramms der Deutschen Rentenversicherung Bund. DRV-Schriften, Band 83, 318-320.

**Miche, E., Roelleke, E., Zoller, B., Huerst, M., Wirtz, U., Tietz, M. & Radzewitz, A.** (2008): Telefonisches Nachsorgeprogramm für Patienten mit chronischer Herzinsuffizienz – Ergebnisse einer einjährigen Betreuung. DRV-Schriften, Band 77, 406.

**Müller-Fahrnow, W., Greitemann, B., Radoschewski, F. M., Gerwinn, H. & Hansmeier, T.** (2005): Berufliche Orientierung in der medizinischen Rehabilitation und Leistungen zur Teilhabe am Arbeitsleben. Die Rehabilitation, 44/5, 287-296.

**Mussgay, L., Schmidt, F., Morad, E. & Rüdell, H.** (2004): Effekte eines aeroben Ausdauertrainings auf die autonome kardiovaskuläre Regulation bei Angst- und Somatisierungspatienten in stationärer psychosomatischer Rehabilitation. DRV-Schriften, Band 52, 551-553.

**Mussgay, L. & Rüdell, H.** (2005): Erschöpfung und

Behandlungserfolg in der stationären Rehabilitation: Förderliche Effekte von Ausdauersport. DRV-Schriften, Band 59, 459-460.

**Pochstein, F.** (2007): Effekte von Handlungsplanung und -kontrolle auf die körperliche Aktivität von pAVK-Patienten. DRV-Schriften, Band 72, 438-439.

**Schliehe, F. & Haaf, H. G.** (1996): Zur Effektivität und Effizienz der medizinischen Rehabilitation. Deutsche Rentenversicherung, Heft 10-11, 666-689.

**Schmitz, B., Schuhler, P., Handke-Raubach, A., Jung, A., Husen, E., Wagner, A., Gönner, S., Limbacher, K. & Vogelgesang, M.** (2002): Die Rehabilitation von Persönlichkeitsstörungen in der stationären verhaltenstherapeutischen Psychosomatik und Suchttherapie. Abschlussbericht.

**Schönberger, C. & v. Kardorff, E.** (2002): Angehörige Krebskranker in der Rehabilitation. Abschlussbericht.

**Schwarz, S., Mangels, M., Holme, M. & Rief, W.** (2007): Langzeitevaluation eines verhaltensmedizinischen Ansatzes in der orthopädischen Rehabilitation – Eine randomisierte, kontrollierte Studie. DRV-Schriften, Band 72, 381-383.

**Schweikert, B., Hahmann, H., Imhof, A., Koenig, W., Kropf, C., Liu, Y., Muche, R., Steinacker, J. & Leidl, R.** (2007): Kosten und Lebensqualität bei ambulanter vs. stationärer kardiologischer Rehabilitation – Gesundheitsökonomische Ergebnisse aus der SARA-Studie. DRV-Schriften, Band 72, 345-347.

**Ströbl, V., Friedl-Huber, A., Küffner, R., Reusch, A., Vogel, H. & Faller, H.** (2007): Vorbereitung eines Zentrums Patientenschulung. Abschlussbericht.

**Waldmann, A., Schwaab, B., Katalinic, A., Poppe, B.,**

**Sheikhzadeh, A. & Raspe, H.** (2008): Mortalität und Morbidität 12 Monate nach akutem koronarem Ereignis: Entlassung aus der Akutklinik versus kardiologische Rehabilitation. DRV-Schriften, Band 77, 425-427.

**Weis, J., Moser, M. T. & Bartsch, H. H.** (2002): Zielorientierte Evaluation stationärer onkologischer Rehabilitationsmaßnahmen. ZESOR-Studie. Abschlussbericht.

**Welk, H., Koch, U., Bergelt, C. & Lehmann, C.** (2005): Krebsrehabilitation und Partnerschaft – Die Rolle des Partners

bei der Inanspruchnahme und im Verlauf der stationären onkologischen Rehabilitation. Abschlussbericht.

**Wittmann, M., Spohn, S. & Petro, W.** (2001): Chronische Bronchitis: Eine Anleitung zur besseren Krankheitsbewältigung. Praxis Klinische Verhaltensmedizin und Rehabilitation, 54, 151-156.

**Zwerenz, R., Knickenberg, R.J., Schattenburg, L. & Beutel, M.E.** (2007): Kurz- und langfristige Behandlungseffekte einer tiefenpsychologisch fundierten Gruppentherapie für beruflich belastete Patienten in der stationären psychosomatischen Rehabilitation. DRV-Schriften, Band 72, 243-245.

**Zwerenz, R., Hillert, A., Koch, S., Wendt, T., Schröder, K. & Beutel, M.E.** (2008): Evaluation des berufsbezogenen Schulungsprogramms „Gesundheitstraining Stressbewältigung am Arbeitsplatz (GSA)“ in der kardiologischen Rehabilitation: Langfristige Behandlungsergebnisse (6 Monate nach Entlassung). DRV-Schriften, Band 77, 229-231.

**Zwingmann, Ch., Buschmann-Steinhage, R., Gerwin, H. & Klosterhuis, H.** (2004): Förderschwerpunkt „Rehabilitationswissenschaften“: Ergebnisse – Umsetzung – Erfolge und Perspektiven. Die Rehabilitation, 43, 260-270.

# Methodische und konzeptionelle Anforderungen an die Ergebnismessung in der medizinischen Rehabilitation<sup>1</sup>

PD Dr. Erik Farin

Universitätsklinikum Freiburg

Abteilung Qualitätsmanagement und Sozialmedizin

## Inhaltsverzeichnis

1.	Einleitung	49
2.	Patientenorientierung im Rahmen der Ergebnismessung	51
3.	ICF-Orientierung im Rahmen der Ergebnismessung	53
4.	Die Verwendung von Modellen der Item-Response-Theorie	57
5.	Die Berücksichtigung der Variabilität von Itembedeutung und Itemschwierigkeit	59
6.	Adaptives Testen	63
7.	Empfehlungen für die Praxis der Ergebnismessung in der medizinischen Rehabilitation	65
8.	Forschungsbedarf	66
9.	Literatur	72

---

<sup>1</sup> Teile des Beitrags stammen aus der Habilitationsschrift des Autors: „Ergebnismessung und Qualitätssicherung in der medizinischen Rehabilitation“, vgl. auch Farin (2008).

## 1. Einleitung

Mit der Ergebnismessung in der medizinischen Rehabilitation wird das Ziel verfolgt, das Resultat der Behandlung zu quantifizieren und damit einer Beurteilung zugänglich zu machen. Entsprechend der Definition von „Messung“ steht dabei die quantitative Bestimmung des Wertes einer Ergebnisgröße im Mittelpunkt. Die Ergebnisgröße soll das Behandlungsergebnis erfassen, also einen postrehabilitativen Zustand oder eine Merkmalsausprägung des Patienten, die auf die vorangegangene Rehabilitationsmaßnahme zurückgeführt wird.

Ergebnismessungen finden unter anderem im Kontext der routinemäßigen, patientenbezogenen Kontrolle individueller Behandlungsverläufe, im Rahmen der Erhebungen externer Qualitätssicherungsprogramme und innerhalb von Evaluationsstudien zur Bewertung rehabilitativer Interventionen statt. Die Frage nach der adäquaten Form der Ergebnismessung stellt sich unabhängig vom jeweiligen Design der Studie, also z.B. sowohl bei randomisiert-kontrollierten Studien als auch bei Beobachtungsstudien ohne Kontrollgruppen. Adäquate Messvorgänge bilden die Voraussetzung Erkenntnis schaffender wissenschaftlicher Studien, da komplexe statistische Analysen keine Aussagekraft besitzen, wenn die erforderliche Messqualität nicht gegeben ist. Bohrnstedt (1983) hat dies wie folgt ausgedrückt: „Measurement is a sine qua non of any science.“

Die Ergebnismessung kann in der Rehabilitation auf verschiedenen methodischen Zugängen basieren, wobei die am häufigsten gewählten (vgl. z.B. Haigh et al., 2001) Fragebögen zur Erfassung von Patient- oder Arzturteilen (z.B. IRES-Fragebogen, Bührlen et al., 2005, Barthel-Index, Mahoney & Barthel, 1965), Leistungstests (z.B. Gehstrecken), apparative Messungen von Körperfunktionsparametern (z.B. Blutdruck) und ärztliche Körperfunktionsuntersuchungen (z.B.

Messung der Gelenkbeweglichkeit in Winkelgraden) darstellen.

Hier sollen Fragebögen im Mittelpunkt der Erörterung stehen. Diese besitzen in der medizinischen Rehabilitation eine besondere Bedeutung, da sie gut geeignet sind, eine patientenseitige Einschätzung der Einschränkungen von Aktivitäten und Teilhabe zu erfassen. Eine patientenseitige Messung ist von hoher Relevanz, da in einem patientenorientierten Versorgungssystem (vgl. Mead & Bower, 2000) die Wahrnehmungen des Patienten ein entscheidendes Ergebniskriterium darstellen. Whiteneck und Gerhart (2001, S. 306) schlussfolgern deshalb: „If what we really care about is the consumer's perspective, survey research is the methodology of choice“. In den USA haben die Forderungen zur Einbeziehung des Patienten in die Ergebnismessung zu nationalen Initiativen geführt, die das Ziel verfolgen standardisierte Itembanken zu entwickeln, mit denen erkrankungsübergreifend so genannte „patient-reported outcomes“ erfasst werden können (PROMIS-Initiative, Cella et al., 2007).

Eine Fokussierung auf Aktivitäten und Teilhabe ist angezeigt, da die zentralen Zielsetzungen der Rehabilitation, die auch die rechtliche Grundlage der Maßnahmengewährung abbilden, die Reintegration in das Berufsleben und die selbständige Bewältigung des Alltagslebens darstellen. Ersteres gilt primär für die von der Rentenversicherung finanzierten Rehabilitationsmaßnahmen (vgl. z.B. Verband Deutscher Rentenversicherungsträger, 2000a), letzteres für das System der gesetzlichen Krankenversicherung (vgl. z.B. Leistner & Beyer, 2005).

Im Folgenden sollen verschiedene konzeptionelle und methodische Anforderungen an die Ergebnismessung in der medizinischen Rehabilitation, die auch wesentliche Diskussionspunkte der aktuellen Forschung wiedergeben (vgl. Dekker et al., 2005; Heinemann, 2005;

Jette & Haley, 2005), näher erörtert werden: Patientenorientierung, ICF-Orientierung, die Verwendung von Item-Response-Modellen, die Berücksichtigung der Variabilität von Itembedeutung und Itemschwierigkeit sowie das adaptive Testen.

## 2. Patientenorientierung im Rahmen der Ergebnismessung

Der gesundheitspolitische Trend der Abkehr von einem versorgerorientierten Gesundheitssystem zu einem patienten- oder nutzerorientierten System, der in den angloamerikanischen Ländern unter den Begriffen „provider-centered“ versus „patient-centered“ bzw. „provider-driven“ versus „consumer-driven“ stattfand (Gauthier & Clancy, 2004) und dort auch Auswirkungen auf die Ergebnismessung besaß (Stryer et al., 2004), hat auch in Deutschland zu entsprechenden veränderten Schwerpunktsetzungen geführt, die von der Politik und politiknahen Gremien (z.B. Gutachten des Sachverständigenrats für die Konzertierte Aktion im Gesundheitswesen, 2003) vermittelt werden.

Das Konzept der Patientenorientierung weist nach Mead und Bower (2000) fünf zentrale Dimensionen auf: 1. die bio-psycho-soziale Perspektive, die den Patienten in seinem komplexen Lebensumfeld betrachtet und neben biologischen auch psychologische und soziale Einflussfaktoren einbezieht, 2. die Berücksichtigung der subjektiven Sichtweise des Patienten und seiner persönlichen Erfahrung von Krankheit, 3. eine gleichberechtigtere Verteilung von Verantwortung und Entscheidungskompetenz als in der traditionellen Patient-Behandler-Beziehung (vgl. das Konzept des „shared decision making“), 4. die Berücksichtigung der Patient-Behandler-Beziehung als wesentlichen Einflussfaktor des Behandlungserfolgs und 5. die Beachtung der Subjektivität des Behandlers.

Die oben beschriebenen Aspekte von Patientenorientierung lassen sich auch auf den Kontext der Ergebnismessung beziehen. Wenn Patienten im medizinischen Behandlungsprozess als „Partner“ (vgl. z.B. Scheibler, 2005) angesehen werden, ist es nur konsequent, diese Forderung im Sinne der partizipatorischen Forschung (vgl. Abma, 2006) auf den Forschungsprozess und damit auch auf den Aspekt der Ergebnismessung zu übertragen. Der Patient soll nicht mehr – wie im traditionellen Verständnis – als Objekt der Forschung, sondern als aktiver Partner des Forschers angesehen werden (vgl. Caron-Flinterman et al., 2005).

Vier Aspekte der Einbeziehung von Patienten in die Ergebnismessung lassen sich unterscheiden:

1. die Erfassung von patientenseitigen Behandlungszielen, Patientenerwartungen und Patientenpräferenzen sowie deren Berücksichtigung bei der Ergebnisbewertung
2. die Beteiligung der Patienten an der Entwicklung von Assessmentinstrumenten
3. die Erfassung von Patientenbewertungen des Behandlungsergebnisses, insbesondere im Hinblick auf die Bereiche, in denen der Patient als „Experte“ anzusehen ist (z.B. Beeinträchtigungen seiner Aktivitäten und deren Auswirkungen auf die individuelle Teilhabe)
4. die Berücksichtigung der Subjektivität des Behandlers sowie der Patient-Behandler-Interaktion als relevante Einflussfaktoren der Ergebnisse.

### 3. ICF-Orientierung im Rahmen der Ergebnismessung

Die im Mai 2001 von der Weltgesundheitsorganisation WHO verabschiedete „Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit“ (ICF) unterscheidet drei Komponenten der Gesundheit (Körperfunktionen und -strukturen, Aktivitäten, Teilhabe) und berücksichtigt als Einflussvariablen der Gesundheit persönliche und umweltbezogene Kontextfaktoren. Mit der ICF liegt ein terminologisches System vor, welches eine international einheitliche Klassifikation für die Beschreibung von Gesundheitszuständen darstellt und die Funktion einer gemeinsamen Sprache der in der medizinischen Rehabilitation tätigen Berufsgruppen übernehmen könnte (Jette, 2006; Stucki et al., 2007). Es liegt deshalb nahe, eine Orientierung an der ICF auch in der Ergebnismessung vorzusehen.

Die Orientierung an der ICF kann eine Patientenorientierung unterstützen, da die von der ICF verwendeten Kategorien – insbesondere auf den Domänen der Aktivitäten und der Teilhabe – alltagsnah und damit passend zur Lebenswelt der Patienten formuliert sind. Dies belegen z.B. Studien, wie die von Coenen et al. (2006), die zeigt, dass Patienten, die mit offenen Fragen im Rahmen von Fokusgruppen befragt wurden (z.B. „Wenn Sie an Ihr alltägliches Leben denken, welches sind Ihre Probleme?“, Übersetzung vom Autor), zu einem großen Teil diejenigen Konstrukte nennen, die auf der Ebene der dreistelligen Kategorien von der ICF abgebildet werden.

Die Berücksichtigung der ICF bei der Ergebnismessung in der Rehabilitation kann prinzipiell auf drei Wegen erfolgen:

1. Die Kodierungen, die in der ICF (DIMDI, 2005)

vorgesehen sind (Quantifizierungen der Kategorien mit Hilfe von fünfstufigen Skalen), werden zur Ergebnismessung herangezogen, indem z.B. der Gesundheitsstatus vor und nach der Rehabilitation auf den für die Erkrankung relevanten Kategorien (vgl. die Bestimmung von Core-Sets, Stucki & Grimby, 2004) kodiert wird.

2. Items bzw. Skalen vorhandener Assessmentinstrumente werden den ICF-Kategorien zugeordnet (vgl. die so genannten „ICF linking rules“, Cieza et al., 2005), damit für die ICF-Kategorien, die für eine bestimmte Erkrankung relevant sind, geeignete Messverfahren ausgewählt werden können.
3. Es werden neue Assessmentinstrumente erarbeitet, die sich bezüglich Terminologie, Struktur und Inhalt unmittelbar an der ICF orientieren.

Die Nutzung der mit der ICF vorgegebenen Kodierungen scheint im Kontext der Ergebnismessung wenig Erfolg versprechend. Die ICF besitzt, wie mehrfach hervorgehoben wurde (Seger et al., 2004; Reed et al., 2005), lediglich die Funktion eines Klassifikationssystems, nicht die eines Assessmentinstruments. Die in der ICF vorgesehene Kodierung eines Gesundheitsproblems erfüllt nicht die wissenschaftlichen Anforderungen, die üblicherweise an ein Assessmentverfahren gestellt werden. So überrascht es auch nicht, dass sowohl die Inter-Rater-Übereinstimmung (z.B. Grill et al., 2007) als auch die Retest-Reliabilität (Okochi et al., 2005) von ICF-Kodierungen eher gering ausfallen. Unter anderem um diesem Problem zu begegnen, hat die American Psychological Association ein Projekt zur Entwicklung eines so genannten „Procedural Manual and Guide for a Standardized Application of the ICF“ gestartet. Das Ziel besteht darin, Personen aus den Gesundheitsberufen Hilfestellung bei der reliablen, validen und klinisch nützlichen ICF-Klassifikation zu bieten (vgl. Reed et al., 2005). Das Vorgehen ist jedoch auf den

Anwendungskreis der industrialisierten, englischsprachigen Länder begrenzt.

Die Zuordnung der Inhalte vorhandener Messverfahren zu den ICF-Kategorien ist nützlich, da damit eine thematische Ordnung erfolgt, die sich an einem einheitlichen Begriffssystem orientiert und vom Forscher in Abhängigkeit von den jeweiligen Core-Sets die passenden Assessmentinstrumente ausgewählt werden können. Der Ansatz kann aber keine Lösung für die Fälle bieten, in denen die vorhandenen Assessmentinstrumente keine adäquate Abdeckung der ICF-Kategorien leisten, was jedoch oft der Fall ist (vgl. z.B. Cieza et al., 2002).

Die Neuentwicklung von Assessmentinstrumenten zu Aktivitäten und Teilhabe, die sich unmittelbar auf die ICF beziehen, scheint eine viel versprechende Möglichkeit der Optimierung der Ergebnismessung in der Rehabilitation, weshalb im Folgenden näher auf die Vorteile und Probleme dieses Ansatzes eingegangen wird. Eine Umsetzung des Konzepts wird z.B. in den Arbeiten von Farin et al. (2006b), Farin et al. (2007b) und Farin & Fleitz (2009) beschrieben.

Die beiden zentralen Vorteile der Neuentwicklung von Assessmentinstrumenten beziehen sich auf die Chance zur Standardisierung der Inhalte von Assessmentinstrumenten sowie auf die Möglichkeit des sinnvollen Einsatzes generischer Verfahren (ggf. in Kombination mit Methoden des adaptiven Testens). Mit der Ausrichtung der Frageninhalte an der ICF ist die Zielvorstellung einer theoriegeleiteten und inhaltlich vereinheitlichten Messung der Beeinträchtigung von Aktivitäten und Teilhabe verbunden. Bisher existiert eine Vielzahl von Fragebögen, die den Anspruch erheben, Aspekte der Einschränkung von Aktivitäten und Partizipation zu erfassen, und vielfach auch für dieselben Erkrankungen entwickelt wurden. Die mangelnde Einheitlichkeit der Verfahren und Frageninhalte basiert weniger auf einem Dissens in grundlegenden Aspekten der Kon-

zeptualisierung von Aktivitäten und Partizipation als vielmehr auf einem fehlenden Rahmen für die Standardisierung der konkreten Operationalisierungen der Fragebogenitems (Long & Pavalko, 2004). Dies führt zu Datensätzen, die zwischen verschiedenen Studien nicht vergleichbar sind, so dass eine Integration der Befunde zur Effektivität der Rehabilitation erschwert wird.

Ein an der ICF orientiertes Assessment trägt darüber hinaus aufgrund der generischen Geltung der ICF der Forderung Rechnung, Verfahren zu entwickeln, die im Sinne eines generischen Ansatzes über verschiedene diagnostische Gruppen hinweg die Effektivität rehabilitativer Maßnahmen abbilden können (Haley et al., 2004a). Es ist nicht einsichtig, warum Patienten mit unterschiedlichen Diagnosen im Sinne der ICD (Internationale Klassifikation der Krankheiten), die aber dieselben Aktivitätseinschränkungen (z.B. Mobilitätseinschränkungen) aufweisen, diesbezüglich mit unterschiedlichen Assessmentinstrumenten befragt werden sollten. Es muss nur sichergestellt sein, dass das entsprechende Messverfahren das ganze Spektrum an Einschränkungen, von sehr leichten Aktivitätsproblemen bis hin zu sehr gravierenden, adäquat abdecken kann. Dies ist z.B. durch Ansätze des adaptiven Testens, bei denen den Befragten in Abhängigkeit vom bisherigen Antwortverhalten unterschiedliche Fragen vorgelegt werden (siehe Abschnitt 6), zu leisten.

Die Idee, die Kategorien der ICF für die Formulierung der Items eines Patientenfragebogens zu nutzen, liegt nahe, da die ICF eine Nähe zur Alltagssprache besitzt (insbesondere auf den Komponenten Aktivitäten und Partizipation) und damit die Verbreitung einer für den Patienten verständlichen Sprache unterstützen kann. Zudem berücksichtigt die Erarbeitung einer ICF-orientierten Patienteneinschätzung des Gesundheitsstatus den Kritikpunkt, dass die ICF die Patientenperspektive und die Rehabilitandenselbsteinschätzung zu wenig berücksichtige (Wade & Halligan, 2003, S. 351).

Der Nachteil der Neuentwicklung von ICF-orientierten Assessmentinstrumenten besteht darin, dass etablierte Verfahren und die damit verknüpften Erfahrungen und Datengrundlagen nicht berücksichtigt werden. Eine Milderung dieses Problems wäre denkbar durch den Ansatz des „Test-Equating“ (Kolen & Brennan, 2004), welcher es auf der Basis von Item-Response-Theorien (vgl. folgender Abschnitt 4) ermöglicht, die Scores verschiedener Instrumente auf einer gemeinsamen Metrik anzuordnen. Dabei muss jedoch vorausgesetzt werden, dass die zu verknüpfenden Instrumente das gleiche Merkmal erfassen.

## 4. Die Verwendung von Modellen der Item-Response-Theorie

Bei der Entwicklung neuer Assessmentverfahren stellt die Anwendung der so genannten „Item-Response-Theorie“ eine Alternative zur klassischen Testtheorie (vgl. z.B. DeVellis, 2006) dar. Nahezu alle Übersichtsarbeiten, Editorials und Positionspapiere zur Ergebnismessung in der Rehabilitation, die in den letzten fünf Jahren erschienen sind (z.B. Jette & Haley, 2005; Heinemann, 2005; Dekker et al., 2005; Ware, 2003; Wade, 2004), betonen die Relevanz der Modelle der Item-Response-Theorie (kurz: IRT-Modelle) und fordern eine stärkere Verbreitung.

IRT-Modelle spezifizieren den Zusammenhang zwischen einer latenten Variable (z.B. das zu messende Persönlichkeitsmerkmal Depressivität) und dem Verhalten bzw. der Verhaltenswahrscheinlichkeit bei der Bearbeitung eines Tests (z.B. Wahrscheinlichkeit des Ankreuzens der Kategorie „trifft vollkommen zu“ bei der Aussage „Ich fühle mich häufig niedergeschlagen“). Die Attraktivität von IRT-Modellen beruht auf ihren spezifischen Eigenschaften (vgl. Hays et al., 2000; Bond & Fox, 2001, und Embretson & Reise, 2000, für

entsprechende Einführungen). Dies soll beispielhaft für das wohl am häufigsten eingesetzte IRT-Modell, das Rasch-Modell (Rasch, 1980; Fischer & Molenaar, 1995), erläutert werden. Das Rasch-Modell besitzt unter anderem die Eigenschaften der Stichprobenunabhängigkeit (spezifische Objektivität), die Intervallskaleneigenschaft und die Eigenschaft, dass sowohl Personenparameter (also die Ausprägungen latenter Personenmerkmale) als auch Itemparameter (also z.B. die Itemschwierigkeit) auf einer Differenzskala liegen. Diese Eigenschaften setzen natürlich die Passung der Daten zum Rasch-Modell voraus, was mit entsprechenden Modellfit-Indices zu belegen ist.

Die Stichprobenunabhängigkeit besagt, dass die Differenzen zweier Personenparameter unabhängig davon sind, ob leichte oder schwere Items (aus einem einmal kalibrierten Item-Pool) eingesetzt werden und auch unabhängig von den Fähigkeiten anderer untersuchter Personen aus dem untersuchten Personen-Pool. Damit ist ein fundamentaler Aspekt von Messung erfüllt, der die Eigenschaften physikalischer Messungen abbildet und in der klassischen Testtheorie (wo Personenfähigkeiten über die relative Lösungshäufigkeit operationalisiert werden und damit stichprobenabhängig sind) nicht erfüllt ist.

Weitere Vorteile, die teilweise aus den beschriebenen Eigenschaften ableitbar sind (vgl. Bond & Fox, 2001; Hays et al., 2000), bestehen darin, dass

- Item- und Personenparameter unabhängig voneinander bestimmt werden können (im Gegensatz zur klassischen Testtheorie, bei der die Personenfähigkeit von der Schwierigkeit der Items abhängt),
- mit Hilfe des Konzepts des „Differential Item Functioning“ Gruppenunterschiede bezüglich der Itemparameter analysiert werden können,
- Skalen aus Items mit unterschiedlichen Antwortkategorien gebildet werden können,

- die Möglichkeit der Anwendung des adaptiven Testens besteht,
- neben dem Item-Fit auch ein Personen-Fit (Passung der Antworten einer Person zum Modell) bestimmt werden kann, mit dem z.B. unachtsam ausfüllende Personen oder das Auftreten kognitiver Fehler identifiziert werden können.

## 5. Die Berücksichtigung der Variabilität von Itembedeutung und Itemschwierigkeit

Das Ideal der exakten Messung besteht darin, dass das Messinstrument unabhängig von Kontextbedingungen stets das gleiche Konstrukt misst (Validität) und dies mit hoher Messgenauigkeit (Reliabilität). Entsprechende Anforderungen scheinen bezüglich physikalischer Messungen auch im Alltag selbstverständlich. Ein Barometer, welches je nach dem Geschlecht dessen, der es aufgehängt hat, etwas anderes anzeigt, oder ein Thermometer, welches nach einem Regenschauer etwas Grundsätzlich anderes misst als vor dem Regen, würde mit deutlicher Skepsis betrachtet werden.

Überträgt man diese Erwartungen ohne kritische Reflexion auf die Situation von Patienten-Selbstberichten in der Ergebnismessung, so kommt man leicht zu falschen Schlussfolgerungen. Die Items eines Fragebogens können z.B. durchaus für Männer eine andere Schwierigkeit aufweisen als für Frauen, und es ist möglich, dass Rehabilitanden mit einem Item vor der Rehabilitationsmaßnahme etwas anderes verbinden als nach der Rehabilitationsmaßnahme, z.B. weil sich kognitive Konzepte oder interne Bezugsmaßstäbe verändert haben. Bei der Ergebnisbewertung auf der Basis von Patienten-Selbstberichten müssen die Komplexität und die Dynamik von Gesundheitswahrnehmungen und Lebensqualitätsbeurteilungen berücksichtigt werden (vgl. Allison et al., 1997).

Die Konsequenz, die sich daraus für die Ergebnismessung in der Rehabilitation ergibt, besteht darin, die Möglichkeit der Variabilität von Itembedeutung und Itemschwierigkeit in Betracht zu ziehen und – wenn möglich – empirisch zu untersuchen. Dieses Thema soll im Folgenden vertieft werden, indem zwei zentrale, in der aktuellen Literatur häufig angesprochene Aspekte erörtert werden: die Phänomene des „Differential Item Functioning“ und des „Response Shifts“.

### 5.1 Differential Item Functioning

Differential Item Functioning (kurz: DIF) bezeichnet das Phänomen, dass Items eines Fragebogens bei verschiedenen Personengruppen (z.B. Frauen und Männer) unterschiedliche Eigenschaften (z.B. Itemschwierigkeiten) aufweisen, auch wenn nach der Ausprägung des Merkmals, welches mit dem jeweiligen Item erfasst werden soll (z.B. Mobilität), adjustiert wird (vgl. Holland & Wainer, 1993; Dorans & Kulick, 2006). Unterschieden werden eine gleichförmige DIF (englisch: uniform DIF) und eine ungleichförmige DIF (englisch: nonuniform DIF). Bei der gleichförmigen DIF ist der Unterschied der Itemeigenschaften (z.B. Itemschwierigkeiten) für alle Ausprägungen des jeweiligen Merkmals konstant, bei der ungleichförmigen DIF hängt der Unterschied von der Ausprägung des Merkmals ab.

Ein gleichförmiges DIF ist z.B. gegeben, wenn das Item „Wie schwer fällt es Ihnen, Geschirr im Spülbecken zu spülen?“, welches Aktivitäten im häuslichen Leben erfassen soll, für Männer (bei gleichem Niveau der Einschränkungen der Aktivitäten im häuslichen Leben, erfasst z.B. über Anker-Items, die kein DIF aufweisen) eine höhere Itemschwierigkeit hat als für Frauen, und zwar unabhängig vom Niveau der Aktivitätseinschränkung. Gilt die unterschiedliche Itemschwierigkeit z.B. nur für schwer Beeinträchtigte oder kehrt sich die Abfolge der Itemschwierigkeit unter Umständen für gering Beeinträchtigte sogar um, so handelt es sich um ein Phänomen des ungleichförmigen DIF.

## 5.2 Response Shift

Während sich das Phänomen des Differential Item Functioning auf die interpersonelle Variabilität von Itemeigenschaften bezieht, erfasst der Response Shift Aspekte der intrapersonellen Variabilität von Itembedeutungen. Ausgangspunkt der Untersuchung von Response Shift waren die in gewisser Weise als paradox bewerteten Befunde,

- dass Patienten mit chronischen Erkrankungen trotz entsprechender Einschränkungen ihre gesundheitsbezogene Lebensqualität (kurz: HRQOL) oft ähnlich gut einschätzen wie Gesunde,
- dass Patienten ihre gesundheitsbezogene Lebensqualität gelegentlich besser bewerten als Behandler,
- dass zum Teil deutliche Diskrepanzen auftreten zwischen Patienten-Selbstberichten und objektiven Indikatoren der Gesundheit (Norman, 2003).

Response Shift bezeichnet das Phänomen, dass Personen, die eine Veränderung ihres Gesundheitsstatus (z.B. bedingt durch eine Intervention) erleben, unter Umständen ihre internen Beurteilungsstandards, ihre Bewertungen von Gesundheitszuständen und ihre Konzeptualisierung von gesundheitsbezogener Lebensqualität verändern (Schwartz & Sprangers, 1999a). Diese Definition spricht drei Formen von Response Shift an, die kurz erläutert werden sollen (vgl. Schwartz & Sprangers, 1999b):

- Rekalibrierung: Veränderung der internen Standards der Messung (z.B. wenn „mäßige Schwierigkeiten“ bei der Frage nach Problemen beim Treppensteigen nach einer Rehabilitation etwas anderes bedeutet als vor der Intervention)
- Repriorisierung: Veränderung in der Bewertung von Gesundheitskomponenten (z.B. wenn soziale Beziehungen nach einer psychotherapeutischen Intervention für die Lebensqualität wichtiger werden als vorher)
- Rekonzeptualisierung: Grundsätzliche Veränderung

des Verständnisses von Gesundheit (z.B. wenn nach einer Beinlähmung das Gehvermögen für die Bewertung der gesundheitlichen Situation irrelevant wird und statt dessen wichtig wird, welche Form der Mobilität mit dem Rollstuhl besteht).

Die drei Formen sind nicht unabhängig voneinander zu sehen, da sich z.B. im Falle einer Rekalibrierung die Intervalle der Skala verändern und somit Ankerpunkte eine neue Bedeutung (Rekonzeptualisierung) erhalten (vgl. Schwartz & Sprangers, 1999b).

Die Relevanz des Response Shifts für die Ergebnismessung besteht darin, dass das Phänomen erhebliche Konsequenzen für die Bewertung von patientenseitig festgestellten Veränderungen hat, da ohne nähere Analysen zur Quantifizierung des Einflusses des Response Shifts unklar ist, ob eine Veränderung eine „wahre“ Veränderung des Gesundheitszustands darstellt oder eine Veränderung der Bewertungsmaßstäbe des Patienten. Dies gilt insbesondere für Patientenurteile mit einer stark subjektiven und aggregierenden Komponente (z.B. allgemeine Zufriedenheit mit dem eigenen Gesundheitszustand). In der medizinischen Rehabilitation und bei der Behandlung chronischer Erkrankungen spielt das Phänomen eine besondere Rolle (vgl. Schwartz et al., 2007), da sich die Ergebnismessung in diesem Bereich wesentlich auf Patientenangaben stützt. Eine Analyse des Phänomens kann deshalb sowohl die Validität der Messung von Effekten der Rehabilitation verbessern als auch das Verständnis der Selbstwahrnehmung von Gesundheit bei chronisch Kranken vertiefen helfen (Schwartz & Sprangers, 1999a).

Response Shift ist nicht notwendigerweise ein unerwünschter Artefakt. Das mit diesem Begriff bezeichnete Phänomen kann Ausdruck eines adaptiven Coping-Verhaltens sein und insofern unter Umständen das Ziel einer psychosozialen Intervention darstellen (vgl. z.B. Sharpe et al., 2005).

## 6. Adaptives Testen

Das konventionelle Vorgehen beim Testen, d.h. z.B. bei der Vorlage von Fragebogen-Items, die die Einschränkungen von Aktivitäten erfassen sollen, besteht darin, dass jeder Person dieselben Fragen in derselben Reihenfolge vorgelegt werden. Die Idee des adaptiven Testens besteht hingegen darin, jeder Person antwortabhängig, also adaptiv, diejenigen Items vorzulegen, die in hohem Maße bei dieser Person Rückschlüsse auf die zu messende Eigenschaft ermöglichen (Wainer, 2000a; Kubinger, 2003). Es werden zwei Varianten des adaptiven Testens unterschieden (vgl. Kubinger, 2003): das „Tailored-testing“, bei dem eine Person „maßgeschneidert“ die für sie passenden Items erhält und das „Branched-testing“, bei dem Items zu Gruppen zusammengefasst werden, so dass sich die adaptive Itemvorgabe darauf beschränkt, nach der Vorgabe einer Itemgruppe antwortabhängig die nächste Itemgruppe vorzulegen. Das „Branched-testing“ ist bezüglich der Anpassung an das individuelle Merkmalsniveau des Befragten im Vergleich zum „Tailored-testing“ suboptimal, es hat aber den Vorteil, dass es auch bei einer Papier-und-Bleistift-Vorgabe anwendbar ist (z.B. mit graphisch dargestellten Sprungbefehlen). Das „Tailored-testing“ erfordert bei einer Ausschöpfung aller Möglichkeiten eine computergestützte Vorgabe der Items und eine mit dem Computer individuell, auf der Basis eines Algorithmus durchgeführte Bestimmung der Auswahl der vorzulegenden Fragen.

Während das adaptive Testen im Bereich der Leistungsdiagnostik auf eine über 30-jährige Geschichte zurückblicken kann (vgl. Meijer & Nering, 1999), sind Anwendungen im Bereich klinischer Tests und in der medizinischen Rehabilitation erst seit einigen Jahren zu verzeichnen. Einhellig sehen jedoch nahezu alle neueren Übersichtsarbeiten zu Entwicklungen der Ergebnismessung in der Rehabilitation (z.B. Hays et al., 2000; Ware, 2003; Jette & Haley, 2005; Heinemann, 2005;

Ware et al., 2005) das adaptive Testen, insbesondere in der computergestützten Form, als einen wichtigen Zukunftstrend an, der mit wesentlichen Vorteilen verbunden ist.

Bezüglich des computergestützten adaptiven Testens (kurz: CAT) im Bereich klinischer Tests liegen auch aktuell im deutschsprachigen Raum nur sehr wenige Studien vor. Zu erwähnen sind die Arbeiten von Walter et al. (2005) zur Entwicklung eines computeradaptiven Tests von Angst, von Fliege et al. (2005) zur Entwicklung eines computeradaptiven Tests für Depressivität sowie ein aktuelles Projekt von Wirtz, Bengel & Gauggel („Entwicklung und Validierung eines adaptiven Assessments für Patienten in der orthopädischen Rehabilitation – RehaCAT“). Der Autor hat einen adaptiven Test zur Erfassung von Mobilität und Selbstversorgung vorgelegt, der aber bisher nur im Sinne des „Branched-testing“ in einer Papier-und-Bleistift-Version (Patienten- und Arzt-Form) vorliegt (Farin et al., 2006b, 2007b; Farin & Fleitz, 2009). Mit der Erstellung einer deutschsprachigen Version der in den USA (u.a. für das adaptive Testen) entwickelten PROMIS-Itembanken (Cella et al., 2007) befasst sich ein anderes aktuelles Projekt des Autors (siehe [www.aqms.de](http://www.aqms.de) unter „Laufende Projekte“).

International arbeiten mehrere Arbeitsgruppen an klinischen CAT-Verfahren, die in der Rehabilitation einsetzbar sind. Hervorzuheben sind insbesondere die Gruppe um Stephen Haley (Haley et al., 2006a, 2006b, 2004a, 2004b), die sich mit einem ICF-orientierten Verfahren zur Erfassung von Aktivitätseinschränkungen (AM-PAC) befasst, die Arbeitsgruppe um John Ware und Jakob Bjorner (Bjorner et al., 2003b, 2003c; Ware et al., 2003), die ein Verfahren zur Messung der Belastungen durch Kopfschmerzen entwickelt hat und die Gruppe um Dennis Hart (Hart et al., 2005, 2006), die sich mit der Messung des funktionalen Status bei Hüft-, Knie-, Fußbehinderungen bzw. bei Rückenproblemen beschäftigt.

## 7. Empfehlungen für die Praxis der Ergebnismessung in der medizinischen Rehabilitation

Im Folgenden sollen aus dem Vorgegangenen Empfehlungen für die Ergebnismessung in der medizinischen Rehabilitation abgeleitet werden. Diese Empfehlungen fokussieren – entsprechend der Zielsetzung der vorliegenden Arbeit – auf Fragebogenerhebungen:

1. Die Erhebung eines patientenseitigen Präferenz-, Ziel- oder Erwartungsparameters zusätzlich zu reinen Ergebnisvariablen scheint notwendig, um die Patientenmeinung bei der Gewichtung der Ergebnisse adäquat berücksichtigen zu können. Ein entsprechender Parameter sollte insbesondere bei Ergebnismessungen, die der Erfassung von Einschränkungen der Aktivitäten und Partizipation betreffen, mitgeführt werden.
2. Wenn Assessmentinstrumente neu entwickelt oder überarbeitet werden, sollten Möglichkeiten der Beteiligung von Patienten geprüft und ggf. genutzt werden. Prüfungen der Akzeptanz und Verständlichkeit scheinen auf jeden Fall angezeigt, ggf. auch die inhaltliche Bewertung der Relevanz der Frageninhalte.
3. Zusätzlich zur Erhebung der statistischen Signifikanz einer Veränderung nach einer rehabilitativen Intervention empfiehlt es sich, die Signifikanz aus Patientensicht zu erheben.
4. Eine Erfassung von Merkmalen der Patient-Behandler-Interaktion und der subjektiven Konzepte der Behandler kann die Aussagekraft der Ergebnisdaten erhöhen, da damit vermutlich relevante Moderatorvariablen erfasst werden.
5. Um die erhobenen Ergebnisdaten auf die ICF beziehen zu können und somit eine bessere Grundlage

für die Vergleichbarkeit und Integration der Daten zu schaffen, sollten bevorzugt Instrumente eingesetzt werden, für die der ICF-Bezug systematisch analysiert wurde bzw. die unmittelbar an der ICF orientiert neu entwickelt wurden.

6. Bei der Neuentwicklung oder Überarbeitung von Assessmentinstrumenten, aber auch bei der klinischen Diagnostik sollten (zumindest zusätzlich zu klassisch-testtheoretischen Verfahren) die Vorteile von Modellen der Item-Response-Theorie (z.B. Rasch-Modell) genutzt werden.
7. Das Differential Item Functioning (DIF) stellt mittlerweile ein Standard-Gütekriterium eines Assessmentinstruments dar. Bei Neuentwicklungen und Überarbeitungen sollte das DIF bestimmt werden, um die Grenzen des generischen Einsatzes des Verfahrens bestimmen zu können.
8. Die Untersuchung des Auftretens eines Response Shifts (z.B. mit dem „Then-Test“ oder Strukturgleichungsmodellen) kann die Aussagekraft der Ergebnisdaten erhöhen und sollte durchgeführt werden, wenn Fragen mit einer deutlichen subjektiven Bewertungskomponente vorgelegt werden.
9. Bei der Neuentwicklung von Assessmentinstrumenten sollten die Möglichkeiten des adaptiven Testens genutzt werden. Wenn die Zielpopulation und das Erhebungssetting es zulassen, empfiehlt sich eine computergestützte Realisierung des adaptiven Testens.

## 8. Forschungsbedarf

Ein Forschungsprogramm zur besseren methodischen Fundierung von Ergebnismessung in der Rehabilitation

und der besseren Interpretierbarkeit von Ergebnisdaten ließe sich durch folgende Aufgabenfelder kennzeichnen, die im Folgenden näher erörtert werden:

- die Verbesserung der methodischen Grundlagen patientenorientierter Messungen,
- die Verbesserung der Zuordnung der Resultate von Ergebnismessungen zu den ICF-Kategorien,
- die Verbesserung der Anwendbarkeit von Item-Response-Modellen und (computergestützten) adaptiven Verfahren auf Probleme der Ergebnismessung in der Rehabilitation,
- die Entwicklung von Strategien zur integrierten Betrachtung und Bewertung von a) direkten und indirekten Veränderungsurteilen sowie von b) patientenseitigen und behandlerseitigen Veränderungsurteilen bei Berücksichtigung der zugrunde liegenden psychologischen Prozesse der Urteilsbildung,
- die Erforschung des Einflusses der Subjektivität des Behandlers und der Patient-Behandler-Beziehung auf die Resultate von Ergebnismessungen.

Patientenorientierte Ergebnismessungen (vgl. Farin & Kriesch, 2008) basieren wesentlich auf der Bestimmung der Präferenzen bzw. Ziele respektive Erwartungen des Patienten im Hinblick auf die Behandlung und den angestrebten Gesundheitszustand. Es mangelt bisher an Arbeiten, die diese verschiedenen Konzepte zur Abbildung der Patientenbedürfnisse systematisch in Beziehung setzen, zum Beispiel durch einen empirischen Vergleich erhobener Präferenzen, Ziele und Erwartungen. Zudem ist die Frage zu beantworten, ob sich der hier postulierte Vorteil des Präferenz-Konstrukts gegenüber den Konzepten „Rehabilitationsziel“ und „rehabilitationsbezogene Erwartung“ empirisch bestätigen lässt. Bezüglich der Patientenpräferenzen lauten zentrale Forschungsfragen: Wie stabil, konsistent und widerspruchsfrei sind die von Patienten formulierten Präferenzen? Ist eine auf die Präferenzen des Patienten eingehende Behandlung (z.B. bei Kongruenz

patientenseitiger Präferenzen und artzseitig bestimmter Reha-Ziele oder bei expliziter Erfassung und Erörterung der Patientenpräferenzen) wirklich effektiver als eine Behandlung, die nicht systematisch auf die Präferenzen des Patienten ausgerichtet ist? Und wenn ja, auf welchen Erfolgsdimensionen?

Bezüglich der ICF-Orientierung in der Ergebnismessung lautet die zentrale Frage, ob der Weg der Neuentwicklung von Assessmentverfahren zielführender ist als der der Zuordnung bestehender Verfahren zu ICF-Kategorien. Unter Umständen ist diese Frage nicht generell zu beantworten, sondern nur in Abhängigkeit vom jeweiligen Anwendungsfeld. Neuentwicklungen werden sich nur durchsetzen können, wenn sie gleichzeitig durch die Verwendung moderner testtheoretischer Ansätze und die Berücksichtigung adaptiver Verfahrenselemente weitere Vorteile gegenüber klassischen Verfahren aufweisen. Die methodische Prüfung der wenigen, bereits vorliegenden Ansätze von ICF-orientierten Neuentwicklungen sollte in vertiefenden Studien erweitert werden; zu neuen, bisher nicht oder nur teilweise abgedeckten ICF-Bereichen (z.B. der Partizipation) sollten weitere Verfahrensentwicklungen erfolgen. Eine unmittelbare Orientierung an der ICF scheint dabei vorteilhafter als eine Orientierung an den ICF-Core-Sets (Stucki & Grimby, 2004), da mit der Orientierung an den ICF-Kategorien eine generische Geltung sichergestellt wird, während ein Bezug auf die ICF-Core-Sets wegen deren erkrankungsspezifischer Konstruktion (die in aktuellen Arbeiten, vgl. MyIntyre & Tempest, 2007, auch kritisch diskutiert wird) nicht mit diesem Vorteil verbunden wäre. Ein generisches, ICF-orientiertes Instrument kann die Passung zu verschiedenen Niveaus der Einschränkungen von Aktivitäten und Teilhabe durch ein adaptives Verfahrensprinzip sicherstellen.

Die Methodologie von Item-Response-Modellen ist im Hinblick auf die Ergebnismessung in der medizinischen

Rehabilitation hinreichend ausgearbeitet, so dass Forschungs- und Entwicklungsbedarf hier eher bezüglich Fragen der Anwendung und Verbreitung besteht (vgl. auch Lohr, 2000). Trotz des zunehmenden Einsatzes auch in der deutschsprachigen Rehabilitationswissenschaften (vgl. z.B. Wirtz & Böcker, 2007; Farin et al., 2006b, Farin, im Druck) sind die grundlegenden Konzepte der Item-Response-Modelle vielen Praktikern, die Ergebnismessungen durchführen, noch nicht hinreichend bekannt, so dass sich die Frage stellt, wie der hinderliche Einfluss der oft sehr technischen Ausdrucksweise dieser Ansätze beseitigt werden kann.

Ein wesentlicher Vorteil der Anwendung von Modellen der Item-Response-Theorie besteht in der Möglichkeit des adaptiven Testens. Eine vollständige Ausschöpfung der diesbezüglichen Optionen ist jedoch nur bei einer computergestützten Vorgabe möglich. Da der Einsatz des computergestützten adaptive Testen (kurz: CAT) in der Rehabilitation noch relativ neu ist, stellt sich eine Vielzahl von Fragen, die zum Gegenstand von entsprechender Studien gemacht werden sollten: Welche Benutzeroberfläche ist für die Akzeptanz von Patienten, die in der Rehabilitation oft recht alt und im Umgang mit Computern wenig geübt sind, adäquat (vgl. Zenisky & Sireci, 2002)? Welchen Einfluss hat eine negative Einstellung gegenüber Computern auf das Messergebnis? Welches ist die optimale Länge des dynamischen Testens, d.h. wie viele Items sollten unter Abwägung von Messgenauigkeit und Messökonomie vorgelegt werden (Ware et al., 2005)? Kann man davon ausgehen, dass die in einem CAT-System implementierten Parameter (Itemschwierigkeiten) hinreichend stabil sind, so dass das System über längere Zeit einsatzfähig ist (vgl. Wainer, 2000b)? Sind Personen- und Itemparameter-schätzungen hinreichend robust für die in CAT-Systemen auftretenden Variationen von Itemreihenfolgen und -positionen? Welche Bedeutung hat das Differential Item Functioning für die Generalisierbarkeit des Einsatzes von CAT-Systemen (Cook et al., 2005)? Können

die Datensätze in den wesentlichen Untergruppen frei sein vom Einfluss gruppenspezifischer Itemschwierigkeiten? Sind Computerversion und Papier-und-Bleistift-Version eines CAT-Verfahrens äquivalente Tests?

Zusätzlich zu den eigentlichen Forschungsfragen stellen sich die Probleme der Finanzierung der Erhebung der großen Datensätze, die für messgenaue Kalibrierungen von Items in einem CAT-System erforderlich sind, sowie des bisherigen Fehlens benutzerfreundlicher Software für die Entwicklung von computergestützten adaptiven Verfahren (Cook et al., 2005).

Für die Interpretierbarkeit von Forschungsarbeiten im Bereich der Ergebnismessung in der Rehabilitation ist es von entscheidender Bedeutung, die Unterschiede zwischen verschiedenen Veränderungsmaßen (direkte versus indirekte Messung, patienten- versus behandlerseitige Urteile) verstehen und erklären zu können, damit in einem zweiten Schritt Strategien zur Integration vorhandener Befunde erarbeitet werden können und die bestmögliche Methodik für zukünftige Studien geplant werden kann. Obwohl sich mehrere Arbeiten mit der Frage von Veränderungsurteilen in der Rehabilitation und bei Patienten mit chronischen Erkrankungen befasst haben (vgl. z.B. Kohlmann & Raspe, 1998; Farin et al 2007c), ist die Befundlage doch in vielen Fällen heterogen. Insbesondere für folgende Fragen konnte noch keine abschließende Antwort gefunden werden: Welche kognitiven Prozesse liegen den verschiedenen Urteilen (ggf. in Abhängigkeit von Fragemerkmale wie Formulierung, Zeitbezug, Ergebnisparameter etc.) zugrunde? Wie sind die Vor- und Nachteile direkter bzw. indirekter Veränderungsurteile in Abhängigkeit vom Studienziel, von der Zielpopulation und vom betrachteten Ergebnisparameter zu gewichten? Wie können direkte und indirekte Veränderungsurteile sinnvoll integriert werden? Welche Bedeutung kommt dabei dem Phänomen des Response Shifts zu? Welches sind die relevanten Auslöser und Moderatoren

des Response Shifts und wie kann man das Phänomen kognitionspsychologisch erklären (z.B. Erklärung der Response Shift – Theoretiker versus Erklärung mit Hilfe impliziter Theorien)? Welches sind die beim Response Shift wirkenden antezedenten Bedingungen und Moderatoren (vgl. z.B. Sprangers & Schwartz, 1999)?

Eine Beantwortung dieser Fragen wird nicht möglich sein ohne die Erforschung der kognitiven Prozesse, die der Urteilsbildung zugrunde liegen. Dafür bieten sich qualitative Verfahren und kognitive Tests an (vgl. Nieuwkerk et al., 2007), die in den Gesundheits- und Rehabilitationswissenschaften international in den letzten fünf bis zehn Jahren vermehrt durchgeführt und publiziert werden (z.B. Daly et al., 2007; Krause, 2006). Im deutschsprachigen Raum hat der Trend der zunehmenden Akzeptanz entsprechender Erhebungsmethoden jedoch bisher wenig Fuß gefasst. Methodisch gute qualitative Arbeiten, also z.B. Arbeiten, die durch kommunikative Validierung (Flick, 2006), Bestimmung der Interrater-Reliabilität der Kodierung und eine transparente Dokumentation des Vorgehens die Glaubwürdigkeit der Befunde sichern, können wesentlich zu einem Erkenntnisfortschritt bezüglich der Bewertung verschiedener Veränderungsurteile beitragen.

Dies gilt auch für die Frage nach der Bedeutung patienten- und behandlerseitiger Veränderungsurteile. Welche unterschiedlichen kognitiven Strategien wenden Patient und Behandler beim Versuch der Beantwortung einer Veränderungsfrage an? Wenn Patienten- und Behandler-Urteile auf grundsätzlich verschiedenen Beobachtungsperspektiven basieren, stellen sich die Fragen, welche spezifischen Merkmale diese Perspektiven kennzeichnen, wie mögliche Unterschiede zu erklären sind und wie die Perspektiven bei einer summarischen Erfolgsbeurteilung zu gewichten sind. Für welche Bereiche ist der Patient der Experte, für welche eher das Rehabilitationsteam und für welche unter Umständen auch (ggf. ergänzend) ein externer Beobachter aus

dem Alltagsleben, wie z.B. der Lebenspartner oder ein anderer nahestehender Angehöriger (vgl. z.B. Elliott et al., 2006)? Einen neuen Ansatz geht diesbezüglich die Arbeit von Farin (im Druck), im Rahmen derer Patienten- und Behandler-Items zu Mobilität und Selbstversorgung zu eindimensionalen Skalen integriert werden, die den Anforderungen des Rasch-Modells genügen. Das resultierende Instrument („MOSES-Combi“) ist für verschiedene Messmodelle (Patienten- und Behandlerdaten vorliegend, nur Patientendaten vorliegend, nur Behandlerdaten vorliegend) einsetzbar ist und liefert aufgrund der Eigenschaft des Rasch-Modells Messwerte, die jeweils auf einer Skala liegen und direkt vergleichbar sind. Damit sind gute Voraussetzungen für die Messung der Ergebnisse auch bei solchen Patienten gegeben, die selber nicht in der Lage sind, ein Urteil bezüglich des Erfolgs der Rehabilitation abzugeben.

## 9. Literatur

**Abma, T. A.** (2006): Patients as partners in a health research agenda setting: the feasibility of a participatory methodology. *Evaluation & The Health Professions*, 29, 424-439.

**Allison, P. J., Locker, D. & Feine, J. S.** (1997): Quality of life: a dynamic construct. *Social Science & Medicine*, 45, 221-230.

**Bjorner, J. B., Kosinski, M. & Ware, J. E.** (2003b): Calibration of an item pool for assessing the burden of headaches: An application of item response theory to the Headache Impact Test (HIT). *Quality of Life Research*, 12, 913-933.

**Bohrnstedt, G. W.** (1983): Measurement. In: P. H. Rossi, J. D. Wright & A. B. Anderson (Hrsg.). *Handbook of survey research* (S. 69-120). New York: Academic Press.

**Bond, T. G. & Fox, C. M.** (2001): Applying the Rasch model: Fundamental measurement in the human sciences. Mahwah, NJ: Lawrence Erlbaum Associates.

**Bührlen, B., Gerdes, N. & Jäckel, W. H.** (2005): Entwicklung und psychometrische Testung eines Patientenfragebogens für die medizinische Rehabilitation (IRES-3). *Die Rehabilitation*, 44, 63-74.

**Caron-Flinterman, J. F., Broerse, J. E. W. & Bunders, J. F. G.** (2005): The experiential knowledge of patients: a new resource for biomedical research? *Social Science & Medicine*, 60, 2575-2584.

**Cella, D., Yount, S., Rothrock, N., Gershon, R., Cook, K., Reeve, B., Ader, D., Fries, J. F., Bruce, B. & Rose, M.** (2007): The Patient-Reported Outcomes Measurement Information System (PROMIS): Promise of an NIH roadmap cooperative group during its first two years. *Medical Care*, 45, 3-11.

**Cieza, A., Geyh, S., Chatterji, S., Kostanjsek, N., Üstün, T. B. & Stucki, G.** (2005): ICF linking rules: an update based on lessons learned. *Journal of Rehabilitation Medicine*, 37, 212-218.

**Coenen, M., Cieza, A., Stamm, T. A., Amann, E., Kollerits, B. & Stucki, G.** (2006): Validation of the International Classification of Functioning, Disability and Health (ICF) Core Set for rheumatoid arthritis from the patient perspective using focus groups. *Arthritis Research & Therapy*, 8.

**Cook, K. F., O'Malley, K. J. & Roddey, T. S.** (2005): Dynamic assessment of health outcomes: Time to let the CAT out of the bag? *Health Services Research*, 40, 1694-1711.

**Daly, J., Willis, K. & Small, R.** (2007): A hierarchy of evidence for assessing qualitative health research. *Journal of Clinical Epidemiology*, 60, 43-49.

**Dekker, J., Dallmeijer, A. & Lankhorst, G.** (2005): Clinimetrics in rehabilitation medicine: current issues in developing and applying measurement instruments 1. *Journal of Rehabilitation Medicine*, 37, 193-201.

**Deutsches Institut für medizinische Dokumentation und Information (DIMDI)** (2005): Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (Stand: Oktober 2005).

**DeVellis, R. F.** (2006): Classical test theory. *Medical Care*, 44, 50-59.

**Dorans, N. J. & Kulick, E.** (2006): Differential item functioning on the Mini-Mental State Examination. An application of the Mantel-Haenszel and standardization procedures. *Medical Care*, 44, 107-114.

**Elliott, D., Lazarus, R. & Leeder, S.** (2006): Proxy respondents reliably assessed the quality of life of elective cardiac surgery patients. *Journal of Clinical Epidemiology*, 59, 153-159.

**Embretson, S. E. & Reise, S. P.** (2000): Item response theory for psychologists. Mahwah, NJ: Lawrence Erlbaum Association.

**Farin, E., Fleitz, A. & Follert, P.** (2006b): Entwicklung eines ICF-orientierten Patientenfragebogens zur Erfassung von Mobilität und Selbstversorgung. *Physikalische Medizin, Rehabilitationsmedizin, Kurortmedizin*, 16, 197-211.

**Farin, E., Fleitz, A. & Frey, C.** (2007b): Psychometric properties of an ICF-oriented, adaptive questionnaire for the assessment of mobility, self care and domestic life. *Journal of Rehabilitation Medicine*, 39, 537-546.

**Farin, E., Dudeck, A., Meffert, C., Glattacker, M., Jäckel, W. H., Beckmann, U. & Böwering, L.** (2007c): Qualitätssicherung in der ambulanten medizinischen

Rehabilitation: Konzeption und Ergebnisse eines Pilotprojekts zur Entwicklung eines Qualitätssicherungsprogramms für die Indikationsbereiche Muskuloskeletale und Kardiologische Erkrankungen. *Die Rehabilitation*, 46, 198-211.

**Farin, E.** (2008): Patientenorientierung und ICF-Bezug als Herausforderungen für die Ergebnismessung in der Rehabilitation. *Die Rehabilitation*, 47, 67-76.

**Farin, E. & Kriesch, M.** (2008): Patientenorientierte Veränderungsmessung. Akzeptanz und Verständlichkeit verschiedener Methoden zur Erfassung von Gesundheitsbewertungen. In: 7. Deutscher Kongress für Versorgungsforschung des Deutschen Netzwerks für Versorgungsforschung e.V., 16.-18.10.2008 in Köln.

**Farin, E. & Fleitz, A.** (2009): The development of an ICF-oriented, adaptive physician assessment instrument of mobility, self care, and domestic life. *International Journal of Rehabilitation Research*, 32 (2), 98-107.

**Farin, E.** (2009): Integration of patient and provider assessments of mobility and self-care results in unidimensional IRT-scales. *Journal of Clinical Epidemiology* online unter DOI: 10.1016/j.jclinepi.2008.11.014.

**Fischer, G. H. & Molenaar, I. W.** (1995): Rasch models: Foundations, recent developments, and applications. New York: Springer-Verlag.

**Flick, U.** (2006): Qualitative Evaluationsforschung. Reinbek: Rowohlt.

**Fliege, H., Becker, J., Walter, O. B., Bjorner, J. B., Klapp, B. F. & Rose, M.** (2005): Development of a computer-adaptive test for depression (D-CAT). *Quality of Life Research*, 14, 2277-2291.

**Gauthier, A. K. & Clancy, C. M.** (2004): Consumer-dri-

ven health care - Beyond rhetoric with research and experience (Editorial). *Health Services Research*, 39, 1049-1054.

**Goverover, Y., Kalmar, J. & Gaudino-Goering, E.** (2005): The relation between subjective and objective measures of everyday life activities in persons with multiple sclerosis. *Archives of Physical Medicine and Rehabilitation*, 86, 2303-2308.

**Grill, E., Mansmann, U., Cieza, A. & Stucki, G.** (2007): Assessing observer agreement when describing and classifying functioning with the International Classification of Functioning, Disability and Health. *Journal of Rehabilitation Medicine*, 39, 71-76.

**Güthlin, C.** (2004): Response Shift: Alte Probleme der Veränderungsmessung, neu angewendet auf gesundheitsbezogene Lebensqualität. *Zeitschrift für Medizinische Psychologie*, 13, 165-174.

**Haigh, R., Tennant, A., Biering-Sorensen, F., Grimby, G., Marincek, C., Phillips, S., Ring, H., Tesio, L. & Thonnard, J.-L.** (2001): The use of outcome measures in physical medicine and rehabilitation within Europe. *Journal of Rehabilitation Medicine*, 33, 273-278.

**Haley, S. M., Coster, W. J., Andres, P. L., Ludlow, L. H., Pengersh, N., Bond, T. L. Y., Sinclair, S. J. & Jette, A. M.** (2004a): Activity Outcome Measurement for Postacute Care. *Medical Care*, 42, I-49-I-61.

**Haley, S. M., Coster, W., Andres, P. L., Kosinski, M. & Pengsheng, N.** (2004b): Score comparability of short forms and computerized adaptive testing: Simulation study with the activity measure for post-acute care. *Archives of Physical Medicine and Rehabilitation*, 85, 661-666.

**Haley, S. M., Siebens, H., Coster, W. J., Tao, W., Black-**

**Schaffer, R. M., Gandek, B., Sinclair, S. J. & Pengsheng, N.** (2006a): Computerized adaptive testing for follow-up after discharge from inpatient rehabilitation: I. Activity outcomes. *Archives of Physical Medicine and Rehabilitation*, 87, 1033-1042.

**Haley, S. M., Ni, P., Hambleton, R. K., Slavin, M. D. & Jette, A. M.** (2006b): Computer adaptive testing improved accuracy and precision of scores over random item selection in a physical functioning item bank. *Journal of Clinical Epidemiology*, 59, 1174-1182.

**Hart, D. L., Mioduski, J. E. & Stratford, P. W.** (2005): Simulated computerized adaptive tests for measuring functional status were efficient with good discriminant validity in patients with hip, knee, or foot/ankle impairments. *Journal of Clinical Epidemiology*, 58, 629-639.

**Hart, D. L., Mioduski, J. E., Werneke, M. W. & Stratford, P. W.** (2006): Simulated computerized adaptive test for patients with lumbar spine impairments was efficient and produced valid measures of function. *Journal of Clinical Epidemiology*, 59, 947-956.

**Hays, R. D., Morales, L. S. & Reise, S. P.** (2000): Item Response Theory and Health Outcomes Measurement in the 21st Century. *Medical Care*, 38, II-28-II-42.

**Heinemann, A. W.** (2005): Putting Outcome Measurement in Context: A Rehabilitation Psychology Perspective. *Rehabilitation Psychology*, 50, 6-14.

**Holland, P. W. & Wainer, H.** (1993): *Differential item functioning*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

**Jette, A. M. & Haley, S. M.** (2005): Contemporary measurement techniques for rehabilitation outcomes assessment. *Journal of Rehabilitation Medicine*, 37, 339-345.

**Jette, A. M.** (2006): Toward a common language for func-

tion, disability, and health. *Physical Therapy*, 86, 726-734.

**Kohlmann, T. & Raspe, H.** (1998): Zur Messung patientennaher Erfolgskriterien in der medizinischen Rehabilitation: Wie gut stimmen „indirekte“ und „direkte“ Methoden der Veränderungsmessung überein? *Die Rehabilitation*, 37, Suppl. 1, 30-37.

**Kolen, M. J. & Brennan, R. L.** (2004): *Test Equating, Scaling, and Linking*. New York: Springer-Verlag.

**Krause, N.** (2006): The use of qualitative methods to improve quantitative measures of health-related constructs [Commentary]. *Medical Care*, 44, 34-38.

**Kubinger, K. D.** (2003): *Adaptives Testen*. In: K. D. Kubinger & R. S. Jäger (Hrsg.). *Schlüsselbegriffe der psychologischen Diagnostik*. Weinheim: Beltz PVU, 1-9.

**Leistner, K. & Beyer, H. M.** (2005): *Rehabilitation in der gesetzlichen Krankenversicherung (GKV). Antragsverfahren unter besonderer Berücksichtigung der ICF*. Landsberg/Lech: Ecomed.

**Lohr, K.** (2000): Health Outcomes Methodology Symposium: Summary and recommendations. *Medical Care*, 39 [9 (Suppl. II)]: II-194-II-208.

**Long, J. S. & Pavalko, E.** (2004): Comparing alternative measures of functional limitation. *Medical Care*, 42, 19-27.

**Mahoney, F. I. & Barthel, D. W.** (1965): Functional evaluation: The Barthel index. *Maryland State Medical Journal*, 14, 61-65.

**McIntyre, A. & Tempest, S.** (2000): Two steps forward, one step back? A commentary on the disease-specific core sets of the International Classification of Functioning, Disability and Health (ICF). *Disability & Rehabili-*

tation 2007, 29 (18), 1475-1479.

**Mead, N. & Bower, P.** (2000): Patient-centredness: A conceptual framework and review of the literature. *Social Science and Medicine*, 51, 1087-1110.

**Meijer, R. R. & Nering, M. L.** (1999): Computerized adaptive testing: Overview and introduction. *Applied Psychological Measurement*, 23, 187-194.

**Nieuwkerk, P. T., Tollenaar, M. S., Oort, F. J. & Sprangers, M. A. G.** (2007): Are retrospective measures of change in quality of life more valid than prospective measures? *Medical Care*, 45, 199-205.

**Norman, G.** (2003): Hi! How are you? Response shift, implicit theories and differing epistemologies. *Quality of Life Research*, 12, 239-249.

**Okochi, J., Utsonomiya, S. & Takahashi, T.** (2005): Health measurement using the ICF: Test-retest reliability study of ICF codes and qualifiers in geriatric care. *Health and Quality of Life Outcome*, 3, 46.

**Rasch, G.** (1980): Probabilistic models for some intelligence and attainment tests. Chicago: University of Chicago Press.

**Reed, G. W., Lux, J. B., Bufka, L. F., Trask, C., Peterson, D. B., Stark, S., Threats, T. T., Jacobson, J. W. & Hawley, J. A.** (2005): Operationalizing the International Classification of Functioning, Disability and Health in clinical settings. *Rehabilitation Psychology*, 50, 122-131.

**Sachverständigenrat für die Konzertierte Aktion im Gesundheitswesen** (2003): Finanzierung, Nutzerorientierung und Qualität (Gutachten 2003). Bonn.

**Scheibler, F.** (2005): 3. Tagung des Förderschwerpunktes „Der Patient als Partner im medizinischen Entschei-

dungsprozess“. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, 48, 796-802.

**Schwartz, C. E. & Sprangers, M. A. G.** (1999a): Introduction to symposium on the challenge of response shift in social science and medicine. *Social Science & Medicine*, 48, 1505-1506.

**Schwartz, C. E. & Sprangers, M. A. G.** (1999b): Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science & Medicine*, 48, 1531-1548.

**Schwartz, C. E., Andresen, E. M., Nosek, M. A. & Krahn, G. L.** (2007): Response shift theory: Important implications for measuring quality of life in people with disability. *Archives of Physical Medicine and Rehabilitation*, 88, 529-536.

**Seger, W., Schian, H.-M., Steinke, B., Heipertz, W. & Schuntermann, M.** (2004): Gesundheits-, sozial-, gesellschafts- und organisationspolitische Auswirkungen der Anwendung der ICF auf eine integrierte Rehabilitation: Vision der Umsetzung und ihrer Folgen. *Gesundheitswesen*, 66, 393-399.

**Sharpe, L., Butow, P., Smith, C., McConnell, D. & Clarke, S.** (2005): Changes in quality of life in patients with advanced cancer. Evidence of response shift and response restriction. *Journal of Psychosomatic Research*, 58, 497-504.

**Sprangers, M. A. G. & Schwartz, C. E.** (1999): Integrating response shift into health-related quality of life research: A theoretical model. *Social Science & Medicine*, 48, 1507-1515.

**Stryer, D. B., Siegel, J. E. & Rodgers, A. B.** (2004): Outcomes research. Priorities for an evolving field. *Medical Care*, 42, III-1-III-5.

**Stucki, G. & Grimby, G.** (2004): Foreword. Applying the

ICF in medicine. *Journal of Rehabilitation Medicine*, Suppl. 44, 5-6.

**Stucki, G., Cieza, A. & Melvin, J.** (2007): The international classification of functioning, disability and health: a unifying model for the conceptual description of the rehabilitation strategy. *Journal of Rehabilitation Medicine*, 39, 279-285.

**Verband Deutscher Rentenversicherungsträger.**

(2000a): Rechtliche und strukturelle Weiterentwicklung der Rehabilitation - Positionen der gesetzlichen Rentenversicherung und Empfehlungen an die Politik. *Die Rehabilitation*, 39, 1-7.

**Wade, D. T. & Halligan, P.** (2003): Editorial. New wine in old bottles: the WHO ICF as an explanatory model of human behaviour. *Clinical Rehabilitation*, 17, 349-354.

**Wade, D. T.** (2004): Assessment, measurement and data collection tools (Editorial). *Clinical Rehabilitation*, 18, 233-237.

**Wainer, H.** (2000a): Computerized adaptive testing. A primer. Mahwah, NJ: Lawrence Erlbaum.

**Wainer, H.** (2000b): Introduction and history. In: H. Wainer (Hrsg.), Computerized adaptive testing. A primer. Mahwah, NJ: Lawrence Erlbaum, 1-22.

**Walter, O. B., Becker, J., Fliege, H., Bjorner, J., Kosinski, M., Walter, M., Klapp, B. F. & Rose, M.** (2005): Entwicklungsschritte für einen computeradaptiven Test zur Erfassung von Angst (A-CAT1). *Diagnostica*, 51, 88-100.

**Ware, J. E.** (2003): Conceptualization and measurement of health-related quality of life: Comments on an evolving field. *Archives of Physical Medicine and Rehabilitation*, 84, Suppl. 2, 43-51.

**Ware, J. E., Kosinski, M., Bjorner, J. B., Bayliss, M. S., Batenhorst, A., Dahlöf, C. G. H., Tepper, S. & Dawson, A.** (2003): Applications of computerized adaptive testing (CAT) to the assessment of headache impact. *Quality of Life Research*, 12, 935-952.

**Ware, J. E., Gandek, B., Sinclair, S. J. & Bjorner, J.** (2005): Item response theory and computerized adaptive testing: Implications for outcome measurement in rehabilitation. *Rehabilitation Psychology*, 50, 71-78.

**Whiteneck, G. G. & Gerhart, K. A.** (2001): Role of survey research in assessing rehabilitation outcome. *American Journal of Physical Medicine and Rehabilitation*, 80, 305-309.

**Wirtz, M. & Böcker, M.** (2007): Eigenschaften und Nutzen des Rasch-Modells in der klinischen Diagnostik. *Die Rehabilitation*, 46, 238-245.

**Zenisky, A. L. & Sireci, S. G.** (2002): Technological innovations in large-scale assessment. *Applied Measurement in Education*, 15 (4), 337-362.

# Anforderungen an ein adäquates System zur Bewertung des Reha-Erfolgs – das Beispiel „Qualitätsbarometer“

**Dipl.-Soz. Wiss. Hans Dörning**

Institut für Qualitätsmanagement und Qualitätssicherung in der Medizinischen Rehabilitation e.V. (IQEM)

## Inhaltsverzeichnis

<b>1. Einführung</b>	<b>85</b>
<b>2. Anforderungen an eine adäquate Bewertung des Reha-Erfolgs</b>	<b>86</b>
<b>3. Literatur</b>	<b>98</b>

## 1. Einführung

Im Rahmen des Qualitätssicherungsprogramms („QS-Programm“) der Deutschen Rentenversicherung werden routinemäßig seit Mitte 1997 vergleichende Analysen zur Struktur-, Prozess- und Ergebnisqualität in rentenversicherungseigenen sowie federführend belegten Rehabilitationseinrichtungen durchgeführt.

Ein Modul des QS-Programms der Deutschen Rentenversicherung ist die standardisierte Rehabilitandenbefragung zur Behandlungszufriedenheit und zur Bewertung der Ergebnisqualität. Im Gegensatz z. B. zu den Qualitätssicherungsprogrammen der gesetzlichen Krankenversicherung und des Heilbäderverbandes basiert die Messung der Ergebnisqualität im QS-Programm der Deutschen Rentenversicherung auf einer retrospektiven Einpunktmessung. D. h. die Rehabilitanden werden einmalig ca. 2-3 Monate nach Beendigung der Rehabilitation zum Reha-Erfolg aus Patientensicht befragt.

Während in den QS-Programmen der gesetzlichen Krankenversicherung und des Heilbäderverbandes, aufgrund einer prospektiven Mehrpunktmessung, die in der Wissenschaft als „Goldstandard“ angesehene indirekte Methode der Veränderungsmessung eingesetzt werden kann, muss im Qualitätssicherungsprogramm der Deutschen Rentenversicherung, bedingt durch das gewählte Verfahren der retrospektiven Einpunktmessung, auf die Methode der direkten Veränderungsmessung zurückgegriffen werden.

Vor diesem Hintergrund befassen sich die folgenden Ausführungen zunächst damit, ob die Ergebnisse der direkten und indirekten Methode der Veränderungsmessung vergleichbar sind. Darüber hinaus werden am Beispiel des „Qualitätsbarometers“, eines vom IQEM entwickelten standardisierten Systems zur systemati-

schen Bewertung der Struktur-, Prozess- und Ergebnisqualität von Rehabilitationseinrichtungen (Bitzer et al. 2004, Dörning et al. 2004, Lorenz & Mackensen 2004), weitere relevante Anforderungen an ein adäquates System zur Bewertung der Ergebnisqualität diskutiert. Diese Diskussion bezieht sich auf die Anzahl notwendiger Erhebungszeitpunkte zur Messung der Ergebnisqualität, auf wesentliche Kriterien zur Gewährleistung eines „fairen“ Vergleichs der Ergebnisqualität unterschiedlicher Rehabilitationseinrichtungen sowie auf besonders zentrale Anforderungen an Instrumente zur Messung der Ergebnisqualität.

## 2. Anforderungen an eine adäquate Bewertung des Reha-Erfolgs

### 2.1 Vergleichbarkeit der indirekten und direkten Veränderungsmessung

Um die durch eine Intervention (in diesem Fall „medizinische Rehabilitation“) erzielten Auswirkungen auf im Zusammenhang mit der Grunderkrankung stehende Beeinträchtigungen, Beschwerden, Symptome, Schmerzen etc. zu ermitteln, wird in der wissenschaftlichen Forschung die prospektive Messung dieser Ergebnisparameter zu mindestens zwei Erhebungszeitpunkten (d. h. vor und nach der Intervention) als Standard angesehen. Diese Form der Veränderungsmessung wird auch als „indirekt“ bezeichnet, da die Probanden zum jeweiligen Erhebungszeitpunkt ausschließlich zu ihrem aktuellen Zustand befragt werden, und Veränderungen „indirekt“ aus dem Vergleich der Messwerte vor und nach der Intervention abgeleitet werden.

Dagegen beruht die „direkte“ Veränderungsmessung auf der einmaligen retrospektiven nach der Intervention durchgeführten Erhebung des Gesundheitszustandes im Vergleich zur Zeit vor der Intervention (Beispiel: „Wie stark haben sich Ihre Beschwerden im Vergleich

zu der Zeit vor der Rehabilitation verändert?"; Antwortkategorien z. B.: „verbessert“, „nicht verändert“, „verschlechtert“) und bildet somit die von den Probanden selbst „direkt“ wahrgenommenen Veränderungen ab.

Bei der indirekten Veränderungsmessung wird (implizit) davon ausgegangen, dass die Erhebung von Beschwerden, Symptomen etc. frei von systematischen Erinnerungsfehlern ist. Dagegen sind bei der direkten Veränderungsmessung Erinnerungsfehler nicht auszuschließen.

Die nachfolgend berichteten Ergebnisse zur Vergleichbarkeit der indirekten und direkten Veränderungsmessung basieren auf einer Studie zur Validität von unterschiedlichen Methoden der Veränderungsmessung am Beispiel der Prozeduren „Leistenbruchoperationen“ (n = 185) (Bitzer et al., 2008b) und „Cholecystektomie“ (Gallenblasenentfernung) (n = 130) in deutschen Akutkrankenhäusern (Bitzer et al., 2008a). Die Ergebnisse lassen sich jedoch problemlos auf die Rehabilitation übertragen.

Vor der Operation (T0) erhielten alle Patienten einer Indikation den gleichen Fragebogen, in dem u. a. die aktuell präoperativ bestehenden Beeinträchtigungen durch indikationsspezifische Beschwerdenlisten erhoben wurden (Beschwerdenliste-Leiste: 9 Items; Beschwerdenliste-Galle: 8 Items) sowie eine Globaleinschätzung zum subjektiven Schweregrad. Diese Angaben bildeten den Ausgangspunkt für die indirekte Veränderungsmessung (prä-Werte).

Zu insgesamt zwei postoperativen Erhebungszeitpunkten (T1: ca. 14 Tage nach Operation; T2: ca. 6 Monate nach Operation) umfassten die Fragebögen wiederum die indikationsspezifischen Beschwerdenlisten sowie die Globaleinschätzung zum subjektiven Schweregrad, mit denen die aktuell zum jeweiligen Erhebungszeitpunkt bestehenden Beeinträchtigungen erhoben wurden (post-Werte für die indirekte Veränderungsmessung). Zusätzlich enthielten die Fragebögen die jeweiligen Items zur

direkten Veränderungsmessung. Die vorgegebenen Antwortmöglichkeiten lauteten dabei „stark verschlechtert“, „leicht verschlechtert“, „nicht verändert“, „leicht gebessert“ und „stark gebessert“.

Zur Bewertung der Vergleichbarkeit der indirekten und direkten Veränderungsmessung wurden die Antwortmöglichkeiten (bei der direkten Methode) bzw. Prä-Post-Differenzen (bei der indirekten Methode) so vercodet bzw. verrechnet, dass sich ein identischer Wertebereich von jeweils -2 bis +2 für jedes Einzelitem ergibt (-2 = „stark gebessert“, -1 = „leicht gebessert“, 0 = nicht verändert“, +1 = „leicht verschlechtert“, +2 = „stark verschlechtert“).

Um Aussagen zur Vergleichbarkeit der Ergebnisse der indirekten und direkten Veränderungsmessung machen zu können, wurden zunächst die mit den beiden Methoden berechneten Veränderungen für die Einzelitems der indikationsspezifischen Beschwerdenlisten deskriptiv bestimmt sowie die mittleren Differenzen über alle Einzelitems. Darüber hinaus wurde, um Hinweise über die Stärke der Übereinstimmung zwischen den Ergebnissen der indirekten und direkten Veränderungsmessung zu erhalten, die Interrater-Reliabilität der beiden Schätzer „indirekte“ und „direkte“ Veränderungsmessung mit Hilfe des Kappa-Koeffizienten berechnet (Bitzer et al., 2003).

Die Ergebnisse zu den berechneten Veränderungen zwischen den Einzelitems der beiden Beschwerdenlisten sind eindeutig und verweisen darauf, dass die Veränderungen im Beeinträchtigungsniveau sowohl zwischen T0 (vor der Operation) und T1 (ca. 14 Tage nach Operation) als auch zwischen T0 und T2 (ca. 6 Monate nach Operation) bei beiden Indikationen in Abhängigkeit von der Methode der Veränderungsmessung durchgängig erheblich differieren.

Ausnahmslos bei allen Items der beiden Beschwerden-

listen liegen die mit Hilfe der direkten Messmethode berechneten Veränderungen deutlich über den Veränderungen, die auf Basis des „Goldstandards“, also der indirekten Messmethode, bestimmt wurden. So beträgt z.B. die mittlere Differenz der indirekten Veränderungsmessung bei Leistenbruchoperationen zwischen T0 und T1 +0,15, d. h. beobachtet wird im Mittel über alle Einzelitems ca. zwei Wochen nach Operation eine leichte Zunahme der Beschwerden. Nach der direkten Veränderungsmessung ergibt sich dagegen eine Abnahme der Beschwerden (-0,69) (bei identischen Wertebereichen zwischen -2 bis +2) (vgl. Tabelle 1).

**Tabelle 1: Indirekte und direkte Veränderungsmessung – Berechnete Veränderungen (Differenzen)**

Leiste	Δ T1-T0		Δ T2-T0	
	Indirekte Methode	Direkte Methode	Indirekte Methode	Direkte Methode
Item 1	-0,02	-0,72	-0,61	-1,33
Item 2	-0,48	-1,03	-1,32	-1,56
Item 3	0,14	-0,81	-0,55	-1,47
Item 4	0,34	-0,91	-0,65	-1,25
Item 5	0,39	-0,59	-0,60	-1,20
Item 6	0,16	-0,48	-0,87	-1,16
Item 7	0,47	-0,64	-0,34	-1,13
Item 8	0,37	-0,67	-0,29	-1,16
Item 9	-0,04	-0,38	-0,09	-0,84
MW*	0,15	-0,69	-0,59	-1,23
<b>Galle</b>				
Item 1	-0,66	-1,33	-1,01	-1,56
Item 2	-0,05	-0,58	-0,25	-0,80
Item 3	-0,33	-0,79	-0,35	-1,08
Item 4	0,13	-0,67	-0,22	-1,00
Item 5	0,53	-0,54	-0,23	-0,96
Item 6	0,72	-0,25	-0,30	-0,96
Item 7	-0,44	-0,63	-0,44	-1,04
Item 8	-0,18	-0,88	-0,38	-0,92
MW*	-0,03	-0,71	-0,40	-1,04

\* MW: Mittlere Differenz über alle Einzelitems

Auch in Bezug auf die gemessenen Veränderungen zwischen T0 und T2 (ca. 6 Monate nach Operation) lässt sich feststellen, dass die direkte Veränderungsmessung die beobachteten Veränderungen auf der Basis der indirekten Methode durchgängig überschätzt. Im Mittel über alle Einzelitems z. B. ergibt sich nach der indirekten Messmethode eine Beschwerdenminderung um -0,59, während nach der direkten Veränderungsmessung die durchschnittliche Beschwerdereduktion gut doppelt so stark ausgeprägt ist (-1,23) (vgl. Tabelle 1).

Die Resultate zur Interrater-Reliabilität der beiden Schätzer „indirekte“ und „direkte“ Messung belegen ebenfalls die mangelnde Übereinstimmung der Ergebnisse zu den Veränderungen im Beeinträchtigungsniveau im zeitlichen Verlauf. So liegt der gemessene Übereinstimmungswert zwischen indirekter und direkter Messmethode (Kappa-Koeffizient) im Zusammenhang mit Leistenbruchoperationen im Mittel aller Einzelitems bei 0,16 (Veränderungen zwischen T0 und T1) bzw. 0,14 (Veränderungen zwischen T0 und T2) und damit nur unwesentlich über der mathematischen Zufallsübereinstimmung von 0. Auch bei der Indikation „Cholelithiasis“ sind lediglich Kappa-Werte von 0,11 bzw. 0,15 feststellbar (vgl. Tabelle 2).

**Tabelle 2: Indirekte und direkte Veränderungsmessung – Berechnete Veränderungen (Differenzen)**

Leiste	Leistenbruchoperation		Gallenblasenentfernung	
	Δ T1-T0	Δ T2-T0	Δ T1-T0	Δ T2-T0
MW Einzelitems*	0,16	0,14	0,11	0,15
GB Beschwerden**	0,12	0,16	0,06	0,12

\* MW Einzelitems: Mittlerer Kappa-Wert über alle Einzelitems der Beschwerdeliste

\*\* GB Beschwerden: Globaleinschätzung der Beschwerden

Bestätigt wird die mangelnde Interrater-Reliabilität bei der Prüfung eines weiteren Outcome-Parameters, der Globaleinschätzung des Beschwerdeniveaus durch die

Probanden. Auch hier liegt der Kappa-Koeffizient zu den kurz- wie auch zu den mittelfristigen Veränderungen bei beiden Indikationen mit Werten zwischen 0,06 und 0,16 nur marginal über dem Zufallswert von 0 (vgl. Tabelle 2).

Insgesamt lässt sich somit festhalten, dass die Ergebnisse der indirekten und der direkten Methode der Veränderungsmessung nicht vergleichbar sind, d. h. sie erfassen vermutlich unterschiedliche Konstrukte. Aufgrund der retrospektiven Einpunktmessung der Ergebnisqualität auf Basis der direkten Veränderungsmessung liegen zudem keine Informationen über den Ausgangsstatus und damit über das Beeinträchtigungsniveau der Probanden zu Beginn der Intervention vor. Das bedeutet aber auch, dass ein äußerst relevanter Prädiktor für den Reha-Erfolg und somit auch ein wesentliches Merkmal für die Fallmix-Adjustierung entfällt (zur Fallmix-Adjustierung vgl. auch Kapitel 2.2).

Auch wenn eine Einpunktmessung deutliche ökonomische Vorteile aufweist, müssen angesichts des hohen Stellenwertes der vergleichenden Qualitätsbewertung, und der aus den Qualitätsergebnissen in Zukunft möglicherweise resultierenden Konsequenzen für die teilnehmenden Rehabilitationseinrichtungen, wie z. B. Nutzung der Ergebnisse für die Belegungssteuerung und/oder für leistungsbezogene Vergütungsverhandlungen, die Limitationen und Verzerrungsmöglichkeiten dieses Ansatzes bedenklich erscheinen.

## **2.2 Weitere Anforderungen an eine adäquate Bewertung des Reha-Erfolgs**

### **Anzahl an Erhebungszeitpunkten**

Wie bereits aus den Ausführungen in Kapitel 2.1 zu entnehmen ist, reicht aufgrund der mangelnden Interrater-Reliabilität der direkten Veränderungsmessung eine Einpunktbefragung zur adäquaten Bewertung des Reha-Erfolgs nicht aus. Minimal sollten zu zumindest zwei

Erhebungszeitpunkten die jeweils relevanten Outcome-Parameter erhoben werden, um Aussagen über Veränderungen in den Zielgrößen treffen zu können.

Sinnvollerweise sollte die erste Erhebung entweder kurz vor Beginn der Reha-Maßnahme oder aber direkt zu Beginn der Rehabilitation erfolgen. Das vom IQEM entwickelte „Qualitätsbarometer“ z.B. sieht als ersten Messzeitpunkt aus Gründen der Praktikabilität und auch der Ökonomie den Beginn des Reha-Aufenthaltes vor (Erhebung des Ausgangsniveaus vor Intervention). Wenn allerdings, wie bei der vergleichenden Bewertung der Ergebnisqualität im Rahmen des QS-Programms der Deutschen Rentenversicherung, ein „Benchmarking“ angestrebt wird und die Ergebnisse des Benchmarking-Verfahrens mit etwaigen Sanktionen verbunden sein könnten, indem sie beispielsweise die Grundlage für die Belegungssteuerung oder für zukünftige leistungsbezogene Vergütungsverhandlungen bilden würden, dann müsste abgewogen werden, ob eher Praktikabilitäts- und finanzielle Erwägungen oder aber bereits durch das Erhebungsverfahren prinzipiell ausgeschaltete Manipulationspotenziale als gravierender einzuschätzen sind.

Als zweiter Erhebungszeitpunkt kommt, um Aussagen über den Reha-Erfolg machen zu können, eine Messung der jeweils relevanten Outcome-Parameter frühestens am Ende des Reha-Aufenthaltes in Betracht. Im „Qualitätsbarometer“ z. B. liegt dieser zweite obligatorische Erhebungszeitpunkt 3 bis 4 Wochen nach Abschluss der Reha-Maßnahme, um dem möglichen Problem der „sozialen Erwünschtheit“ und damit der Gefahr positiv gefärbter Antworten zu entgehen. Bei einer vergleichenden Qualitätsbewertung sollte zudem, wie bereits beim ersten Erhebungszeitpunkt, die grundsätzliche Möglichkeit weiterer Beeinflussungspotenziale bei der Terminierung des zweiten Messzeitpunktes bedacht werden.

Auf der Basis der beiden genannten Messzeitpunkte ist es grundsätzlich möglich, fundierte Aussagen über den

Reha-Erfolg zu machen. Diese Aussagen können sich allerdings nur auf den kurzfristigen Erfolg beziehen und erlauben noch keine Rückschlüsse über die Nachhaltigkeit der während der Rehabilitation erzielten Erfolge. Aus diesem Grund ist beispielsweise bei Einsatz des „Qualitätsbarometers“ fakultativ ein dritter Erhebungszeitpunkt vorgesehen. Dieser dritte Messzeitpunkt liegt je nach Indikationsgebiet 5 bis 8 Monate nach Entlassung aus der Reha-Einrichtung und ist schwerpunktmäßig auf den mittelfristigen Reha-Erfolg ausgerichtet.

#### **Kriterien zur Gewährleistung eines „fairen“ Vergleichs**

Eine weitere wesentliche Anforderung an ein System zur Bewertung des Reha-Erfolgs ist bei vergleichenden Qualitätsanalysen die Gewährleistung eines „fairen“ Vergleichs und damit der adäquate Umgang mit einem möglicherweise unterschiedlichen Fallmix in den einzelnen Reha-Einrichtungen.

Dieser Anforderung wird beispielsweise im „Qualitätsbarometer“ durch die Verwendung angemessener und dem gegenwärtigen Stand der Wissenschaft entsprechender multivariater Verfahren der Risikoadjustierung und unter Berücksichtigung der unterschiedlichen Fallzusammensetzungen in den Einrichtungen entsprochen. Dabei wird simultan für verschiedene Störgrößen, sogenannte Confounder, kontrolliert, d. h. für Einflussgrößen, die das Reha-Ergebnis u. U. beeinflussen und die nicht oder nur unwesentlich von den Reha-Einrichtungen selbst beeinflusst werden können.

Im „Qualitätsbarometer“ werden beispielsweise in die Fallmix-Adjustierung u. a. verschiedene soziodemographische Merkmale (z. B. Alter, Geschlecht etc.), die Reha-Motivation zu Beginn der Rehabilitation, bestehende Risikofaktoren und Komorbiditäten sowie das Krankheitsmanagement (z. B. bisherige Reha-Aufenthalte, Antrag auf EU-/BU-Berentung) einbezogen. Darüber hinaus wird als weiterer, von der Rehabilitation nicht beeinflussbarer wesentlicher Prädiktor für die

Ergebnisqualität, das gesundheitliche Beeinträchtigungsniveau zu Beginn der Rehabilitation berücksichtigt.

Bisherige Erfahrungen verdeutlichen jedoch, dass trotz einer am derzeitigen Stand der Wissenschaft ausgerichteten Fallmix-Adjustierung keine generelle „faire“ Vergleichbarkeit über alle Reha-Indikationen hinweg möglich ist. Ein „fairer“ Vergleich erscheint nur auf der Ebene von Indikationsbereichen möglich.

#### **Instrumente zur Messung des Reha-Erfolgs**

Generell sollte bei der Auswahl von Instrumenten zur Messung des Reha-Erfolgs darauf geachtet werden, dass diese Instrumente bereits im rehabilitationswissenschaftlichen Kontext eingesetzt worden sind. Außerdem sollten sie hinsichtlich relevanter psychometrischer Gütekriterien geprüft sein und in einer validierten deutschsprachigen Version zur Verfügung stehen. Sie sollten zudem nicht zu zeitintensiv und relativ einfach von Rehabilitanden bearbeitbar sein sowie möglichst geringe Lizenzkosten verursachen.

Im Rahmen der Messung der Ergebnisqualität sollten sowohl indikations- bzw. indikationsbereichsspezifische Instrumente als auch ein mehrere Dimensionen des gesundheitlichen Befindens thematisierendes generisches (indikationsübergreifendes) Instrument, zu dem bevölkerungsbezogene Referenzdaten zur Verfügung stehen, zum Einsatz kommen.

Ein oder mehrere indikations- bzw. indikationsbereichsspezifische Instrumente sollten vor allem verwendet werden, weil sie in der Regel spezifischer auf die jeweilige Indikation bzw. den jeweiligen Indikationsbereich zugeschnitten sind, was zumeist auch mit einer höheren Änderungssensitivität verbunden ist. Das heißt, dass reale Änderungen im Beeinträchtigungsniveau mit diesen Instrumenten in der Regel auch exakter abbildbar sind.

So weist z. B. das im „Qualitätsbarometer“ im Indikationsbereich Orthopädie eingesetzte indikations-spezifische Instrument „Funktionsfragebogen Hannover“ (FFbH), das vor allem auf funktionelle Aspekte ausgerichtet ist, die u. a. aufgrund von Schmerzen eingeschränkt sein können, eine etwas höhere Änderungssensitivität auf (0,73) als die beiden am ehesten korrespondierenden Subskalen „Physische Mobilität“ und „Schmerz“ des generischen Nottingham Health Profile (NHP) (Rudolph et al., 2007).

**Tabelle 3: Änderungssensitivität NHP-Subskalen und FFbH**

Nottingham Health Profile	ES* (adjustiert)
- Energie	- 0,38
- Schmerz	- 0,54
- Emotionale Reaktion	- 0,32
- Schlaf	- 0,29
- Soziale Isolation	- 0,27
- Physische Mobilität	- 0,69
Funktionsfragebogen Hannover	0,73

\* ES: Effektstärke

Zusätzlich sollte durchgängig in allen Indikationsbereichen ein generisches Instrument, zu dem bevölkerungsbezogene Referenzdaten zur Verfügung stehen, eingesetzt werden, da mit Hilfe eines derartigen Instrumentes wertvolle Zusatzinformationen aus mehreren Perspektiven gewonnen werden können.

Bei allen Indikationsbereichen wird beispielsweise im „Qualitätsbarometer“ das Nottingham Health Profile (NHP) als generisches Erhebungsinstrument verwendet, für das bevölkerungsbezogene Referenzwerte verfügbar sind. Das NHP setzt sich aus sechs Skalen mit insgesamt 38 Items zusammen, die verschiedene Aspekte des psychischen Befindens, der körperlichen Verfassung, der sozialen Beziehungen und der funktionellen Kompetenz thematisieren.

Die Referenzdaten des Nottingham Health Profile ermöglichen es, die Einschränkungen im gesundheitlichen Befinden bzw. in der gesundheitsbezogenen Lebensqualität der Rehabilitanden zu Beginn der Rehabilitation im Vergleich zur bundesdeutschen Bevölkerung abzuschätzen. Darüber hinaus können die Dimensionen des gesundheitlichen Befindens identifiziert werden, in denen besonders ausgeprägte Einschränkungen vorhanden sind. Vor allem in diesen Dimensionen sind in der Regel positive Veränderungen durch die Rehabilitation zu erwarten. Umgekehrt gibt der Vergleich mit den Referenzdaten auch Aufschluss über die Dimensionen des Befindens, in denen keine oder lediglich marginale Einschränkungen bei den Rehabilitanden zu Beginn der Rehabilitation bestehen. In diesen Dimensionen sind auch keine oder allenfalls leichte Verbesserungen durch die Rehabilitation erwartbar.

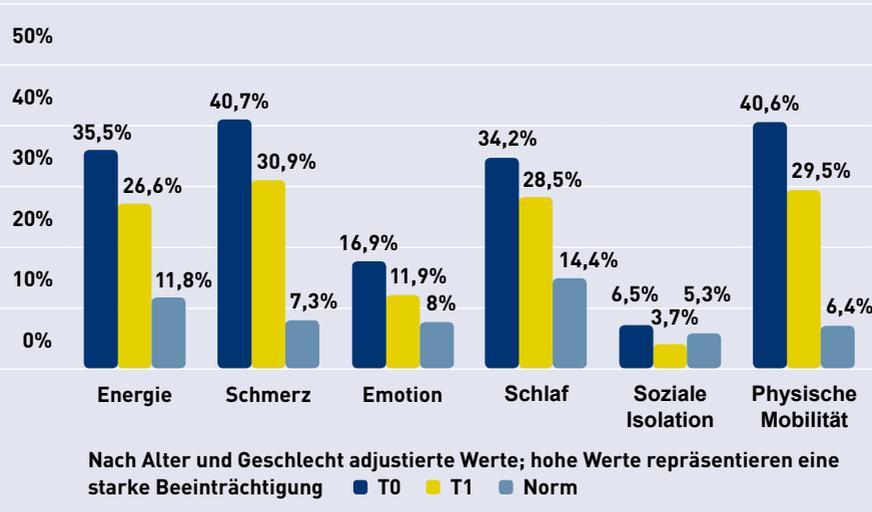
Ein Beispielergebnis, das auf gepoolten Daten aus mehreren Rehabilitationseinrichtungen beruht, soll die Vorgehensweise verdeutlichen: Patienten der orthopädischen Rehabilitation sind zu Beginn der Rehabilitation in nahezu sämtlichen Bereichen des gesundheitlichen Befindens im Vergleich zur Referenzbevölkerung erheblich eingeschränkt (ein Wert von 100% steht jeweils für die maximale Beeinträchtigung, ein Wert von 0% für keinerlei Einschränkung). Die stärksten Einschränkungen bestehen, sowohl absolut als auch relativ im Vergleich zur Referenzbevölkerung, in der „Physischen Mobilität“ und bei „Schmerzen“ (vgl. Abbildung 1).

Durch die Rehabilitation kommt es, mit Ausnahme der Dimension „Soziale Isolation“, in der bereits zu Beginn der Rehabilitation keine statistisch signifikanten Unterschiede zur Referenzbevölkerung bestehen, in allen übrigen Bereich des gesundheitlichen Befindens zu substantziellen und statistisch signifikanten Verbesserungen. Am stärksten sind diese Verbesserungen im Bereich der „Physischen Mobilität“ und bei „Schmerzen“ ausgeprägt: Einschränkungen der körperlichen Mobilität neh-

men um 11,1 Prozentpunkte ab und die Belastung durch „Schmerzen“ geht um 9,8 Prozentpunkte zurück.

Die Ausführungen zu „weiteren Anforderungen an eine adäquate Bewertung des Reha-Erfolgs“ verdeutlichen, dass neben der Problematik der Auswahl einer angemessenen Methode zur Messung der Ergebnisqualität rehabilitativer Interventionen (indirekte vs. direkte Veränderungsmessung) eine Reihe weiterer Aspekte

**Abbildung 1: Werte des Nottingham Health Profile zu Beginn (T0) und 3 Wochen nach Entlassung aus der Rehabilitation (T1) im Vergleich zur bundesdeutschen Referenzbevölkerung (Norm)**



bei der zielgerechten Ausgestaltung eines fundierten Qualitätssicherungssystems zur Bewertung des Rehabilitationsergebnisses zu berücksichtigen sind. Beispielfähig wurden dazu die besonders relevant erscheinenden Anforderungen an die Anzahl notwendiger Erhebungzeitpunkte zur Messung der Ergebnisqualität, an die Gewährleistung eines „fairen“ Vergleichs unterschiedlicher Rehabilitationseinrichtungen sowie an die Auswahl des einzusetzenden Instrumentariums zur Ergebnismessung diskutiert.

Die Ausführungen im Rahmen dieses Beitrags sollten sich aber nicht nur auf eine Problematisierung wesentlicher Aspekte zur Bewertung des Reha-Erfolgs beschränken, sondern auch beispielhaft Möglichkeiten zu einem anforderungsgerechten Umgang mit den problematisierten Aspekten aufzeigen.

### 3. Literatur

**Bitzer, E. M., Lorenz, C. & Dörning, H.** (2003): Comparison of indirect and direct measures of change in symptoms after elective surgery. *Quality of Life Research*, 12 (7), 243.

**Bitzer, E. M., Dörning, H. & Sampath Kumar, D.** (2004): Das Qualitätsbarometer - Hintergrund und Ausgestaltung. *DRV-Schriften*, Band 52, 127-128.

**Bitzer, E. M., Lorenz, C., Nickel, S., Dörning, H. & Trojan A.** (2008a): Assessing patient-reported outcomes of cholecystectomy in short-stay surgery. *Surgical Endoscopy*, 22 (12), 2712-2719.

**Bitzer, E. M., Lorenz, C., Nickel, S., Dörning, H. & Trojan, A.** (2008b): Patient-reported outcomes in hernia repair. *Hernia*, 12 (4), 407-414.

**Dörning, H., Bitzer, E. M., Sampath Kumar, D. & Lorenz, C.** (2004): Das Qualitätsbarometer – Ein Qualitätsbewertungssystem für die Medizinische Rehabilitation. In: Hellmann, W. (Ed.) *Praxis Klinischer Pfade*. Landsberg/Lech: Ecomed Verlagsgesellschaft, 1-23.

**Lorenz, K. & Mackensen, I.** (2004): Ergebnisqualität in der Rehabilitation nach Hüft- und Kniegelenkersatz – Das Qualitätsbarometer zur Erkennung von Optimierungspotenzial. *DRV-Schriften*, Band 52, 135-136.

**Rudolph, A., Bitzer, E. M., Dörning, H. & Sampath**

**Kumar, D.** (2007): Das Nottingham Health Profile: Ein geeignetes Instrument zur Erhebung der Ergebnisqualität in der kardiologischen und orthopädischen Rehabilitation? In: Wichmann, H. E., Nowak, D. & Zapf, A. (Eds.). Abstractband zum Kongress „Medizin und Gesellschaft 2007“ Augsburg. 17. - 21. September 2007. Mönchengladbach, Rheinware-Verlag, 289.

## Die Operationalisierung von „Rückkehr an den Arbeitsplatz“ als unterschätzte Herausforderung an die Abbildung der Ergebnisqualität

**Prof. Dr. Matthias Morfeld<sup>1</sup>**

University of Applied Sciences Magdeburg und Stendal  
Angewandte Humanwissenschaften  
Rehabilitationspsychologie

### Inhaltsverzeichnis

<b>1. Hintergrund</b>	<b>101</b>
<b>2. Methode</b>	<b>105</b>
<b>3. Ergebnisse</b>	<b>107</b>
<b>4. Diskussion</b>	<b>109</b>
<b>5. Literatur</b>	<b>110</b>

---

<sup>1</sup> Unter Mitwirkung von K. Hofreuter-Gätgens, Universitätsklinikum Hamburg-Eppendorf, Institut für Medizinische Psychologie.

# 1. Hintergrund

Für alle Formen der Leistungen zur Teilhabe (soziale, berufliche, medizinische) steht die (Re-)Integration in das soziale Umfeld und in das Erwerbsleben im Vordergrund. Dies betrifft grundsätzlich alle Akteure und Bereiche des Rehabilitationsprozesses: den Versicherten selbst, den Leistungsträger, den Leistungserbringer und nicht zuletzt auch das persönliche und Erwerbsumfeld des Versicherten. Für Deutschland ist das trägerspezifisch besonders festgelegt. Es gilt die so genannte einheitliche Risikoordnung, die deutlich macht, dass der Träger die Maßnahme finanziert, der auch das Risiko ihres Scheiterns trägt (Morfed et al., 2007).

Neben den psycho-sozial orientierten Verlaufs- und Zielkriterien (Bengel et al., 2008) die Aussagen über den Verlauf und den Erfolg einer Rehabilitationsmaßnahme bzw. besonderen Interventionen im Rahmen dieser machen und klinischen Parametern steht immer auch das Kriterium im Vordergrund, ob eine mehr oder weniger erfolgreiche Rückkehr in das Erwerbsleben stattgefunden hat.

Die Zielsetzung der beruflichen Reintegration gilt für die Rentenversicherung in besonderem Maße, da die Verhinderung bzw. Verzögerung vorzeitiger Berentungen Kosteneinsparungen bei den Rentenzahlungen bedeuten. Angestrebt wird demnach aus ökonomischen Gründen eine „möglichst dauerhafte“ Wiedereingliederung<sup>2</sup>.

Generell setzt eine Wiedereingliederung eine Erwerbsfähigkeit voraus. Darunter wird die Fähigkeit eines Versicherten verstanden, „unter Ausnutzung der

Arbeitsgelegenheiten, die sich ihm nach seinen Kenntnissen und Erfahrungen sowie seinen körperlichen und geistigen Fähigkeiten im ganzen Bereich des wirtschaftlichen Lebens bieten, Erwerbseinkommen zu erzielen“ (VDR, 2003, Haldorsen et al., 1998). Dies soll durch die ambulante oder stationäre Rehabilitation erreicht werden.

Die alleinige Wiederherstellung der Erwerbsfähigkeit fasst den Begriff der erfolgreich dauerhaften Wiedereingliederung jedoch nicht vollständig. Was aber genau darunter zu verstehen ist, bleibt offen, da es sich bei der beruflichen Reintegration um einen unbestimmten Rechtsbegriff handelt (Bürger, 1997). Eine allgemein akzeptierte, endgültige und empirisch fassbare Definition ist bislang in der Literatur noch nicht vorhanden.

In der bisherigen Forschungsliteratur wird die Rückkehr an den Arbeitsplatz oftmals mit dem bloßen Status der Erwerbstätigkeit oder Arbeitsfähigkeit gleichgesetzt (z. B. Maurischat & Mittag, 2004; Haldorsen et al., 1998; Hildebrandt et al., 1997). Diese Kriterien sind allerdings nicht hinreichend, da sie den beruflichen Werdegang der Rehabilitanden nicht erfassen können. Dies ist mit substantiellem Informationsverlust und -verzerrungen verbunden. z.B. könnte ein Rehabilitand zwar arbeitsfähig, jedoch aufgrund seiner Erkrankung langfristig arbeitslos geworden sein. Oder im Falle der Erwerbstätigkeit ist die Möglichkeit einer bereits beantragten Rente gegeben.

Der Begriff der dauerhaften Wiedereingliederung benötigt solche Kriterien, die ein möglichst genaues Abbild des Verlaufs der bisherigen und weiteren Erwerbstätigkeit wiedergeben. Bürger et al. (2002) versuchen, dieser Forderung nachzukommen und schlagen folgende Kriterien einer erfolgreichen dauerhaften Wiedereingliederung vor:

1. Aufnahme der Erwerbstätigkeit,

<sup>2</sup> SGB VI, §9 Aufgabe der Rehabilitation. (1) Die Rentenversicherung erbringt medizinische, berufsfördernde und ergänzende Leistungen zur Rehabilitation, um 1. den Auswirkungen einer Krankheit oder einer körperlichen, geistigen oder seelischen Behinderung auf die Erwerbsfähigkeit der Versicherten entgegenzuwirken oder sie zu überwinden und 2. dadurch Beeinträchtigungen der Erwerbsfähigkeit der Versicherten oder ihr vorzeitiges Ausscheiden aus dem Erwerbsleben zu verhindern oder sie möglichst dauerhaft in das Erwerbsleben wiederinzugliedern.

2. Fehlzeiten von unter 12 Wochen und
3. keinen Rentenantrag gestellt oder geplant.

Grundsätzlich ist die Aufnahme der Erwerbstätigkeit Voraussetzung einer erfolgreichen beruflichen Wiedereingliederung. Das zweite Kriterium resultiert aus Studienergebnissen, denen zufolge Fehlzeiten von über 12 Wochen die Wahrscheinlichkeit einer Gefährdung der langfristigen Wiedereingliederung beinhaltet (Bürger et al., 2001). Die dauerhafte Reintegration fordert außerdem die Ausübung der Erwerbstätigkeit möglichst bis zum gesetzlich festgelegten Renteneintrittsalter. Eine bereits beantragte Rente oder ein geplanter Rentenantrag stünde dem entgegen. Mithilfe dieser Kriterienauswahl können vergangene wie auch zukünftige Verläufe eruiert und einfach messbar gemacht werden.

Hinsichtlich des Zeithorizonts der Beurteilung der Rückkehr gibt es in der Literatur ebenfalls unterschiedliche Vorgehensweisen und Auffassungen. Während einige Untersuchungen das Zielkriterium direkt nach Rehabilitationseende erfassen, wird in anderen Untersuchungen der Evaluationszeitpunkt um 6, 12, 24 oder mehr Monate verschoben (z. B. Skonen et al., 2002; Mau et al., 2002; Bendix et al., 1998; Budde & Keck, 2001). Auch hier gilt, dass eine erfolgreiche dauerhafte Wiedereingliederung nicht kurzfristig, also direkt nach Beendigung der Maßnahme, beurteilt werden kann, sondern erst im Verlauf nach der Entlassung, wenn der Patient wieder ausreichend lange im Erwerbsleben ist.

Im Folgenden soll auf eine erste einfache Differenzierung des Begriff ‚Rückkehr an den Arbeitsplatz‘ eingegangen werden. Diese umfasst zum einen die inhaltliche Dimension – was bedeutet es eigentlich, wieder an den Arbeitsplatz zurückgekehrt zu sein? – und zum anderen erste Überlegungen für den Zeitpunkt der Rückkehr und seine Ausgestaltung.

### **1.1 Erfolgreiche Rückkehr zur Arbeit – die inhaltliche Dimension**

Aus Sicht der Gesetzlichen Rentenversicherung ist unter einer Wiederherstellung der Erwerbsfähigkeit eine dauerhafte Behebung der Minderung der Leistungsfähigkeit im Erwerbsleben zu verstehen (VDR, 2003). Dauerhaft ist dabei allerdings sozialrechtlich nicht näher definiert. Oftmals geschieht in diesem Kontext eine Gleichsetzung von Rückkehr an den Arbeitsplatz mit Arbeitsfähigkeit. Dies muss als problematisch angesehen werden, da offen bleibt, wie viele der nach der Rehabilitationsmaßnahme als arbeitsfähig Entlassenen tatsächlich an den Arbeitsplatz zurückkehren. Die Herstellung der Arbeitsfähigkeit ist dabei lediglich als eine Voraussetzung für die berufliche Wiedereingliederung zu sehen. Unklar bleibt darüber hinaus der Umgang mit Phasen der Arbeitslosigkeit, einem Wechsel der Berufstätigkeit, einer Reduktion des Stundenumfangs bei Teilzeitbeschäftigung oder qualitative Veränderungen der Arbeitstätigkeit an sich.

### **1.2 Erfolgreiche Rückkehr zur Arbeit – die zeitliche Dimension**

In der Literatur finden sich unterschiedliche Vorgehensweisen und Auffassungen bzgl. der Operationalisierung und Erfassung der Rückkehr an den Arbeitsplatz. Aus methodischer Perspektive der Rehabilitationsforschung ist jedoch klar zu folgern, dass eine erfolgreiche dauerhafte Wiedereingliederung nicht kurzfristig, also direkt nach Beendigung einer Rehabilitationsmaßnahme erhoben werden kann. Dabei sind Evaluationszeitpunkte von mindestens 6 Monaten erforderlich, wünschenswert jedoch sind deutlich längere (VDR, 1999). Grundsätzlich muss in diesem Kontext auch auf die zu erfragende Absicht des Versicherten, wann er denn überhaupt wieder vorhat, an den Arbeitsplatz zurück zu kehren, eingegangen werden. Aus verschiedenen theoretischen Modellen aus der Gesundheitspsychologie (z. B. HAPA, Ralf, 1999) ist bekannt, dass der Intention unter bestimmten Bedingungen eine substantielle Prognostizität zukommt.

Zusammenfassend kann hier festgehalten werden, dass das Konstrukt ‚Return to Work‘ eine substantielle Verwendung als Ergebnisparameter darstellt, jedoch auch verschiedene Konzepte seiner Operationalisierung impliziert. Grundsätzlich muss jedoch eine eher diffuse Operationalisierung der Rückkehr an den Arbeitsplatz in rehabilitationswissenschaftlichen Studien konstatiert werden.

Im Folgenden soll von daher anhand einer rehabilitationswissenschaftlichen Untersuchung exemplarisch aufgezeigt werden, wie differenziert die Operationalisierung von ‚Rückkehr an den Arbeitsplatz‘ umgesetzt werden kann und sollte (Morfeld et al., 2006).

## 2. Methode

In der vorliegenden Untersuchung soll aufgrund dieser Überlegungen die Begriffsbestimmung der erfolgreichen dauerhaften Wiedereingliederung nach Bürger et al. (2002) verwendet werden. Außerdem wird der Zeitpunkt 12 Monate nach Ende der Rehabilitation gewählt, zum einen da eine ‚stabile‘ Eingliederung auch eine langfristige impliziert und zum anderen aus forschungspraktischen Gründen. Gleichzeitig kann somit die Vergleichbarkeit der Ergebnisse dieser Arbeit mit denen anderer Untersuchungen sichergestellt werden. Zur Prüfung der Wirksamkeit eines neu konzipierten kognitiv-verhaltenstherapeutischen Programms (Back To Balance) wurde eine kontrollierte quasi-experimentelle prospektive Therapiestudie in zwei großen Rehabilitationseinrichtungen durchgeführt. Eingeschlossen wurden Patienten mit Zustand nach erster Bandscheibenoperation, im Alter zwischen 20 und 60 Jahre sowie entsprechend dem Hauptzielkriterium das Vorhandensein eines festen Arbeitsplatzes. Ausschlusskriterien waren ein bereits gestellter Rentenantrag sowie Versteifungen im LWS-Bereich. Die Stichprobenbeschreibung zeigt die folgende Tabelle 1.

**Tabelle 1: Soziodemografischer Kennwerte der Untersuchungstichprobe**

		<b>% (N)</b>
Geschlecht	weiblich	54,8% (143)
	männlich	45,2% (118)
	Alter i. Jahren im Mittel [niedrigstes - höchstes Alter]	41,7 J. [21 – 64]
Familienstand	ledig	29,0% (75)
	verheiratet	56,0% (145)
	verwitwet	0,8% (2)
	geschieden	12,0% (31)
	getrennt lebend	2,3% (6)
Schulabschluss	Ohne Schulabschluss	2,3% (6)
	Haupt-/Volksschule	30,9% (80)
	Realschule / Mittlere Reife	34,0% (88)
	Polytechnische Oberschule	3,5% (9)
	Fachhochschulreife	5,8% (15)
	Allgemeine / fachgebundene Hochschulreife / Abitur	22,0% (57)
	andere	1,5% (4)
Ausbildung	keine Berufsausbildung	6,4% (16)
	Beruflich-betriebliche Ausbildung (Lehre)	58,0% (145)
	Beruflich-schulische Ausbildung	10,4% (26)
	Fachschule, Meister-, Technikerschule, Berufs- oder Fachakademie	9,6% (24)
	Fachhochschul-Abschluss	8,0% (20)
	Hochschulabschluss	6,8% (17)
	andere	0,8% (2)
Träger der Rehabilitationsmaßnahme	BfA	23,9% (53)
	LVA	65,8% (146)
	GKV	2,7% (6)
	Freie Heilfürsorge	3,2% (7)
	Privat	2,7% (6)
	sonstige	1,8% (4)

### 3. Ergebnisse

Die zentrale Frage der hier zugrunde gelegten Studie ist die nach der erfolgreichen, d.h. stabilen Wiedereingliederung in den Beruf. Gemäß den theoretischen Ausführungen ist eine stabile Wiedereingliederung dann erfolgt, wenn die Versicherten ein Jahr nach Beendigung der Rehabilitation an den Arbeitsplatz zurückgekehrt sind, sie weder einen Rentenanspruch gestellt noch die Intention dazu haben und Fehlzeiten von weniger als 12 Wochen im Jahr nach der Reha vorliegen. Die Abbildung 1 veranschaulicht, dass 95% (n = 205) der Rehabilitanden innerhalb eines Jahres nach Reha-Ende in die Erwerbstätigkeit zurückkehren. Die übrigen 5% sind entweder arbeitslos, befinden sich in Umschulung oder sind Hausfrau/Hausmann geworden.

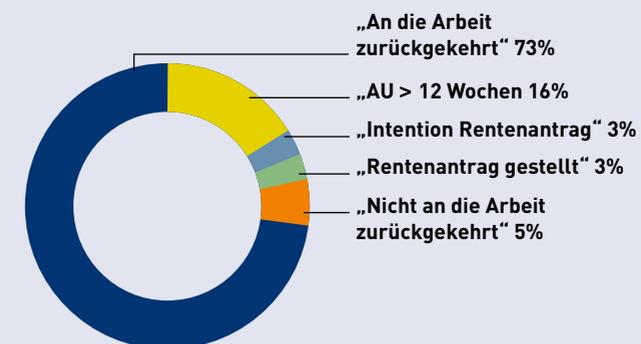
**Abbildung 1: Anteil der Patienten, die ein Jahr nach der medizinischen Reha-Maßnahme wieder an den Arbeitsplatz zurückgekehrt sind**



Die Quote derjenigen, die an den Arbeitsplatz zurückgekehrt sind (n = 205), reduziert sich jedoch aufgrund der zu berücksichtigenden Messkriterien. In dem einen Jahr nach der Reha haben 2,5% (n = 5) einen Rentenanspruch gestellt, 2,7% (n = 5) bekunden die Intention, einen Rentenanspruch zu stellen und 15,6% (n = 32) weisen eine längere AU-Dauer als 12 Wochen auf.

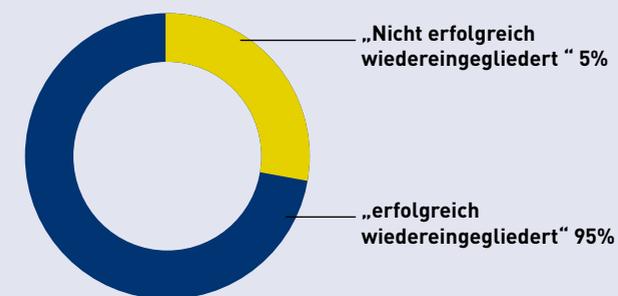
Darunter befinden sich vier Patienten, die durchgängig seit der Reha krankgeschrieben sind. Somit ergibt sich ein Anteil von 73% erfolgreich wiedereingegliederten Patienten (siehe Abbildung 2).

**Abbildung 2: Anteil der erfolgreich wiedereingegliederten Patienten ein Jahr nach Reha-Ende**



Demnach verringert sich die ursprüngliche Wiedereingliederungsquote von 95% unter den gewählten Kriterien um 21%. Von der Stichprobe zu t3 (n = 205) konnten insgesamt 73% der Rehabilitanden erfolgreich eingegliedert werden (vgl. Abbildung 3).

**Abbildung 3: Anteil stabil an den Arbeitsplatz zurückgekehrter Patienten 1 Jahr nach Rehabilitationsende**



Insgesamt verdeutlichen die gefundenen Resultate, welche unterschiedlichen Aussagen möglich sind, wenn man die Operationalisierung der Rückkehr an den Arbeitsplatz unterschiedlich gewichtet und ausgestaltet.

## 4. Diskussion

Die gezeigten Ergebnisse machen deutlich, dass eine methodisch gut abgesicherte und mit sozialrechtlichen Kriterien hinterlegte Operationalisierung von ‚Rückkehr an den Arbeitsplatz‘ nicht trivial ist. Deutlich wird aber auch, dass bei unterschiedlicher Gewichtung der einzelnen Erhebungsbausteine durchaus unterschiedliche Ergebnisse erzielt werden können. In diesem Zusammenhang ist vergleichbar mit den Empfehlungen zum Einsatz standardisierter Assessmentverfahren für die subjektive Gesundheit und Funktionsfähigkeit auch eine im Konsens aller Beteiligten (Kostenträger, Leistungserbringer, Rehabilitationsforscher) getroffene Empfehlung für die Abbildung einer Rückkehr an den Arbeitsplatz nach erfolgter Rehabilitation zu treffen. Aus Sicht der Autoren sollte bei einer solchen Harmonisierung auch zugleich eine Passung für den Bereich der Leistungen zur Teilhabe am Arbeitsleben und der Gemeinschaft mitgedacht werden. Dies vor allem auch aus dem Grunde da hier z. T. andere Kostenträger involviert sind, die einen solchen Harmonisierungsprozess durchaus befruchten würden und zudem schließlich und endlich ein sektorenübergreifendes Instrumentarium vorliegen würde.

Darüber hinaus ergeben sich für die Autoren eine ganze Reihe weiterer Arbeitsschritte hin zu einer weiteren Harmonisierung von ‚Rückkehr an den Arbeitsplatz‘. Hierzu gehört natürlich an erster Stelle die Entwicklung eines umfassend akzeptierten Erhebungsinstrumentes und seiner psychometrischen Prüfung, so wie es verschiedene Autoren auch schon angedeutet haben (Maurischat & Mittag, 2004; Mittag & Raspe, 2003). Ferner

sollte in einer Arbeitsgruppe geprüft werden, welche Daten bzgl. der Frage nach Rückkehr an den Arbeitsplatz bei den Kostenträgern geprüft und für Weiterentwicklungen genutzt werden können.

Was in der hier vorgelegten Arbeit nicht geklärt werden konnte, sind eher inhaltliche Fragen nach der Prognostizität – das bedeutet, welche Faktoren sind denn eigentlich unter welchen Bedingungen für eine mehr oder weniger erfolgreiche Rückkehr an den Arbeitsplatz verantwortlich. Eine Reihe von eher theoretischen (Hanson & Kopjar, 2003) aber auch empirischer (Muche et al., 2000) Arbeiten weist auf den Klärungsbedarf an dieser Stelle hin. Ebenfalls einer Prüfung bedarf aus Sicht der Autoren die Frage nach der Schnittmenge, also nach der Versichertengruppe die auf den psychosozialen Skalen eher auch aus prognostischer Sicht ‚gute‘ Werte erreicht haben jedoch nicht wieder an den Arbeitsplatz zurückgekehrt sind. Oder aus entgegen gesetzter Perspektive: Bedeutet die nach der hier vorgeschlagenen Definition von Rückkehr an den Arbeitsplatz tatsächlich erfolgte erfolgreiche Wiederaufnahme der Erwerbstätigkeit auch gleichzeitig eine Verbesserung der subjektiven Gesundheit?

## 5. Literatur

**Bendix, A. F., Bendix, T. & Haestrup, C.** (1998): Can it be predicted which patients with chronic low back pain should be offered tertiary rehabilitation in a functional restoration program? A search for demographic, socioeconomic, and physical predictors: Functional restoration. *Spine*, 23 (16), 1775-1783.

**Bengel, J., Wirtz, M. & Zwingmann, Ch.** (2008): Diagnostische Verfahren in der Rehabilitation. Göttingen: Hogrefe-Verlag.

**Budde, H.-G. & Keck, M.** (2001): Prädiktoren der beruflichen Wiedereingliederung nach stationärer kardiologischer Rehabilitation im Rahmen der Arbeiterrentenversicherung. *Die Rehabilitation*, 40, 208-216.

**Bürger, W. et al.** (1997): Arbeit, Psychosomatik und medizinische Rehabilitation: Eine Längsschnittuntersuchung. Bern: Hans Huber-Verlag.

**Bürger, W. et al.** (2001): Multiperspektivische Einschätzungen zur Wahrscheinlichkeit der Wiedereingliederung von Patienten im Erwerbsleben nach orthopädischer Rehabilitation: Ergebnisse und prognostische Relevanz. *Die Rehabilitation*, 40, 217-225.

**Bürger, W. et al.** (2002): Ambulante und stationäre orthopädische Rehabilitation: Ein Vergleich von Strukturmerkmalen, Wirksamkeit und Kosten. Abschlussbericht Hamburg: Abteilung für medizinische Psychologie.

**Haldorsen, E.M.H. et al.** (1998): Predictors for outcome of a multi-modal cognitive behavioral treatment program for low back pain patients: a 12-month follow-up study. *European Journal of Pain*, 293-307.

**Hanson, B.P. & Kopjar, B.** (2003): Klinische Studien, Outcome-Parameter – Ein Paradigma-Wechsel? *Chirurg*, 74, 1034-1039.

**Hildebrandt, J. et al.** (1997): Prediction of Success From a Multidisciplinary Treatment Program for Chronic Low Back Pain: Functional Restoration. *Spine*, 22 (9), 990-1001.

**Mau, M. et al.** (2002): Prognose der sozialmedizinischen Entwicklung ein Jahr nach teilstationärer oder stationärer Rehabilitation wegen Dorsopathie. *Die Rehabilitation*, 41, 160-166.

**Maurischat, C. & Mittag, O.** (2004): Erfassung der beruflichen Rollenqualität und ihre prognostische Bedeutung für die Wiederaufnahme der Berufstätigkeit. *Die Rehabilitation*, 43, 1-9.

**Mittag, O. & Raspe, H.** (2003): Eine kurze Skala zur Messung der subjektiven Prognose der Erwerbsfähigkeit: Ergebnisse einer Untersuchung an 4.279 Mitgliedern der gesetzlichen Arbeiterrentenversicherung zu Reliabilität (Guttman-Skalierung) und Validität der Skala. *Die Rehabilitation*, 42, 169-174.

**Morfeld, M., Man, w., Jäckel, W. H. & Koch, M.** (2007): Im Querschnitt – Rehabilitation, physikalische Medizin und Naturheilverfahren. München: Urban & Fischer-Verlag.

**Morfeld, M. et al.** (2006): Kognitiv-verhaltenstherapeutische Intervention bei Patienten nach erster Bandscheiben Operation und die Rückkehr zur Arbeit. *Verhaltenstherapie & Verhaltensmedizin*, 27 (3), 338-362.

**Muche, R. et al.** (2000): Entwicklung und Validierung eines Prognosemodells zur Vorhersage der Arbeitsfähigkeit nach Rehabilitation anhand routinemäßig erhobener Parameter. *Die Rehabilitation*, 39 (5), 262-267.

**Ralf, S.** (1999): Self-regulatory Processes in the Adoption and Maintenance of Health Behaviors: The Role of Optimism, Goals, and Threats. *Journal of Health Psychology*, 4 (2), 115-127.

**Skouen, J.S. et al.** (2002): Relative cost-effectiveness of extensive and light multidisciplinary treatment programs versus treatment as usual for patients with chronic low back pain on long-term sick leave: Randomized controlled study. *SPINE*, 27 (9), 901-909.

**Verband Deutscher Rentenversicherungsträger** (1999): Förderschwerpunkt „Rehabilitationswissen-

schaften“ – Empfehlungen der Arbeitsgruppen „Generische Methoden“, „Routinedaten“ und „Reha-Ökonomie“. DRV-Schriften, Band 16, Frankfurt/Main.

**Verband Deutscher Rentenversicherungsträger**  
(2003): Sozialmedizinische Begutachtung für die gesetzliche Rentenversicherung. Berlin/Heidelberg: Springer-Verlag.

## Die Bedeutung subjektiver Erfolgsmaße für die Ergebnisqualität<sup>1</sup>

**Dr. Thorsten Meyer**  
Universität zu Lübeck  
Universitätsklinikum Schleswig-Holstein  
Institut für Sozialmedizin  
Campus Lübeck

### Inhaltsverzeichnis

<b>1. Problemhintergrund</b>	<b>115</b>
<b>2. Objektivität und Messung subjektiver Erfolgsmaße</b>	<b>121</b>
<b>3. Beispiele von Ergebnissen mit subjektiven Erfolgsmaßen in der Rehabilitation</b>	<b>124</b>
<b>4. Resümee</b>	<b>127</b>
<b>5. Literatur</b>	<b>127</b>

---

<sup>1</sup> Dieser Beitrag basiert auf einem Vortrag vom 25. November 2008 in der Deutschen Rentenversicherung Bayern-Süd.

# 1. Problemhintergrund

Die Deutsche Rentenversicherung setzt seit 1997 subjektive Erfolgsmaße zur Bewertung von Leistungen zur medizinischen Rehabilitation im Rahmen ihrer Qualitätssicherung ein (Egner, Gerwinn & Buschmann-Steinhage, 2006). Eine Zufallsstichprobe von Rehabilitanden aus Kliniken, die von der Deutschen Rentenversicherung betreut werden, erhält etwa acht bis zwölf Wochen nach ihrem Rehabilitationsaufenthalt einen umfangreichen Fragebogen zugeschickt, der neben Fragen zur Rehabilitandenzufriedenheit, zu Leistungen während des Aufenthaltes sowie Fragen zum aktuellen Gesundheitszustand auch Fragen zum Erfolg der Rehabilitation umfasst. Im Folgenden möchte ich mich auf diese bisher in der Darstellung der Rehabilitandenbefragung eher vernachlässigten Erfolgsindikatoren beziehen. Es handelt sich – neben der Rehabilitandenzufriedenheit – um Kriterien der Ergebnisqualität. In der Rehabilitandenbefragung kommt die Methode der direkten Veränderungsmessung zum Einsatz (vgl. Kohlmann & Raspe, 1998), d.h. die Betroffenen werden nach erfolgter Maßnahme gebeten anzugeben, ob bzw. inwieweit sich Veränderungen in einem bestimmten Bereich (z.B. „Kreuzschmerzen“) ergeben haben<sup>2</sup>.

Welche Bedeutung kommt diesen Indikatoren der Ergebnisqualität zu? Welchen Folgen möchten wir ihnen zugestehen? Gegenwärtige Praxis ist die regelmäßige, halbjährliche Rückmeldung der Ergebnisse an die einzelnen von der Deutschen Rentenversicherung betreuten Kliniken. Sie beinhaltet einen expliziten Vergleich zum mittleren Wert vergleichbarer Kliniken und zum besten Wert, den eine Klinik mit vergleichbarem Indikationsspektrum erreicht hat. Mit dieser Rückmeldung

---

<sup>2</sup> Die direkte Veränderungsmessung wird abgegrenzt von der indirekten Veränderungsmessung, die auf der Bildung der Differenzen von Statusmessungen zu mindestens zwei Messzeitpunkten basiert.

erhält das interne Qualitätsmanagement die Möglichkeit, potenzielle Schwachstellen im eigenen Leistungsspektrum zu identifizieren und Qualitätsprobleme zielgerichteter bearbeiten zu können. Die Daten über den Rehabilitationserfolg können aber nicht nur intern verwendet werden. Prinzipiell eröffnen sie die Möglichkeit, auch zukünftigen Rehabilitanden als Information zur Verfügung gestellt zu werden. Die Ergebnisse könnten in der Außendarstellung der Klinik genutzt werden, in der z.B. auf besondere Stärken und Erfolge der Klinik hingewiesen wird. Den Rehabilitanden wurde zudem mit der Einführung des SGB IX im Rahmen des Wunsch- und Wahlrechts nach §9 das Recht zugestanden, selbst eine Klinik auszuwählen und der Leistungsträger ist aufgefordert, diesem Wunsch zu entsprechen, wenn er sich nicht als ungerechtfertigt erweist. Um diesem Recht nachkommen zu können, bedarf der angehende Rehabilitand angemessene, vergleichende Informationen. Die Erfolgsraten einer Klinik könnten dabei eine zentrale Entscheidungsgrundlage darstellen. Dieser Gedanke ist auf die Interessen der Rehabilitanden ausgerichtet. Er ist weniger attraktiv für diejenigen Kliniken, die in diesen Vergleichen eher schlechter abschneiden (Meyer et al., 2008, 2009).

Auch der Leistungsträger in Form der Deutschen Rentenversicherung kann die Daten zur externen Qualitätssicherung und -kontrolle nutzen. Für den Bereich der Rehabilitandenzufriedenheit hat er diese Möglichkeit schon genutzt<sup>3</sup>. Auch für den Bereich der Rehabilitationserfolge liegen erste vergleichende Studien vor (Meyer et al., 2009). Aus Sicht der Leistungsträger könnten die Erfolgsdaten zusätzlich dazu verwendet werden, Rehaprozesse zu steuern, z.B. über höhere Zuweisungsquoten für erfolgreichere Kliniken oder auch prinzipiell in der Entwicklung von am Rehaerfolg orientierten

---

<sup>3</sup> Hier hatte sich über die Jahre gezeigt, dass der Bereich „Rehaplan und -ziele“ wiederholt schlechter beurteilt wurde als die anderen Zufriedenheitsbereiche. Zu dieser Problemstellung wurde ein wissenschaftliches Projekt gefördert, um Gründe für die vergleichsweise schlechtere Bewertung zu identifizieren (Meyer et al., 2008, 2009).

monetären Anreizsystemen („pay-for-performance“). Subjektive Erfolgsmaße als Kriterium für die Auswahl von Kliniken durch die Rehabilitanden? Als Kriterium der Zuweisung von Rehabilitanden? Als Grundlage der Vergütung? Dies ist prinzipiell möglich und hat vor dem Hintergrund der Diskussionen um Qualitätsentwicklung in der Rehabilitation einen nicht von der Hand zu weisenden Reiz. Aber spätestens an dieser Stelle regt sich der Widerspruch in uns:

**Angaben von Rehabilitanden können nicht ein Ausmaß an Gültigkeit für sich beanspruchen, um diese Folgen zu rechtfertigen! Oder etwa doch?**

Um sich dieser Frage anzunähern, möchte ich auf einige Aspekte eingehen, die uns bei der kritischen Würdigung subjektiver Erfolgsmaße Bauchschmerzen bereiten können. Dazu gehören das Problem der Subjektivität, der Objektivität und Messbarkeit von Patientenangaben und ein Blick auf Beispiele zur Plausibilität von subjektiven Erfolgsmaßen.

**Subjektivität subjektiver Erfolgsmaße**

Was bedeutet Subjektivität in diesem Kontext? Um welche Art von Subjektivität geht es hier? Eine Ebene der Subjektivität bezieht sich auf den Rehabilitanden als Informationsquelle. Die Food and Drug Administration der US-amerikanischen Regierung hat den Begriff der „Patient-reported Outcomes“ (PRO) geprägt (2006). Ein PRO sei demnach die Messung jedweden Aspekts der Gesundheit eines Patienten, der direkt vom Patienten selber komme, d.h. ohne die Interpretation der Antwort des Patienten durch einen Arzt oder jemand anderen. Eine andere Ebene von Subjektivität ist gemeint, wenn es um die Erfassung von Zuständen geht, die direkt nur dem Betroffenen selber zugänglich sind: z.B. Gefühle, Stimmungen, Einstellungen. Sie bezieht sich demnach nur auf einen Teil der PRO-Definition. Diese Unterscheidung ist insbesondere wichtig für die Möglichkeit und Form, mit der die Angaben der Befragten prinzipiell

validiert werden können. Es gibt noch eine dritte wichtige Ebene von Subjektivität, die in der o.g. PRO-Definition nicht klar getrennt wird: die Bewertungsinstanz liegt in der befragten Person selber. Insbesondere in der Lebensqualitätsforschung wurde auf die Bedeutung dieses Unterschiedes hingewiesen (Schwartz & Raplein, 2004; Meyer, 2004). Es ist ein wichtiger Unterschied, ob ein Patient angeben soll, inwieweit er noch körperlich in der Lage ist, bestimmte Handlung auszuführen (z.B. zwei Stockwerke Treppen zu steigen), oder ob nach seiner Zufriedenheit mit der körperlichen Leistungsfähigkeit gefragt wird. Im ersten Fall wird der Befragte als Informationsquelle für einen prinzipiell objektiv überprüfbaren Sachverhalt angesehen, im letzteren Fall ist nach einer Einschätzung gefragt, die nicht nur von der tatsächlichen oder wahrgenommenen Leistungsfähigkeit abhängt, sondern von weiteren psychischen Eigenschaften (z.B. Anpassung des eigenen Anspruchsniveaus). Diese Unterscheidung („reporting“ vs. „rating“) ist neben der schön erwähnten Erfassung von Lebensqualität insbesondere auch für die Erfassung der Patientenzufriedenheit von Bedeutung. Subjektive Erfolgsmaße können sich prinzipiell auf alle drei Ebenen beziehen. Sie stellen unterschiedliche Formen von Subjektivität dar, was bei der Interpretation der Erfolgsmaße berücksichtigt werden muss.

Das mag ja so sein, werden Sie sich nun denken. Bedeutet es nicht, dass die Ergebnisse aus solchen Befragungen unweigerlich durch persönliche Faktoren „verzerrt“ sind, z.B. durch mangelnde Fähigkeiten, Fragebögen auszufüllen, schwierige Persönlichkeiten, negative Einstellungen zur Rehabilitation, Über- oder Unterschätzung der eigenen Leistungsfähigkeit, inadäquate Krankheitstheorien o.ä.? Ja, wenn wir an dem Ideal einer von außen bestimmbaren objektiven Wahrheit festhalten. Damit schließen wir allerdings die zweite und dritte aufgeführte Ebene der Subjektivität von Erfolgsmaßen kategorisch aus. Eine andere Perspektive eröffnet sich, wenn die Angaben des Rehabilitanden als eine Ebene

der Wirklichkeit betrachtet werden, so, wie der Rehabilitand sich zum Zeitpunkt der Befragung selber sieht.

Diese subjektive Ebene der Wirklichkeit kann sehr bedeutsam sein. Das musste einmal mein (ehemaliger!) Zahnarzt erfahren, der mir mit den Worten „Das kann gar nicht wehtun“ eine zusätzliche Betäubung verwehrt (so genannte berühmte letzte Worte...). Mir war es, eingequetscht in den Zahnarztstuhl, mit aufgerissenem Mund, ziemlich egal, wie meine Nervenbahnen im Mundraum verlaufen müssten und ob ich theoretisch Schmerzen haben könnte oder nicht. Ich hatte sie, sie waren für mein Leiden und mein Verhalten verantwortlich. Sie stellten für mich meine eigene, höchst relevante Ebene der Wirklichkeit dar.

Dass dieser Vorfall mehr ist als eine Anekdote, wissen wir aus Untersuchungen zur Bewertung des eigenen Gesundheitszustands (subjektive Gesundheit). Die einfache Frage „Wie würden Sie Ihren Gesundheitszustand im Allgemeinen beschreiben?“ mit den Antwortalternativen *exzellent, gut, zufriedenstellend, schlecht*<sup>4</sup> steht in Zusammenhang mit der Überlebensdauer der Befragten – und zwar unabhängig von ihrem zum Zeitpunkt der Befragung „objektiv“ erhobenen Gesundheitszustand. Z.B. berichteten Heidrich et al. (2002) im Rahmen der MONICA-Studie aus Augsburg standardisierte Mortalitätsraten der untersuchten Kohorte in Abhängigkeit der Bewertung ihres allgemeinen Gesundheitszustandes. Demnach steigt die Mortalitätsrate mit schlechter Bewertung der eigenen Gesundheit: von 4,5 (pro 1.000 Personenjahre) bei den Befragten mit „exzellenter“ Gesundheit, über 8,4 bei Befragten mit „guter“ Gesundheit, 16,0 bei „zufriedenstellender“ und 13,5 mit als „schlecht“ bewerteter Gesundheit. Unabhängig vom in der MONICA-Studie ausführlich erhobenen klinischen Status der Befragten ist das Risiko zu versterben für

die Personen, die ihren Gesundheitszustand mit zufriedenstellend oder schlechter bewerten, immer noch um den Faktor 1,5 höher als gegenüber den Personen mit „gutem“ oder „exzellentem“ Gesundheitszustand (hazard rate ratio = 1,5; 95%-Konfidenzintervall 1,1-2,2). Der negative Zusammenhang zwischen selbstbeurteilter Gesundheit und Überlebenszeit gilt als einer der konsistentesten Befunde der sozialwissenschaftlichen Gesundheitsforschung, er wurde schon in den 1990er Jahren in einer systematischen Übersichtsarbeit zusammenfassend dargestellt (Idler & Benyami, 1997).

Auch im Bereich der Rehabilitation haben wir ein schönes Beispiel für die – prognostische – Bedeutung subjektiver Wirklichkeit vorliegen. Die Skala zur Erfassung der subjektiven Prognose der Erwerbstätigkeit (kurz: SPE-Skala) besteht aus den folgenden drei Items (Mittag & Raspe, 2003):

**Wenn Sie an Ihren derzeitigen Gesundheitszustand und Ihre berufliche Leistungsfähigkeit denken ...**

1. Glauben Sie, dass Sie Ihre jetzige (letzte) Tätigkeit bis zum Erreichen des Rentenalters ausüben können? (Bitte nur ein Kästchen ankreuzen)

sicher       eher ja       unsicher       eher nein       auf keinen Fall

2. Sehen Sie durch Ihren derzeitigen Gesundheitszustand Ihre allgemeine Erwerbsfähigkeit dauerhaft gefährdet?

nein       ja

3. Tragen Sie sich zur Zeit mit dem Gedanken, einen Rentenantrag aus Gesundheitsgründen (Erwerbsminderungsrente) zu stellen?

nein       ja       habe bereits einen Rentenantrag gestellt

<sup>4</sup> Für diese Frage liegen alternative Formulierungen sowohl für den Itemstamm als auch für das Antwortformat vor.

Das erste Item wird zur Auswertung dichotomisiert („sicher“/„eher ja“ = 0; sonst = 1), die zustimmenden Antworten von Frage 2 und 3 erhalten den Wert 1, alle drei Werte werden zu einem Summenwert addiert, der von 0 (gute Prognose) bis 3 (schlechte Prognose) variiert. In einer Bevölkerungsstichprobe arbeitslosenversicherter Personen zur Prognose eines Rentenantrags und einer vorzeitigen Berentung ergab sich für den Zeitraum von vier bis sechs Jahren unter Berücksichtigung verschiedener prognostisch relevanter Faktoren<sup>5</sup> eine mehr als doppelt so hohe Wahrscheinlichkeit für einen Rentenantrag von Personen mit SPE = 3 gegenüber solchen mit SPE = 0 (adj. Hazard Ratio = 2,378; 95%-Konfidenzintervall; 1,518-3,726) und eine um den Faktor 1,8 erhöhte Wahrscheinlichkeit für eine tatsächliche vorzeitige Berentung (adj. Hazard Ratio = 1,808; 95%-Konfidenzintervall; 0,966-3,385) (vgl. dazu Mittag et al., 2006).

## 2. Objektivität und Messung subjektiver Erfolgsmaße

Der gedankliche Ansatz, die subjektiven Angaben des Rehabilitanden als eine – relevante – Ebene der Wirklichkeit zu betrachten, hat sich also als durchaus fruchtbar erwiesen. Tatsächlich ist es so, dass subjektive Erfolgsmaße auf ihre Objektivität untersucht werden können. Hiermit ist nicht gemeint, dass wir den Angaben des Rehabilitanden denen des Arztes gegenüberstellen<sup>6</sup>. Wir haben bei der Bestimmung des Begriffs Subjektivität gesehen, dass wir damit bestimmte Aspekte der subjektiven Wirklichkeit der Betroffenen systematisch ausgrenzen würden. Vielmehr ist mit die-

5 Alter, Geschlecht, Anzahl chronischer Erkrankungen, Dauer der Arbeitsunfähigkeit, Grad der Behinderung, allgemeiner Gesundheitszustand, berufliches Leistungsvermögen (Selbstangaben).

6 Grande (1998) bezeichnete diese Form von Objektivität als das „medizinische Modell“.

sem Begriff die Objektivität der Messung angesprochen, d.h. Objektivität als ein Gütekriterium der Messung von Personeneigenschaften<sup>7</sup>. Sie ist definiert als das Ausmaß, in der das Ergebnis der Messung unabhängig von jeglichen Einflüssen außerhalb der untersuchten Person ist (Lienert & Raatz, 1995).

Messung? Können wir bei der Erfassung subjektiver Erfolgsmaße von Messung sprechen? Wenn wir den Begriff Messung verwenden, denken wir an erster Linie an ein Messgerät, oder etwas altmodischer an einen Zollstock oder Lineal. Mit so einer Messlatte können wir beispielsweise die Körpergröße einer Person messen. Wir könnten auch die betroffene Person danach fragen, wie groß sie ist. Die Gültigkeit der Antwort würde davon abhängen, ob der Befragte seine eigene Größe kennt – und ob er nicht ggf. dazu neigt, seine eigene Größe eher zu überschätzen, was bei Männern durchaus vorkommen soll. Wenn wir nun darauf zielen, die Körpergröße mit einem Fragebogen zu messen, müssen wir Aussagen

oder Fragen formulieren, deren Beantwortung von der Körpergröße einer Person abhängt. Beispiele für mögliche Aussagen finden sich in Abbildung 1<sup>8</sup>. Auf der linken Seite ist die interessierende, die zu messende Größe dargestellt. Auf der rechten Seite sind Indikatoren, die mit der Ausprägung der Körpergröße in Zusammenhang stehen, d.h. es sind Indikatoren der Körpergröße. Wenn z.B. jemand die Aussage „Auf Flugreisen sitze ich normalerweise bequem“ bejaht, ist dies ein Indikator für eine mittlere bis kleinere Körpergröße. Aus den Antworten auf die einzelnen Items lässt sich auf die Körpergröße zurück schließen – wenn das Modell, hier das Messmodell, gilt, d.h. die Items tatsächlich Indikatoren der Körpergröße darstellen. Genau genommen

7 Zur „klassischen Trias“ der Gütekriterien der Messung psychischer Eigenschaften gehören außerdem die Zuverlässigkeit („Reliabilität“, Messgenauigkeit) und Gültigkeit („Validität“).

8 Dieses Beispiel ist einem Vortrag von C.A.W. Glas zum Thema „IRT modelling of longitudinal data“ am Institut für Pädagogik der Naturwissenschaften an der Christian-Albrechts Universität zu Kiel am 23.08.2007 entnommen.

lässt sich nicht auf den absoluten Wert der Körpergröße schließen, wohl aber auf die Beziehung von verschiedenen gemessenen Personen untereinander.

**Abbildung 1: Messmodell eines hypothetischen Fragebogens zur Messung der Körpergröße (latente Variable) durch Indikatoren (manifeste Variablen) in Form von Fragebogenitems**



Denken Sie sich nun anstelle der Körpergröße andere interessierende Variablen, die Sie gerne erfassen, „messen“ möchten, z.B. Schmerzintensität, Schmerzverarbeitung, Depressivität, Funktionskapazität, Arbeitsfähigkeit, Rehamotivation, Lebensqualität, Patientenzufriedenheit, soziale Unterstützung, subjektive Prognose der Erwerbsfähigkeit. All diese „Konstrukte“ können Sie nicht direkt beobachten, d.h. sie stellen latente Variablen dar. Sie müssen über manifeste Indikatoren erschlossen werden. Wie gut eine solche Messung tatsächlich ist, hängt von der Güte der Auswahl bzw. Formulierung der Indikatoren ab, im Falle von Fragebögen den Fragebogenitems. Die Güte einer solchen Messung lässt sich mithilfe von psychometrischen Analyseverfahren (z.B. Modellpassungstests im Rahmen konfirmatorischer Faktorenanalysen) überprüfen. Keine Messung ist perfekt, auch die diagnostischen Bewertungen des Arztes lassen sich als Messung

verstehen und auf ihre Güte untersuchen. D.h. die Erfassung von subjektiven Erfolgsmaßen ist prinzipiell nicht weniger objektiv als die Einschätzung eines Arztes. Die Objektivität der Messung, sowie ihre Zuverlässigkeit und Gültigkeit, muss sich auf der Basis eines Messmodells empirisch bewähren – auch die Einschätzungen von Ärzten.

### 3. Beispiele von Ergebnissen mit subjektiven Erfolgsmaßen in der Rehabilitation

Gehen wir also davon aus, dass subjektive Erfolgsmaße in der Lage sind, einen eigenen, potenziell bedeutsamen Teil der Wirklichkeit abzubilden, und dass wir sie prinzipiell genauso gut oder schlecht wie andere Erfolgsmaße messen können. Diese Annahmen sollen Sie bei der Betrachtung der folgenden Ergebnisse leiten. In einer systematischen Literaturübersicht zur Wirksamkeit der stationären Rehabilitation bei Patienten mit chronischen Rückenschmerzen in Deutschland sind Hüppe & Raspe (2005) u.a. zu folgenden Ergebnissen gekommen: Der gemittelte kurzfristige Effekt (dargestellt in Form von Effektstärken, ES)<sup>8</sup> der Rehabilitationsmaßnahme beträgt für die Schmerzintensität ES = 0,54, für die Depressivität ES = 0,70, für die Funktionskapazität ES = 0,30. Für die Interpretation dieser Effektstärken ist es hilfreich zu wissen, dass eine ES = 0,5 als mittlerer Effekt<sup>9</sup> betrachtet wird, ES = 0,2 als kleiner und eine ES von 0,8 als großer Effekt. D.h. es ist ermutigend zu sehen, dass die Rehabilitation es kurzfristig schafft, im Mittel relevante Veränderungen der Schmerzen und auch depressiver Symptome zu erzielen, allerdings

<sup>8</sup> Die Effekte werden in Übersichtsarbeiten in der Regel in Form von Effektstärken (ES) dargestellt. Dabei wird die Differenz zweier Mittelwerte ins Verhältnis zu ihrer Standardabweichung gesetzt. Für die Betrachtung eines Vergleichs von Werten vor der Reha mit Werten nach der Reha bedeutet das: Um wie viel Standardabweichungseinheiten hat die das Erfolgsmaß verändert? Eine ES von 0,5 bedeutet beispielsweise, dass die Veränderung im Mittel eine halbe Standardabweichung beträgt.

<sup>9</sup> Eine Übersichtsarbeit hat gezeigt, dass ab ein Wert von ES = 0,5 von einer für den Betroffenen bedeutsame Veränderung gesprochen werden kann (Mittag et al., 2006).

eher geringere Veränderungen im Bereich der Funktionskapazität. Dieses Beispiel macht es einem nicht schwer, die Ergebnisse als bedeutungsvoll zu interpretieren.

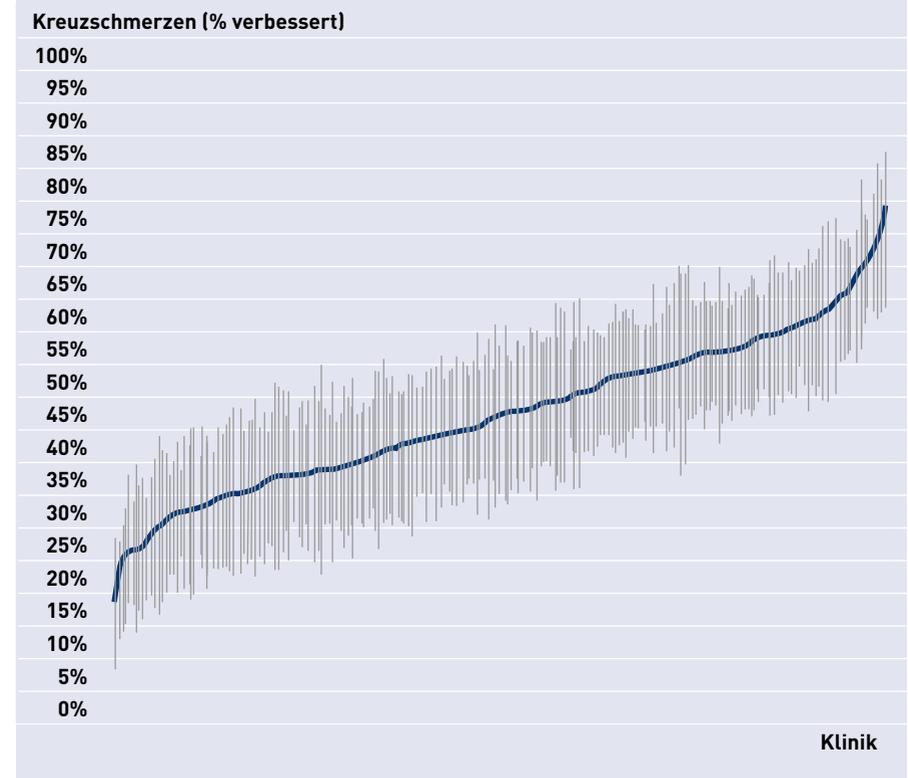
Ein Blick auf die mittelfristigen Erfolge nach drei bis zwölf Monaten ist allerdings ernüchternd: der Effekt für die Schmerzintensität liegt bei  $ES = 0,25$ , für die Depressivität bei  $ES = 0,16$ , für die Funktionskapazität bei  $ES = 0,01$ . Wenn diese Ergebnisse mit der gleichen Ernsthaftigkeit wie die der kurzfristigen Effekte betrachtet werden, müssen wir schlussfolgern, dass in den genannten Bereichen mittelfristig keine für die Rehabilitanden bedeutsamen Effekte zu erwarten sind, allenfalls kleine Verbesserungen der Schmerzintensität. Mit der Akzeptanz dieser Ergebnisse haben die Rehabilitationskliniker in der orthopädischen Rehabilitation seit dem Erscheinen der Übersichtsarbeit von Hüppe & Raspe zu kämpfen.

Ein weiteres Beispiel: Im Rahmen der Rehabilitandenbefragung in der Qualitätssicherung werden die Betroffenen acht bis zwölf Wochen nach der Rehamaßnahme u.a. gebeten anzugeben, ob sich ihre Kreuzschmerzen verändert haben. Es ist eine einfache Frage, die beim Vorliegen von Kreuzschmerzen eine einfache Antwort auf einer dreistufigen Skala (verbessert, gleich geblieben, verschlechtert) verlangt. Die Analyse dieses subjektiven Erfolgsmaßes ergab, dass 46,1% ihre Kreuzschmerzen als „gebessert“, 44,7% als „nicht verändert“ und 9,2% als „verschlechtert“ beurteilten. Dieses Ergebnis verweist auf eine etwas optimistischere Bewertung der medizinischen Rehabilitation bei chronischen Rückenschmerzen. Wenn wir gewillt sind, dieses Ergebnis ernst zu nehmen, dann müssen wir uns auch mit folgendem Ergebnis auseinandersetzen (Abbildung 2): Die Erfolgsquote variiert zwischen den Kliniken in einem beträchtlichen Ausmaß. Die Klinik mit der höchsten Erfolgsrate kann auf über drei Viertel aller Patienten als verbessert verweisen, die Klinik mit dem geringsten Anteil liegt bei 17%. Die statistische Berücksichtigung

unterschiedlicher Patienteneigenschaften zwischen den Kliniken, die so genannte Fall-Mix-Adjustierung, kann diese Unterschiede übrigens nur zum kleinen Teil erklären (Meyer et al., 2009).

Ebenso konnte aufgezeigt werden, dass die Wahrscheinlichkeit für eine erfolgreiche Rehabilitation entschei-

**Abbildung 2: Kliniken im Vergleich: Anteil von Rehabilitanden mit verbesserten Kreuzschmerzen (gegenüber dem Anteil von Rehabilitanden mit unveränderten bzw. verschlechterten Kreuzschmerzen) pro Klinik (k=210) mit 95%-Konfidenzintervall (vgl. dazu Meyer et al., 2009)**



dend mit der vorangegangenen Dauer der Arbeitsunfähigkeitszeiten, einer Rentenanspruchsstellung, dem Versicherungszweig (Angestellten vs. Arbeiterrentenversicherung),

sowie unabhängig davon auch der Schulbildung zusammenhängt. Ein Rehabilitand mit Abitur hat, unabhängig von seiner sozialmedizinischen Problemlage (!), eine um den Faktor 1,8 höhere Wahrscheinlichkeit für eine Verbesserung seiner Kreuzschmerzen (Mittag et al., 2006). Zum Vergleich: Der Umfang der Bewegungstherapie, dokumentiert mithilfe der Klassifikation Therapeutischer Leistungen, erreicht einen Faktor von 1,4.

## 4. Resümee

Ich möchte folgende Punkte am Ende festhalten: Subjektive Erfolgsmaße bilden eine eigenwertige Perspektive auf Rehabilitationserfolg ab. Sie können genauso gut – oder genauso schlecht – gemessen werden wie andere Erfolgsmaße (Outcomes) auch. Sie haben allerdings den großen Vorteil, dass sich über die Fragebogenmethode eine große, repräsentative Anzahl von Rehabilitanden effizient befragen lässt. Und: Subjektive Erfolgsmaße sind in dem Ausmaß bedeutsam, in dem wir ihnen Konsequenzen zugestehen.

## 5. Literatur

- Egner, U., Gerwin, H. & Buschmann-Steinhage, R.** (2006): Stand der Qualitätssicherung in der Rehabilitation der gesetzlichen Rentenversicherung. *Die Rehabilitation*, 45, 221-231
- Grande, G.** (1998): Gesundheitsmessungen. Unterschiedliche Gültigkeitsbereiche von Arzt- und Patientenangaben. Heidelberg: Asanger.
- Heidrich, J., Liese, A., Löwel, H. & Keil, U.** (2002): Self-rated health and its relation to all-cause and cardiovascular mortality in Southern Germany. Results from

the MONICA Augsburg Cohort Study 1984-1995. *Ann. Epidemiol.*, 12, 338-345.

**Hüppe, A. & Raspe, H.** (2005): Zur Wirksamkeit von stationärer medizinischer Rehabilitation in Deutschland bei chronischen Rückenschmerzen: Aktualisierung und methodenkritische Diskussion einer Literaturübersicht. *Die Rehabilitation*, 44, 1-10.

**Idler, E. & Benyami, Y.** (1997): Self rated health and mortality: a review of twenty-seven community studies. *Journal of Health and Social Behavior*, 38, 21-37.

**Kohlmann, T. & Raspe, H.** (1998): Zur Messung patientennaher Erfolgskriterien in der medizinischen Rehabilitation: Wie gut stimmen „indirekte“ und „direkte“ Methoden der Veränderungsmessung überein? *Die Rehabilitation*, 37 (Suppl. 1), 24-31.

**Lienert, G. & Raatz, U.** (1995): Testaufbau und Testanalyse (5. Aufl.). Weinheim: Beltz.

**Meyer, T.** (2004): Vorstellungen schizophrener Menschen über Lebensqualität. Hamburg: Kovac-Verlag.

**Meyer, T., Brüggemann, S. & Widera, T.** (2008): Rehabilitation braucht Zielorientierung- aktuelle Lehren aus der Rehabilitandenbefragung. *RVaktuell*, 3, 94-99.

**Meyer, T., Pohontsch, N., Maurischat, C. & Raspe, H.** (2008): Patientenzufriedenheit und Zielorientierung in der Rehabilitation. Reihe: Rehabilitationsforschung. Lage: Jacobs-Verlag.

**Meyer, T., Pohontsch, N. & Raspe, H.** (2009): Zielfestlegungen in der stationären somatischen Rehabilitation – die Herausforderung bleibt. *Die Rehabilitation*, 48, 128-134.

**Meyer, T., Pohontsch, N., Maurischat, C. & Raspe, H.**

(2009): Abschlussbericht zum Projekt „Analyse der Wirksamkeit der stationären medizinischen Rehabilitation bei Patienten mit chronischen Rückenschmerzen“ für die Deutsche Rentenversicherung Bund.

**Meyer, T., Pohontsch, N., Maurischat, C. & Raspe, H.**

(2009): Warum beurteilen Rehabilitanden die Rehaplanung und Zielorientierung weniger positiv als andere Aspekte der medizinischen Rehabilitation? Physikalische Medizin, Rehabilitationsmedizin, Kurortmedizin / Journal of Physical and Rehabilitation Medicine, 19, 85-92.

**Meyer, T., Pohontsch, N. & Raspe, H.** (zur Publ. angen.): Zielfestlegungen in der stationären somatischen Rehabilitation – die Herausforderung bleibt. Die Rehabilitation.

**Mittag, O. & Raspe, H.** (2003): Eine kurze Skala zur Messung der subjektiven Prognose der Erwerbstätigkeit: Ergebnisse einer Untersuchung an 4279 Mitgliedern der Gesetzlichen Arbeiterrentenversicherung zur Reliabilität (Guttman-Skalierung) und Validität der Skala. Die Rehabilitation, 42, 169-174.

**Mittag, O., Meyer, T., Glaser-Möller, N., Matthis, C. & Raspe, H.** (2006): Vorhersage der Erwerbstätigkeit in einer Bevölkerungsstichprobe von 4.225 Versicherten der LVA über einen Prognosezeitraum von fünf Jahren mittels einer kurzen Skala (SPE-Skala). Gesundheitswesen, 68, 294-302.

**Norman, G. R., Sloan, J. A. & Wywich, K. W.** (2003): Interpretation of changes in health-related quality of life. The remarkable universality of half a standard deviation. Med Care, 41, 582–592.

**Schwartz, C. D. & Rapkin, B. D.** (2004): Reconsidering the psychometrics of quality of life assessment in light of response shift and appraisal. Health and Quality of Life Outcomes, 2, 16.

**U.S. Department of Health and Human Services, Food and Drug Administration et al.** (2006): Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims.

# Adaptives Assessment: Perspektiven für die Ergebnisqualität

**Prof. Dr. Markus Wirtz<sup>1</sup>**

Pädagogische Hochschule Freiburg  
Abteilung für Forschungsmethoden

## Inhaltsverzeichnis

<b>1. Zusammenfassung</b>	<b>133</b>
<b>2. Einleitung</b>	<b>133</b>
<b>3. Die Eigenschaften der Item-Response-Theorie (IRT)</b>	<b>135</b>
<b>4. Die Entwicklung eines adaptiven Tests</b>	<b>138</b>
<b>5. Das Vorgehen beim adaptiven Testen</b>	<b>147</b>
<b>6. Zusammenfassung und Ausblick</b>	<b>152</b>
<b>7. Literatur</b>	<b>152</b>

---

<sup>1</sup> Unter Mitwirkung von Ch. Frey, E. Prinz & E. Müller, Abteilung für Forschungsmethoden, Pädagogische Hochschule Freiburg; Dr. T. Forkmann & Dr. M. Böcker, Institut für Medizinische Psychologie und Medizinische Soziologie, Universitätsklinikum der RWTH Aachen.

# 1. Zusammenfassung

Ergebnismessungen sind in den meisten Einrichtungen der Akut- und Rehabilitationsversorgung sehr umfangreich, insbesondere wenn sie den Anspruch haben, möglichst alle für die Patienten relevanten Bereiche abzudecken. Dies bedeutet sowohl für Patienten als auch für Personal einen hohen zeitlichen Aufwand. Das adaptive Testen bietet hier neue Möglichkeiten bei der Ergebnisqualitätsmessung. Unter dem adaptiven Testen versteht man ein Verfahren, bei dem sich die Auswahl der Fragen, die Patienten gestellt wird, an Eigenschaften des Patienten selbst orientiert. Das bedeutet: Der Test wird auf die Situation des Patienten zugeschnitten. Umgesetzt wird dies meistens im Rahmen eines Computer-adaptiven Test (CAT).

Adaptives Testen unterscheidet sich von klassischen Assessmentinstrumenten vor allem in seinen methodischen Voraussetzungen und den daraus folgenden Implikationen. So kann ein gut umgesetzter adaptiver Test mit nur einem Bruchteil der Fragen eine Dimension mit ähnlicher oder gar deutlich besserer Messgüte wie in einem klassischen Testverfahren erheben.

Im Folgenden werden die messtheoretischen Grundlagen, die Voraussetzungen und Umsetzungen adaptiver Tests dargestellt.

## 2. Einleitung

Was ist adaptives Testen?

### Definition

Von adaptivem Testen spricht man dann, wenn sich die Auswahl des Items, welches bei einer Messung einer Person vorgelegt werden, an dem Antwortverhalten

dieser Person orientiert. Bei der Auswahl eines neuen Items wird das bereits gezeigte Antwortverhalten berücksichtigt.

Beim adaptiven Testen geht es also darum, Personen Items vorzulegen, die sich an den Kenntnissen oder Fähigkeiten der Personen orientieren. Ziel ist es, Items zur Beantwortung vorzugeben, die auf die Person zugeschnitten sind (Frey, 2007).

Das adaptive Testen ist in einigen Bereichen der Merkmalsmessung schon eingesetzt worden oder wird aktuell umgesetzt. Im deutschen Sprachraum wurden beispielweise adaptive Testmodelle zur Messung von Angst (Becker et al., 2002) und zur Erfassung von Stresserleben (Kocalevent, 2005) entwickelt, international ist hier das PROMIS-Projekt des National Institutes of Health (NIH) zu nennen. Im Rahmen des PROMIS-Projektes (z.B. DeWalt et al., 2007; Reeve et al., 2007; Hays et al., 2007) wurden adaptive Testprozeduren entwickelt, die zum Ziel haben, für Patienten mit chronischen Erkrankungen die Bereiche psychische, physische und soziale Gesundheit abzudecken.

Das adaptive Testen wird in dem vorliegenden Beitrag am Beispiel der Umsetzung in dem Projekt RehaCAT<sup>2</sup> für den Bereich der orthopädischen Rehabilitation zur

Erfassung von Funktionsfähigkeit im Alltag, im Beruf und der Depression dargestellt. Das Projekt RehaCAT wird aktuell an der PH Freiburg in Kooperation mit der Universität Freiburg, Institut für Psychologie und mit der RWTH Aachen umgesetzt.

---

<sup>2</sup> Wirtz, M., Bengel, J. & Gauggel, S.. Entwicklung und Validierung eines adaptiven Assessments für Patienten in der orthopädischen Rehabilitation – RehaCAT. Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Förderkennzeichen WI 3210/2-1.

### 3. Die Eigenschaften der Item-Response-Theorie (IRT)

Um Aussagen über die Leistungsfähigkeit einer Person im Rahmen der klassischen Testtheorie (KTT) machen zu können, ist es erforderlich, dass jedes Individuum alle Items bearbeitet, die zur Erfassung der Leistungsfähigkeit vorgesehen sind. Auch nur über den daraus errechenbaren Score (z.B. einer Skala) können die Leistungen verschiedener Personen miteinander verglichen werden. Der Summenwert einer Skala der KTT kann dabei durch verschiedene Antwortmuster zustande gekommen sein (siehe grünes und rotes Antwortmuster in Abbildung 1). Während die ersten drei Items das grün markierte Belastungsniveau als geringer ausgeprägt ausweisen, zeigt das letzte Item hier eine geringere Belastung an. Beide Patientenprofile in Abbildung 1 würden hier also denselben Summenwert ergeben.

**Abbildung 1: Auswirkungen zweier Antwortmuster auf den Summenwert**

Bitte kreuzen Sie an, in welchem Maße Sie in den letzten 4 Wochen Schwierigkeiten hatten, folgendes zu tun	(1) Unmöglich	(2) starke Schwierigk.	(3) mäßige Schwierigk.	(4) geringe Schwierigk.	(5) keine Schwierigk.
Eine halbe Stunde körperlich arbeiten (wie z. B. Staub saugen, Rasen mähen, etc.)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Einen schweren Gegenstand über 5 Meter tragen	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
30 Minuten ohne Unterbrechung stehen (z. B. in einer Warteschlange)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zu einer kulturellen Veranstaltung oder einem Fest gehen	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Das letzte Item erweist sich in empirischen Untersuchung als wenig testfair, da es von Kontexteinflüssen

systematisch beeinflusst wird (Wirtz et al., 2005). Der Summenwert bildet diese inhaltliche Struktur nicht ab und es resultiert eine Kenngröße, die nicht eindeutig im Sinne des Belastungsniveaus des Patienten interpretiert werden kann. Um alle Informationen, die in einem Summenwert stecken auszuschöpfen, müsste somit das Antwortmuster mit berücksichtigt werden.

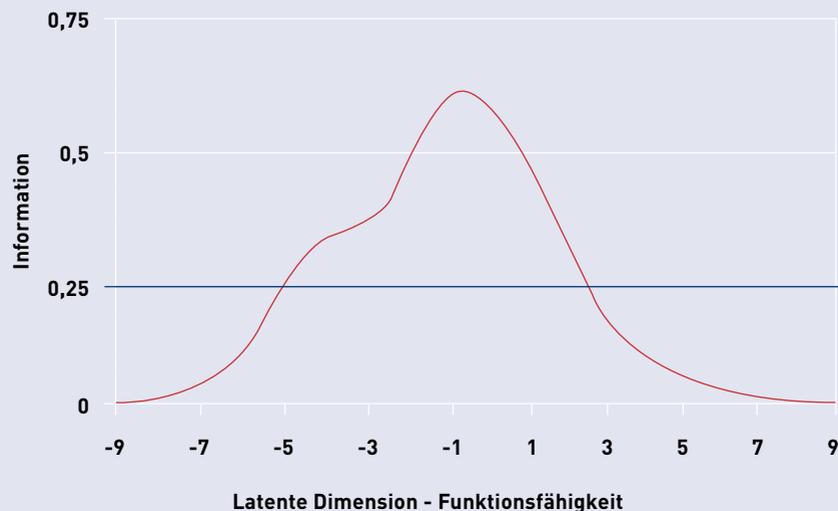
Genügt eine Itemgruppe hingegen den Annahmen des Rasch-Modells, das der Familie der Item Response Modelle (IRT) entstammt, so kann die Ambiguität bei der Interpretation des Summenwerts vermieden werden, da dieser alle Informationen über das Ankreuzverhalten einer Person enthält (Bond & Fox, 2001). Ferner kann jedes einzelne Item für Vergleichszwecke separat herangezogen werden. Es muss nicht die komplette Skala vorgelegt werden, um Aussagen über die Leistungsfähigkeit einer Person machen zu können (Spezifische Objektivität). Durch diese Eigenschaft der IRT wird adaptives Testen möglich. Es können mehrere Items, die dasselbe Merkmal erfassen (Unidimensionalität) zu Diagnostik flexibel ausgewählt werden, ohne Einbußen in den klassischen Gütekriterien für Mess-Skalen hinnehmen zu müssen.

#### 3.1 Reliabilität

In der KTT wird der Reliabilitätskoeffizient aus allen Items eines Tests geschätzt. Dieser Koeffizient bezieht sich auf alle Personen, die den Test ausfüllen, unabhängig davon, welche Leistungsfähigkeit diese Personen aufweisen. Betrachtet man die Reliabilität in der KTT jedoch genauer, wird ersichtlich, dass die Reliabilität nicht in allen Schwierigkeitsbereichen des Tests gleich gut sein kann. Dies hängt in wesentlichem Maße von der Zusammensetzung der Items ab. Meistens befinden sich in einem Test viele Items mit mittlerer Schwierigkeit, und wenige mit hoher oder geringer Schwierigkeit. In diesem Fall ist eigentlich die Reliabilität für einen mittleren Schwierigkeitsbereich höher als für extreme Schwierigkeitsbereiche. Abbildung 2 stellt

diesen Zusammenhang dar. Die rote Linie verdeutlicht, wie bei einer statischen Skala die Information, die die Items bzw. der Summenwert liefern, in den Extrembereichen abnimmt. Personen mit einer sehr hohen bzw. sehr geringen Fähigkeit auf dem zu messenden latenten Merkmal können nur wenig reliabel eingestuft werden, die Information der Items ist zu gering.

**Abbildung 2: Informationskurven für einen statischen (rot) und einen adaptiven (blau) Test**



In der IRT wird dieser Beobachtung Rechnung getragen: Die Reliabilität hängt von der Zusammenstellung der Items ab bzw. von der Fähigkeit der Person, die diese bearbeitet. Dies ist vor allem für adaptives Testen sehr interessant. So kann für jede Kombination von Items ein separater Reliabilitätskoeffizient errechnet werden. Der bestmögliche Reliabilitätskoeffizient ist dann erreichbar, wenn die Schwierigkeit der Items exakt zu der Fähigkeit des Probanden passt (Wright & Stone, 1979). Also wenn z.B. ein Test mit schwierigen Items von einer Person mit hoher Fähigkeit bearbeitet wird. Aus diesen Überlegungen wird auch ersichtlich,

dass ein adaptiver Test, um eine bestmögliche Reliabilität zu erreichen, die Schwierigkeit der Itemvorgaben dem Antwortverhalten des Probanden anpassen muss.

Hat der adaptive Test alle Schwierigkeitsbereiche gleich gut abgedeckt, sind die Informationen, die die Items oder die Skala eines adaptiven Tests liefern, für alle Personenfähigkeiten gleich gut (Abb. 2, blaue Linie). Hieraus ergeben sich insbesondere Vorteile für die Veränderungsmessung, da auch niedrige und hohe Belastungsniveaus reliabel erfasst werden. Außerdem werden Boden- und Deckeneffekte vermieden (Bühner, 2006).

### 3.2 Validität

In der KTT bezieht sich der Validitäts- wie der Reliabilitätskoeffizient auf den gesamten Test bzw. die gesamte Skala. Wird das Instrument verändert (z.B. ein Item kommt hinzu oder eins wird weggelassen) muss die Validität neu ermittelt werden.

Die Validität der IRT wird auf Grundlage der Itembank (Gesamtheit aller Items, die zum Assessment eines Merkmals herangezogen werden kann) errechnet und ist gültig für alle Subsets aus dieser Itembank. Da ein adaptiver Test aus einer Auswahl an Items aus der Itembank besteht, gelten diese Überlegungen insbesondere für diese Assessmentmethode.

## 4. Die Entwicklung eines adaptiven Tests

### Was sind die Voraussetzungen für adaptives Testen?

Die beschriebenen Vorzüge der Anwendung von Skalen, die den Annahmen des Rasch-Modells genügen, stellen die Basis für adaptives Testen dar. Nur Itemgruppen, deren Eigenschaften durch das Rasch-Modell angemessen modelliert werden können, besitzen die Eigenschaft der strikten Eindimensionalität und der spezifischen

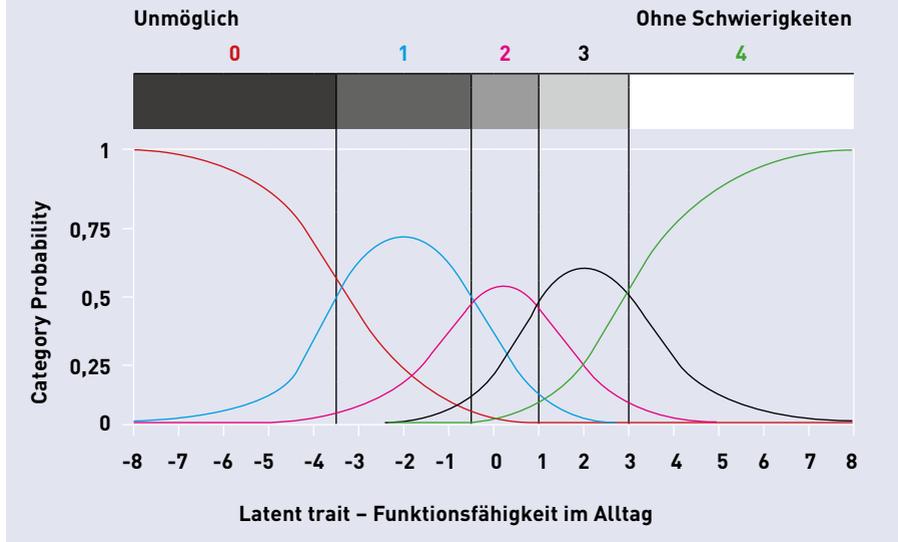
Objektivität. Eindimensionalität bedeutet, dass alle Items desselben Tests dasselbe zugrunde liegende Merkmal erfassen. Spezifische Objektivität bedeutet, dass die diagnostischen Daten trotz unterschiedlicher Itemvorgaben vergleichbare Informationen zu Ausprägungen desselben Merkmals liefern sollen. Auch andere IRT-Modelle können beim adaptiven Testen zur Anwendung kommen. Dann sind allerdings weitere Informationen zu berücksichtigen. Die folgende Darstellung bezieht sich nur auf das Rasch-Modell, das optimale psychometrische Eigenschaften besitzt.

Die Eindimensionalität auf Itemebene erfordert, dass ein monoton ansteigender Zusammenhang zwischen der Fähigkeit einer Person (latente Merkmalsausprägung) und der Wahrscheinlichkeit der Auswahl einer bestimmten Antwortkategorie eines Items (manifestes Merkmal) bestehen muss. Das bedeutet, dass alle Items des Tests dasselbe zugrundeliegende Merkmal erfassen, und diese Items durch keine weiteren Merkmale konfundiert sein sollen.

Des Weiteren muss es einen monoton ansteigenden Zusammenhang geben zwischen der Fähigkeit einer Person (latentes Merkmal) und dem Antwortverhalten auf ein bestimmtes Item hin (manifestes Merkmal). Abbildung 3 stellt die Wahrscheinlichkeiten für die einzelnen Antwortalternativen dar (Ordinate) und verdeutlicht damit diesen Zusammenhang: Je größer die Funktionsfähigkeit im Alltag (Abszisse) ist, desto eher wird die Antwort auf die Frage „Einen schweren Gegenstand über 5m tragen“ in Richtung „ohne Schwierigkeiten“ ausfallen. Hat z.B. eine Person in diesem Beispiel eine Fähigkeit von +2, so sind die Wahrscheinlichkeiten für die Antworten „Unmöglich“ bzw. „0“ (rote Wahrscheinlichkeitsverteilung) und „1“ (blau) sehr gering. Die Antwortoptionen „2“ (rosa) oder „4“ (grün) sind mit 0,22 beide gleich wahrscheinlich. Am ehesten wird diese Person aber die Antwortoption „3“ (schwarz) wählen: hier beträgt die Wahrscheinlichkeit ungefähr 0,55.

Das Antwortverhalten der Probanden hängt von der latenten Fähigkeit ab, je größer die Fähigkeit, desto wahrscheinlicher sind geringe Schwierigkeiten beim Lösen des Items. Und dieser Zusammenhang muss in einer streng monotonen Weise für alle Items einer Itembank eines adaptiven Tests gegeben sein.

**Abbildung 3: Wahrscheinlichkeiten der Auswahl der jeweiligen Antwortoptionen in Abhängigkeit von der Personenfähigkeit für ein einzelnes Item einer Rasch-homogenen Skala**



Neben der geforderten Unidimensionalität bei einer Realisierung im Rahmen des Rasch-Modelles gibt es auch Modelle für multidimensionales adaptives Testen, diese sind aber noch nicht ausreichend elaboriert (Frey, 2007), bieten aber für die Zukunft interessante Erweiterungsmöglichkeiten für das adaptive Testen.

#### Vom Itempool zur Itembank

Um das adaptive Testen zu realisieren, werden Items benötigt, die dem Probanden potentiell vorgelegt werden können. Die Vorgabe der Items orientiert sich an dem Antwortverhalten der Probanden. Da dieses sehr vielfäl-

tig sein kann und meist ein sehr breites Fähigkeitspektrum erfasst werden soll, ist es sinnvoll, viele Items im Hintergrund zu haben.

In einem ersten Schritt ist zu klären, welches Konstrukt erhoben werden soll und wie dieses Konstrukt theoretisch umrissen werden kann. In dem Projekt RehaCAT ist eines der Konstrukte die „Funktionsfähigkeit im Alltag“, die in Anlehnung an die ICF als Ausmaß definiert ist, in dem eine Person mit Behinderung selbständig alltägliche Aktivitäten durchführen kann. Die Aktivitäten werden u.a. in die Bereiche „Mobilität“, „Selbstversorgung“ und „Häusliches Leben“ aufgliedert. Die ICF gibt dabei sehr genau vor, wie die einzelnen Kategorien mit Inhalten zu füllen sind.

Die genauen Inhalte der Items sollte auf Basis einer umfassenden Literaturanalyse und Expertenbefragung erfolgen. Hier kann geklärt werden, welche Instrumente schon vorhanden sind, um Aspekte des interessierenden Konstruktes zu erfassen. Dieser wichtige Schritt kann ebenfalls noch die Definition des Konstruktes bereichern bzw. erweitern.

Da sich schon einige Assessmentinstrumente mit dem Thema „Funktionsfähigkeit im Alltag“ beschäftigt haben, können diese Instrumente einen Beitrag für den Itempool leisten: Die Items aus anderen Tests wurden angepasst (übersetzt, umformuliert, das Antwortformat neu festgelegt) und in die Sammlung der Items aufgenommen. Da die Items der bestehenden Testverfahren nicht alle Bereiche des interessierenden Merkmals abdecken, wurden weitere Items hinzugefügt, um den Bereich „Funktionsfähigkeit im Alltag“ inhaltsvalide abzudecken. Der Itempool besteht aus all den gesammelten Items, die laut Literatur und eigener Überlegungen zu Erfassung des interessierenden Bereiches nötig sind. Ein besonderes Augenmerk bei der Zusammenstellung der Items ist auf die gleichmäßige Verteilung der Schwierigkeiten zu legen: Es sollen sowohl Items vorhanden sein, die auch mit geringen

Fähigkeiten auf Seiten der Probanden erfüllt werden können und solche, die nur von wenigen Probanden mit besonders guten Fähigkeiten erfüllt werden können.

Um von dem bestehenden, idealerweise breit gefüllten Itempool zu einer gut ausgewählten Itembank zu gelangen, sind mehrere Schritte notwendig. Im Rahmen einer qualitativen Itemanalyse empfiehlt es sich zum einen, Experten zu befragen, welche Items für das betreffende Gebiet relevant sind (Augenscheinvalidität), zum anderen sollte auch schon in einem sehr frühen Stadium der Itemselektion Verständlichkeit und Sinnhaftigkeit der Items an einer kleinen Stichprobe der Zielpopulation vorgetestet werden. Durch Experten- und Betroffenenbefragung können schon erste Hinweise auf Verbesserungen oder Veränderungen von Items des Pools erhalten werden.

Der wichtigste Schritt, um eine Itembank mit allen Voraussetzungen für adaptives Testen zu erhalten, ist die Bearbeitung aller Items durch eine größere Stichprobe. Ziel dieser Befragung ist es, Items zu identifizieren, die die Voraussetzungen für ein adaptives Assessment erfüllen (Rasch-Homogenität), und die empirischen Schwierigkeitsschätzungen für jedes Item zu ermitteln.

### **Die Schwierigkeitsindices**

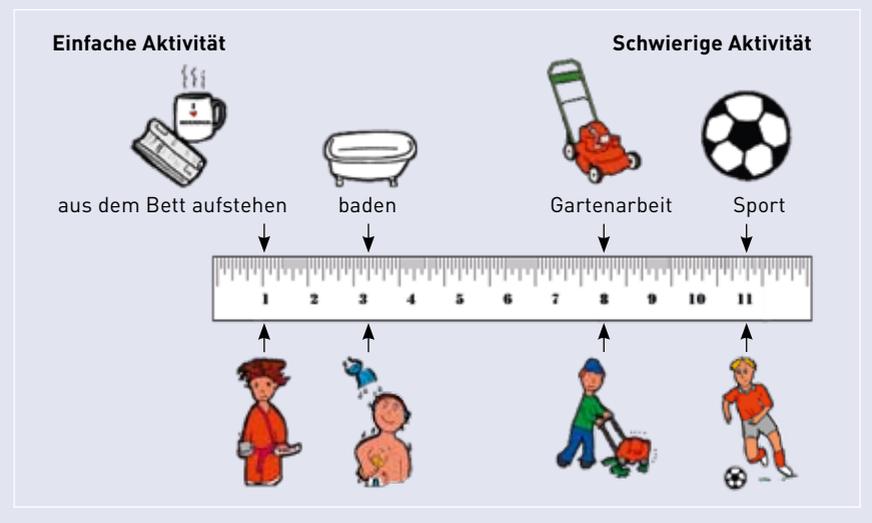
Die IRT geht davon aus, dass Personen, die eine bestimmte Fähigkeit haben mit hoher Wahrscheinlichkeit alle Items lösen können, die Schwierigkeiten aufweisen, die unterhalb dieses Fähigkeitsniveaus liegen. Je weiter die Schwierigkeit eines Items unterhalb dieser Fähigkeit liegt, desto größer ist die Wahrscheinlichkeit. Analog verhält es sich bei Schwierigkeiten, die über den Fähigkeiten der Probanden liegen. Je weiter die Schwierigkeit über der Fähigkeit liegt, desto niedriger die Wahrscheinlichkeit. Die Information, die ein Item über die Fähigkeit einer Person liefert, ist nun abhängig von der Passung zwischen Schwierigkeit und Fähigkeit. Bei

einer fehlenden Passung ist der Zugewinn an Information eher gering: Eine Person mit einer geringen Fähigkeit kann ein sehr schweres Item wahrscheinlich nicht lösen, diese Information war zu erwarten und liefert keine neuen Erkenntnisse.

Bei einem adaptiven Test kommt es zu einer möglichst hohen Passung zwischen der Schwierigkeit eines Items und der Fähigkeit (bzw. der aufgrund der vorangegangenen Antworten geschätzten Fähigkeit) einer Person. Dies liefert den maximal möglichen Informationsgewinn über die tatsächliche Fähigkeit.

Sind für alle Items, die die Voraussetzungen für adaptives Testen erfüllen, Schwierigkeitsindices empirisch ermittelt worden, so lassen sie sich an einem Kontinuum mit den Polen „einfache Aktivität“ und „schwierige Aktivität“ anordnen (siehe Abbildung 4).

**Abbildung 4: Beispielhafte Ordnung von Items aus dem Bereich „Funktionsfähigkeit im Alltag“ nach Schwierigkeiten (Quelle: Mallinson, 2001)**



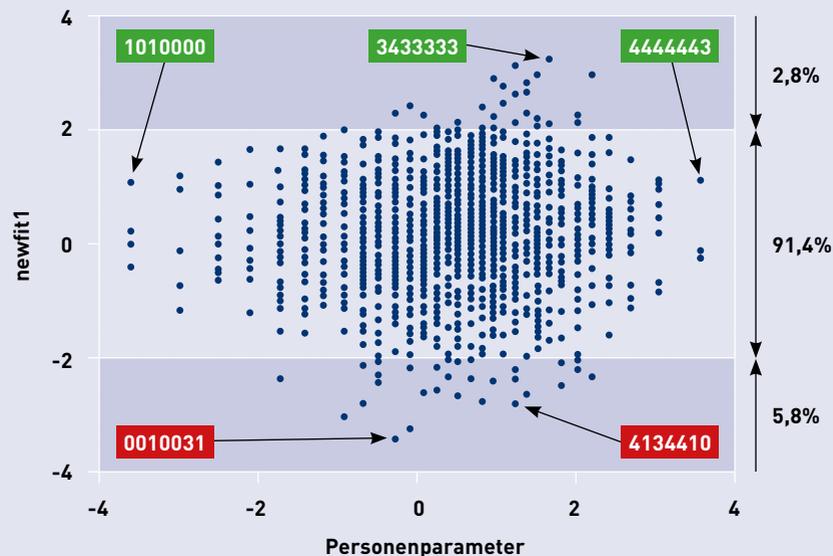
Nicht nur die Tätigkeiten (unten in Abbildung 4), sondern auch die Personen, können auf der so ermittelten Merkmalsdimension eingeordnet werden. Jeder Patient kann also genau durch diejenigen Tätigkeiten charakterisiert werden, die er noch bewältigen kann. Eine steigende Funktionsfähigkeit drückt sich also in der besseren Ausführbarkeit schwerer Tätigkeiten aus.

In dem Projekt RehaCAT sind dies z.B. Items des Bereiches „Funktionsfähigkeit im Alltag“ für Patienten der orthopädischen Rehabilitation. Es wird davon ausgegangen, dass sich das Antwortverhalten innerhalb dieser Personengruppe nicht unterscheidet. Tätigkeiten, die für orthopädische Patienten leicht zu bewältigen sind, sollten von vielen Patienten gelöst werden. Items, die Tätigkeiten beschreiben, die von diesen Patienten nur sehr schwer durchgeführt werden können, sind auch nur von wenigen zu lösen. Diese Beobachtung sollte sich in den Antwortmustern der einzelnen Personen wiederfinden. Ordnet man Items nach Schwierigkeit (leichte zuerst, das schwerste zuletzt), so können sich die in Tabelle 1 dargestellten Muster ergeben (bei Antwortoptionen von „0“ bzw. „unmöglich“ bis zu „4“ bzw. „ohne Schwierigkeiten“):

**Tabelle 1: Antwortmuster von Personen mit unterschiedlicher Einschränkung bei leichten bis schweren Items**

Patient	Leichtes Item	Eher leichtes Item	Eher schweres Item	Schweres Item
Modellkonformes Antwortmuster				
Stark eingeschränkt	1	1	0	0
eingeschränkt	3	2	3	1
Wenig eingeschränkt	4	4	3	3
Kein modellkonformes Antwortmuster				
Stark eingeschränkt	1	0	0	3
eingeschränkt	1	1	2	4
Wenig eingeschränkt	2	3	4	4

**Abbildung 5: Scatterplot mit modellkonformem [grüne Kästchen] und nicht modellkonformem Antwortverhalten [rot]. (Abszisse: Funktionsfähigkeit des Patienten; Ordinate: Modellpassung; Quelle: Wirtz & Böcker, 2006)**



So können einzelne Personen (Tabelle 1, Abbildung 5) oder Personengruppen (Abbildung 6 und 7) identifiziert werden, die ein Antwortverhalten zeigen, das nicht den Vorgaben des Modells entspricht.

**Abbildung 6: Beispielitems aus dem Bereich „Funktionsfähigkeit im Alltag“ zur Verdeutlichung unterschiedlichen Antwortverhaltens von Gruppen**

	Momentan kann ich folgende Aktivität ohne fremde Hilfe durchführen ...	ohne Schwierigkeiten	mit geringen Schwierigkeiten	mit mittleren Schwierigkeiten	mit großen Schwierigkeiten	unmöglich
<b>Item 1</b>	Einen leichten Gegenstand vom Boden anheben, z.B. einen gefüllten Papierkorb	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>Item 2</b>	Einen leichten Gegenstand über 5 Meter tragen, z.B. einen gefüllten Papierkorb	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

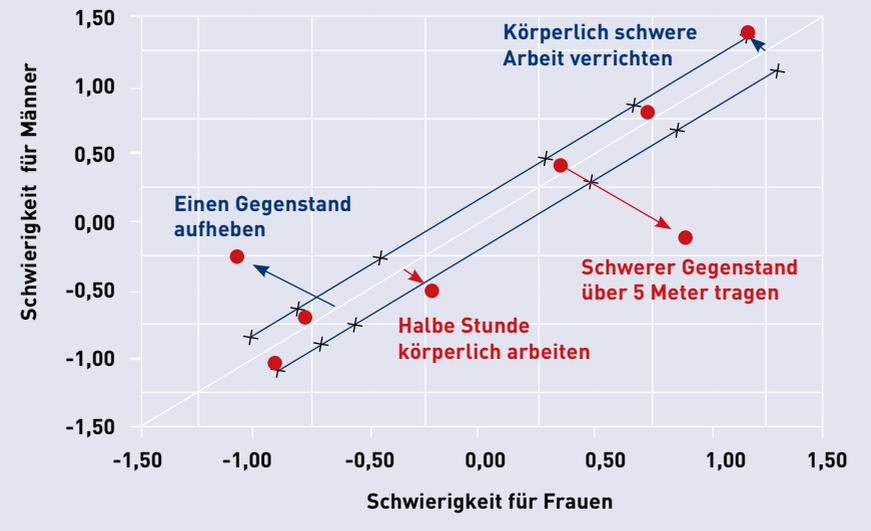
Personengruppen können auf Items unterschiedlich reagieren. Beispielhaft können hier zwei Items aus dem Bereich Funktionsfähigkeit im Alltag aufgeführt werden, um dies zu verdeutlichen:

Für Personen mit einer Hüft-Endoprothese ist es meist schwieriger, Gegenstände vom Boden aufzuheben als Gegenstände zu tragen. So handelt es sich für diesen Personenkreis bei Item 1 um ein schwierigeres Item als bei Item 2. Bei Item 1 werden mehr Personen „große Schwierigkeiten“ bzw. „unmöglich“ angeben.

Für Patienten mit anderen orthopädischen Erkrankungen (z.B. nach einer Knie-OP) kann die Ordnung der Schwierigkeiten dieser beiden Items anders aussehen: Hier ist z.B. das Anheben eines Gegenstandes kein großes Problem, das Transportieren über eine kurze Strecke dagegen schon. Dieses Analyseergebnis kann anzeigen, dass u.U. unterschiedliche Problemstrukturen erfasst werden, oder aber die Definition des erfassten Konstruktes zwischen Patientenpopulationen systematisch variiert.

Abbildung 7 zeigt den Vergleich von Schwierigkeiten einzelner Items bei Frauen und Männern. Beide Geschlechter zeigen ein unterschiedliches Antwortverhalten: Das Item „schwere Gegenstände über 5m tragen“ weist für Frauen eine unerwartet hohe (modellinkompatible) Schwierigkeit im Vergleich zu den Männern auf, genau umgekehrt verhält es sich bei dem Item „einen Gegenstand aufheben“.

Abbildung 7: Vergleich der Schwierigkeiten einzelner Items für Frauen und Männer (Quelle: Wirtz & Böcker, 2007)



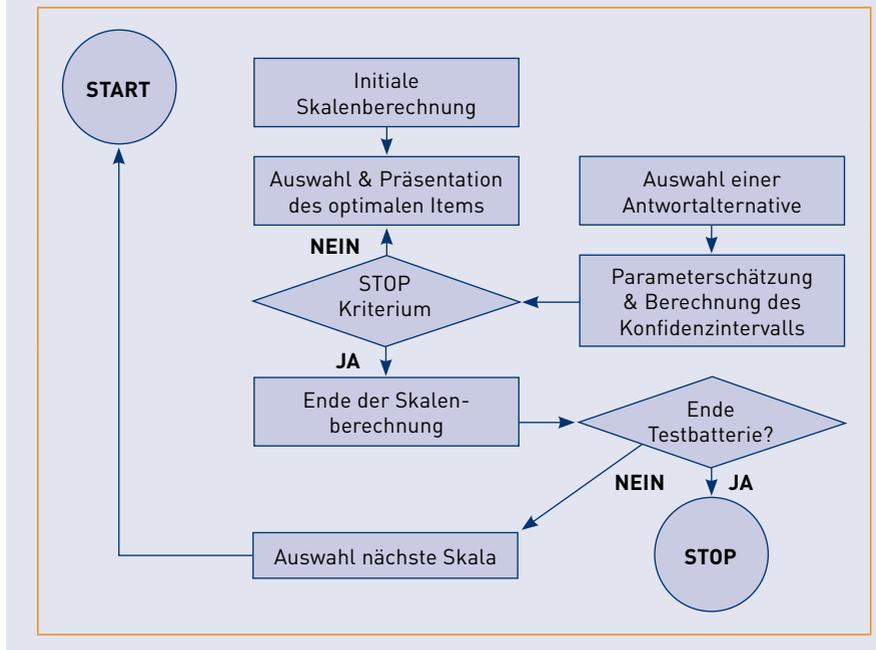
Dies wirft die Frage der Fairness von (adaptiven) Tests auf. Um fair zu sein, müssen Tests für unterschiedliche Personen vergleichbar sein. Auf Gruppenebene muss sichergestellt werden, dass für unterschiedliche Teilgruppen die Items dasselbe Merkmal erfassen. Abbildung 7 zeigt, dass dies im Rahmen einer Rasch-Modellierung anschaulich überprüft werden kann.

## 5. Das Vorgehen beim adaptiven Testen

Ein optimale adaptive Testprozedur erfordert eine gute Passung der Schwierigkeitsindizes der Items und der Fähigkeitsausprägungen der zu testenden Patienten. In allen Schritten des adaptiven Tests liegt das Hauptaugenmerk darauf, wie die Fähigkeit der Person am besten geschätzt werden kann. Es muss demnach ein adaptiver Algorithmus spezifiziert werden (vgl. Abbildung 8), der die Itemauswahl zu Beginn und während

des Tests spezifiziert und der genau festlegt, unter welchen Bedingungen ein Test beendet wird.

Abbildung 8: Algorithmus für die Itemauswahl eines adaptiven Tests



### 5.1 Beginn der Testung

Die erste Frage, die ein Proband im Rahmen des adaptiven Tests präsentiert bekommt, soll, sofern keine weiteren Informationen über die Person vorliegen, eine mittlere Schwierigkeit aufweisen. Das bedeutet, dass in diesem Fall das erste Item zufällig aus denen mit mittlerem Schwierigkeitsindex ausgewählt wird.

Meistens gibt es aber schon vorab weitere Informationen über die Person, die den Test bearbeitet. Diese weiteren Informationen helfen, die Fähigkeit der Person besser schätzen zu können. Der Proband bekommt dann idealerweise genau das Item vorgelegt, das der durchschnittlichen Fähigkeit der genauer definierten Subpo-

pulation (z.B. mit entsprechendem Alter und Geschlecht) entspricht. Als Beispiel für ein solches Vorgehen sei hier ein kurzer Vortest erwähnt, durch den die Einstufung des Probanden in eine von mehreren vorher definierten Leistungsstufen geschieht (Frey, 2007). Eine wichtige Information, die entscheidend für die Einschätzung der Fähigkeit ist, ist das Indikationsgebiet des Probanden. Wie oben bereits erläutert, hängt z.B. die Fähigkeit, einen Gegenstand anzuheben davon ab, ob eine Person mit orthopädischer Indikation hauptsächlich Wirbelsäulen-, Hüft- oder Knieprobleme hat.

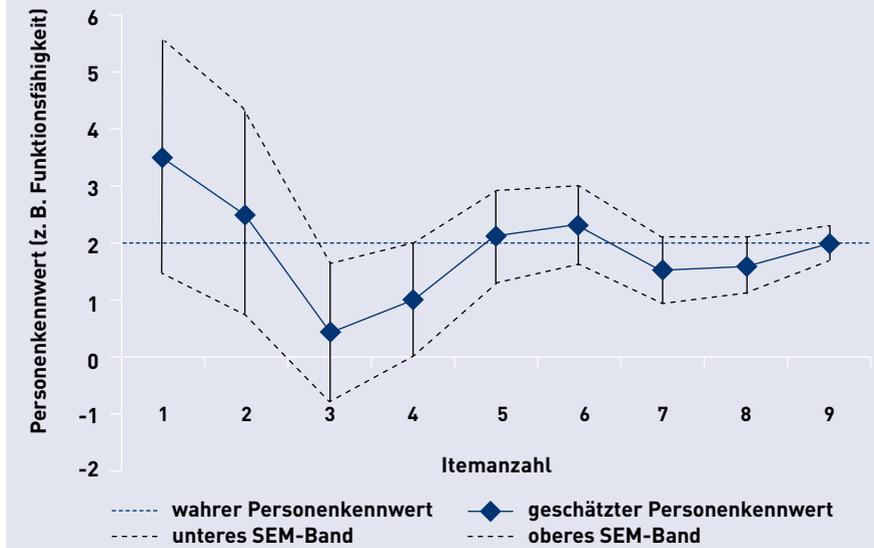
## 5.2 Verlauf der Testung

Das entscheidende Kriterium für die Itemauswahl während der Testung ist die Schätzung der Fähigkeit der Person auf der Grundlage der Angaben des Probanden in schon vorgelegten und beantworteten Items. Allgemein kann das dahinterliegende Rational folgendermaßen beschrieben werden:

- Kann ein Proband eine Aufgabe nicht lösen (weil seine Einschränkungen zu groß sind), ist das nächste Item weniger schwierig,
- Kann ein Proband eine Aufgabe lösen, so wird als nächstes ein Item präsentiert, das eine höhere Schwierigkeit aufweist.

Ziel ist es dabei, sich der wahren Fähigkeit der Person bzw. dem wahren Personenparameter immer mehr zu nähern (siehe Abbildung 9). Die Abbildung zeigt die Schwierigkeiten der sukzessive ausgewählten Testitems und den aufgrund des bisherigen Antwortverhaltens des Patienten geschätzten Standardfehler der Messung (SEM). Die Itemschwierigkeit konvergiert gegen den zugrunde liegenden Fähigkeitsparameter des Patienten, und die Messung wird nach jeder Patientenangabe präziser.

Abbildung 9: Annähern der Schätzung des Personenparameters an den tatsächlichen Wert



## 5.3 Ende der Testung

Wird eine spezifizierter, hinreichend kleiner Standardfehler unterschritten, so liegt eine hinreichend reliable Schätzung vor und die Testprozedur wird beendet. Über dieses gebräuchlichste Kriterium hinaus können auch weitere Bedingungen für einen Testabbruch definiert werden. Dies kann sich z.B. an der Itemanzahl bzw. an der Testzeit orientieren. Insgesamt kann das Vorgehen beim adaptiven Testen wie folgt zusammengefasst werden (vgl. Abbildung 9):

- das erste Item liegt mit seiner Schwierigkeit in einem mittleren Bereich (wenn über die Person keine genaueren Informationen vorliegen)
- kann der Proband ein Item nicht lösen, wird ein leichteres, kann er es lösen, ein schwereres Item vorgelegt
- gleichzeitig wird auf Grundlage der Antworten der Person ein Maß für die Personenfähigkeit (bzw.

→ Einschränkung) und deren Schätzgüte berechnet die letzten beiden Schritte werden so oft wiederholt, bis es zu einer Schätzung des Personenmerkmals kommt, das ein definiertes Gütekriterium erreicht.

Abbildung 10 veranschaulicht das in Abbildung 9 dargestellte iterative Vorgehen nochmals für die Itemschwierigkeiten.

Gestartet wird mit einem Item mittlerer Schwierigkeit, dies kann von dem Patienten aufgrund seiner Einschränkungen nicht gelöst werden. Das zweite Item ist ebenfalls zu schwierig, das dritte kann gelöst werden. Die Schwierigkeit des Items liegt unter der Fähigkeit des Patienten. Das nächste Item (Item 4) ist schwieriger als Item 3 aber leichter als Item 2 und kann ebenfalls gelöst werden. Bei Item 5 ist die Passung zwischen Personenfähigkeit und Itemschwierigkeit dann so gut, dass die Testung hier beendet wird, wenn zudem die Messpräzision zufrieden stellend ist.

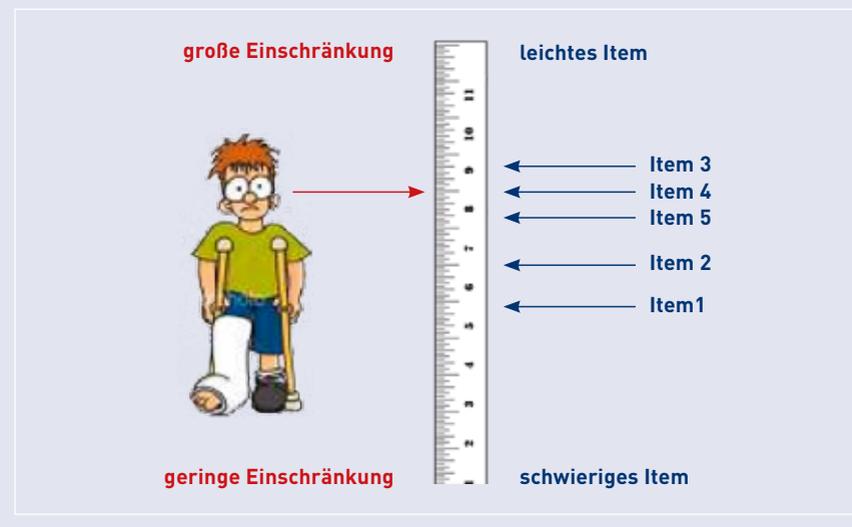
## 6. Zusammenfassung und Ausblick

Ziel dieses Beitrags war es, die grundlegenden Prinzipien des adaptiven Testens im Hinblick auf dessen Anwendung in der Diagnostik in der Gesundheitswissenschaft zu verdeutlichen. Grundlage für das adaptive Testen bilden empirisch ermittelte eindimensionale Itembanken, die das zu erfassende Konstrukt (im Beitrag: ‚Funktionsfähigkeit im Alltag‘) in allen empirisch relevanten Ausprägungsgraden differenziert erfassbar machen. Alle Testitems müssen den Annahmen des gewählten IRT-Modells genügen (z.B. Rasch-Modell). Zudem müssen subpopulationenspezifische psychometrische Eigenschaften identifiziert und ggf. bei der Testung berücksichtigt werden. Das adaptive Testen erfolgt dann durch die sukzessive Auswahl von Items, die die bereits vorliegenden Informationen über den Zustand des Patienten optimal berücksichtigen. Adaptive Testprozeduren ermöglichen somit eine sehr ökonomische Diagnostik, die insbesondere in den Extrembereichen des Merkmalspektrums eine höhere Messpräzision als klassische Erhebungsinstrumente gewährleistet. Hierdurch wird auch für Evaluationsfragen aufgrund der verbesserten Sensitivität für Veränderungen eine viel versprechende Grundlage geschaffen. Derzeit werden die Eigenschaften des adaptiven Testens für RehaPATIENTEN im Projekt RehaCAT im deutschsprachigen Raum erstmals umfassend untersucht.

## 7. Literatur

Becker, J., Walter, O. B., Fliege, H., Bjorner, J., Ravens-Sieberer, U., Walter, M., Klapp, B. F. & Rose, M. (2002): Using the item response theory to develop a computer adaptive test for anxiety. Quality of Life Research, Volume 11, Number 7, 670.

Abbildung 10: Annäherung der Itemschwierigkeiten an die Einschränkung der Person



**Bond, T. G. & Fox, C. M.** (2001): Applying the Rasch model: Fundamental measurement in the human sciences. Mahwah, New Jersey: Lawrence Erlbaum.

**Bühner, M.** (2006): Einführung in die Test- und Fragebogenkonstruktion. München: Pearson

**DeWalt, D. A., Rothrock N., Yount S. & Stone A. A.** (2007): Evaluation of item candidates: the PROMIS qualitative item review. *Medical care*, 45 (5 Suppl 1), 12-21.

**Fliege, H., Becker, J., Walter, O.B., Bjorner, J.B., Klapp, B.F. & Rose, M.** (2005): Development of a Computer-adaptive Test for Depression (D-CAT). *Quality of Life Research*, 14, 2277-2291.

**Frey, A.** (2007): Adaptives Testen. In: Moosbrugger, H. & Keleva, A., *Testtheorie und Fragebogenkonstruktion*, Berlin: Springer-Verlag, 261-278.

**Hays, R. D., Liu, H., Spitzer, K. & Cella, D.** (2007): Item Response Theory Analyses of Physical Functioning Items in the Medical Outcomes Study. *Medical care*, 45 (5) Suppl. 1, 32-38.

**Kocalevent, R. D.** (2007): Stress-CAT - Entwicklung eines computeradaptiven Tests zur Erfassung von Stresserleben. Saarbrücken: VDM Verlag Dr. Müller.

**Mallinson, T.** (2001): Measuring the Impact of Fatigue on Everyday Activities during Chemotherapy. Paper presented at COMET, Chicago.

**Reeve, B. B. et al.** (2007): Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical care*, 45 (5 Suppl 1), 22-31.

**Wirtz, M. & Böcker, M.** (2007): Das Rasch-Modell – Eigenschaften und Nutzen für die diagnostische Praxis. *Die Rehabilitation*, 46 (4), 238-245.

**Wirtz, M., Farin, E., Bengel, J., Jäckel, W., Hämmerer, D. & Gerdes, K.** (2005): IRES-24 Patientenfragebogen – Entwicklung der Kurzform eines Assessmentinstrumentes in der Rehabilitation mittels der Mixed-Rasch-Analyse. *Diagnostica*, 51 (2), 75-87.

**Wright, B.D. & Stone, M.** (1979): *Best test design*, Chicago: MESA Press.

# Möglichkeiten und Grenzen der Bestimmung der Ergebnisqualität in der Rehabilitation durch Therapeuteneinschätzungen am Beispiel des SF-8

**PD Dr. Holger Schulz<sup>1</sup>**

Universitätsklinikum Hamburg-Eppendorf

Zentrum für Psychosoziale Medizin

Institut und Poliklinik für Medizinische Psychologie

## **Inhaltsverzeichnis**

<b>1. Einleitung</b>	<b>157</b>
<b>2. Methoden</b>	<b>158</b>
<b>3. Ergebnisse</b>	<b>159</b>
<b>4. Diskussion</b>	<b>162</b>
<b>5. Literatur</b>	<b>165</b>
<b>6. Anhang</b>	<b>167</b>

---

<sup>1</sup> Unter Mitwirkung von T. Harfst, S. Andreas, S. Kawski, U. Koch & S. Rabung, Universitätsklinikum Hamburg-Eppendorf, Zentrum für psychosoziale Medizin, Institut und Poliklinik für Medizinische Psychologie.

# 1. Einleitung

Für die Rehabilitation einer Vielzahl körperlicher, aber auch psychischer Erkrankungen stellt die Lebensqualität ein zentrales Kriterium für das Behandlungsergebnis dar: Die Fortschritte in der Diagnostik und Therapie von Erkrankungen, wie z.B. die sehr häufigen, oft lebensbedrohlichen Herz-Kreislauf-, Krebs-, Nieren- und Lebererkrankungen führen zu einer Verlängerung der Krankheitsverläufe und Behandlungsdauern sowie zu einer größeren Anzahl behandlungsfähiger und auch längerfristig überlebender Patienten. Für den Bereich der Qualitätssicherung ist darüber hinaus die zentrale Frage zu klären, inwieweit dieser Outcomebereich in der Routine durch die behandelnden Ärzte und Therapeuten oder durch den Patienten selbst eingeschätzt werden kann.

Eine umfassende Abbildung der Gesundheit, die sich am bio-psycho-sozialen Modell von Engel (1977) orientiert, setzt auch die subjektiven Einschätzungen der Betroffenen im Selbstbericht voraus. Damit ist ein Experte für Gesundheit nicht nur der behandelnde Arzt, sondern ebenso die einzelne Person selbst. In diesem Zusammenhang ist ein Synonym für die subjektive Gesundheit der Ausdruck „gesundheitsbezogene Lebensqualität“. Daraus folgt: Die Lebensqualität als objektive Realität gibt es nicht, Lebensqualität ist vielmehr ein nicht direkt beobachtbares psychologisches Konstrukt. Zur Frage der Dimensionen der gesundheitsbezogenen Lebensqualität kommen Cella et al. (2005) in einer aktuellen Analyse einer größeren Stichprobe von Krebs- und HIV-Patienten zu dem Ergebnis, dass sich bisher vor allem die zwei Dimensionen körperliches und psychisches Wohlbefinden eindeutig empirisch haben nachweisen lassen.

Hieran anknüpfend lassen sich hinsichtlich der Formen der Lebensqualitätsmessung ein- von mehrdimensionalen

Verfahren unterscheiden: Erstere finden sich vor allem in Studien, in den gesundheitsökonomische Fragestellungen im Kontext der Lebensqualität untersucht werden, weil hierfür zumeist ein einziger Ergebniswert benötigt oder bevorzugt wird. Sehr viel häufiger werden allerdings mehrdimensionale Instrumente eingesetzt. Eines der weltweit sehr verbreiteten Instrumente ist der SF-36 Fragebogen (Bullinger & Kirchberger, 1998; Ware, 2000). Der SF-36 ist ein Beispiel für ein krankheitsübergreifendes, d.h. ein sogenanntes generisches Verfahren. Er besteht aus acht Skalen, die sich zu einer körperlichen und einer psychischen Summenskala zusammenfassen lassen.

Schließlich kann hinsichtlich der Erhebungsquellen noch unterschieden werden zwischen Fragebogen zur Selbsteinschätzung durch den Patienten auf der einen Seite sowie Fremdeinschätzungen, vornehmlich durch Ärzte oder Familienangehörige, auf der anderen Seite. Bisher liegen jedoch nur vergleichsweise wenige Studien zum Zusammenhang zwischen Selbst- und Fremdeinschätzungen der gesundheitsbezogenen Lebensqualität vor: Daher sollen nachfolgend empirische Ergebnisse zum Zusammenhang beider Einschätzungsarten präsentiert werden.

# 2. Methoden

Als empirische Grundlage für die Analysen zum Zusammenhang von Selbst- und Fremdeinschätzungen der Lebensqualität dienen Daten, die in der Pilotphase eines Qualitätssicherungsprogramms der Gesetzlichen Krankenversicherung im Indikationsbereich psychische und psychosomatische Erkrankungen in elf Rehabilitations-Fachkliniken erhoben wurden (Kawski & Koch, 2004). Die Stichprobe umfasst 1.812 Patienten, von denen verwertbare Daten zu Behandlungsbeginn und -ende vorliegen.

Für die Erfassung der gesundheitsbezogenen Lebensqualität wurde mit dem Fragebogen SF-8 (Ware, Kosinski, Dewey & Gandek, 2000) eine von den amerikanischen Originalautoren entwickelte Kurzform des Fragebogens zur gesundheitsbezogenen Lebensqualität SF-36 verwendet, welche ebenfalls die zwei Summenskalen (Psychische und Körperliche Summenskala, PSK bzw. KSK) enthält. Der SF-8-F wurde als ergänzendes Fremdeinschätzungsverfahren zur therapeutenseitigen Erfassung der Beeinträchtigung der gesundheitsbezogenen Lebensqualität eingesetzt. Eingeschätzt werden sollte der Zeitraum der zurückliegenden Woche, es existiert jedoch auch eine Version mit einem 4-Wochen-Zeitfenster. Der SF-8-F ist eine von Schulz et al. (in Vorbereitung) entwickelte deutschsprachige Anpassung des Selbsteinschätzungsbogens SF-8 an die Erfordernisse eines Fremdeinschätzungsinstruments und lehnt sich an das bestehende Fremdeinschätzungsverfahren des SF-36 an.

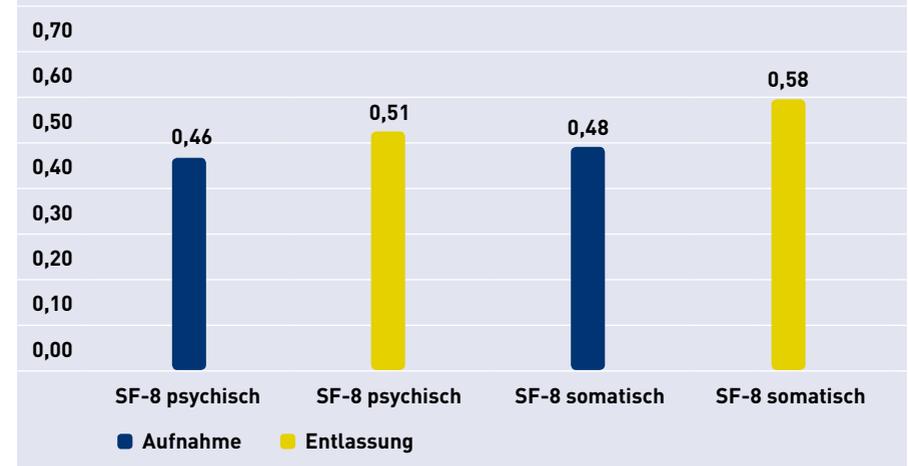
Ergänzend wurden verschiedene klinische, soziodemografische und sozialmedizinische Variablen erhoben sowie weitere standardisierte Selbst- und Fremdeinschätzungsverfahren. Die Erhebungen (Selbsteinschätzungen, Fremdeinschätzungen durch die Bezugstherapeuten) fanden in der Regel in den ersten drei Tagen nach Aufnahme in der Klinik sowie in den letzten drei Tagen vor der Entlassung aus der Klinik statt. Die Fragestellungen wurden mit Hilfe von Produkt-Moment-Korrelationen überprüft. Die Interpretation der Höhe des Zusammenhangs orientiert sich an Cohen (1988), mit  $r > .10$  als kleine,  $r > .30$  als mittlere und  $r > .50$  als große Effektstärke.

### 3. Ergebnisse

Die Zusammenhänge zwischen Selbst- und Fremdeinschätzung auf den beiden Summenskalen des SF-8 und

SF-8-F sind für die Gesamtstichprobe (siehe Abbildung 1) sowohl für den Aufnahme-, als auch für den Entlassungszeitpunkt von großer (KSK) bzw. annähernd großer Effektstärke (PSK). Für beide Skalen erweist sich erwartungsgemäß die Korrelation zur Entlassung als signifikant höher gegenüber der zur Aufnahme.

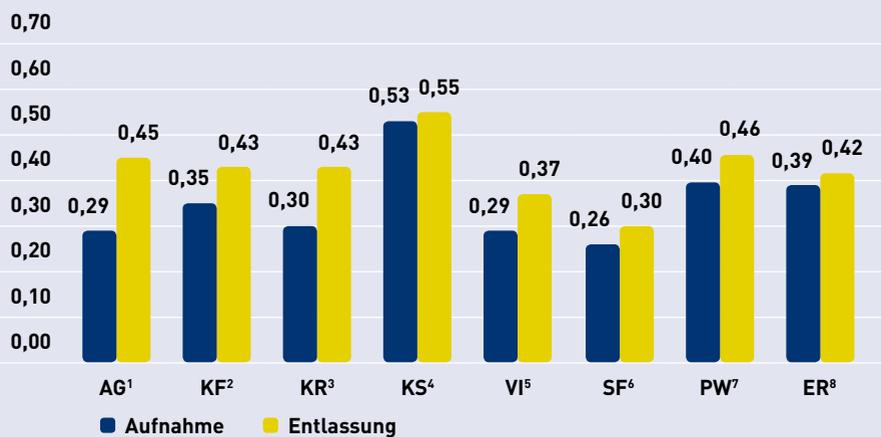
**Abbildung 1: Zusammenhang zwischen Selbst- und Fremdeinschätzungen auf den Summenskalen des SF-8 bzw. SF-8-F (Produkt-Moment-Korrelationen)**



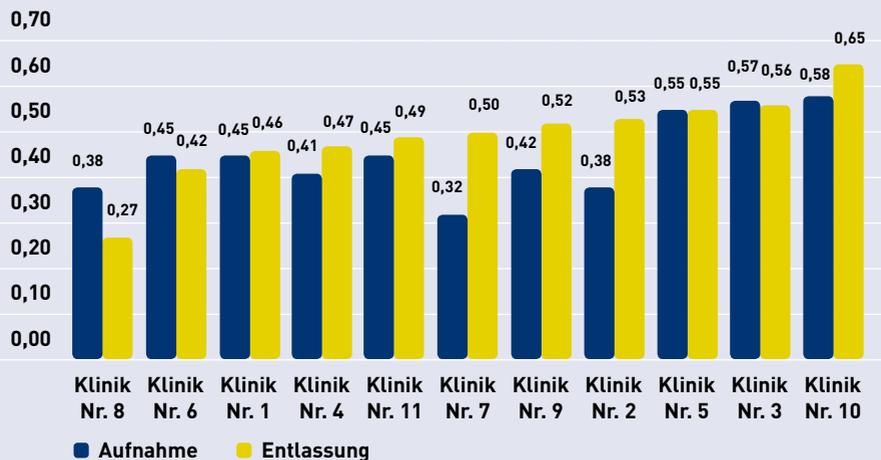
Auf der Ebene der einzelnen Items (siehe Abbildung 2) findet sich für das Item 4 (Schmerz) die höchste Übereinstimmung zu beiden Zeitpunkten, bei den drei anderen Items der Körperlichen Summenskala findet sich eine Zunahme der Zusammenhänge zur Entlassung. Bei der Psychischen Summenskala werden die Items 7 (Psychisches Wohlbefinden) und 8 (Emotionale Rollenfunktionen) zu beiden Zeitpunkten übereinstimmender eingeschätzt als die Items 5 (Vitalität) und 6 (Soziale Funktionsfähigkeit).

Eine wichtige Quelle für verschieden hohe Übereinstimmungen zwischen den Einschätzungen der Patienten und den Fremdeinschätzungen der Therapeuten könnten die Kliniken seien, in denen die Patienten behandelt

**Abbildung 2: Zusammenhang zwischen Selbst- und Fremdeinschätzungen der Items des SF-8 bzw. SF-8-F (Produkt-Moment-Korrelationen)**



**Abbildung 3: Zusammenhang zwischen Selbst- und Fremdeinschätzungen auf der psychischen Summenskala des SF-8 bzw. SF-8-F (Produkt-Moment-Korrelationen) getrennt nach Klinik**



- 1 Allgemeine Gesundheitswahrnehmung
- 2 Körperliche Funktionsfähigkeit
- 3 Körperliche Rollenfunktionen
- 4 Körperliche Schmerzen
- 5 Vitalität
- 6 Soziale Funktionsfähigkeit
- 7 Psychisches Wohlbefinden
- 8 Emotionale Rollenfunktion

wurden: Es zeigte sich für die PSK, dass zur Aufnahme die Werte je nach Klinik zwischen .38 und .58 und zur Entlassung zwischen .27 und .68 liegen (siehe Abbildung 3). Eine vergleichbare Spannweite findet sich auch für die KSK, hier liegen die Werte zur Aufnahme zwischen .48 und .66 sowie zur Entlassung zwischen .42 und .67.

## 4. Diskussion

Es sollte untersucht werden, inwieweit die zwei zentralen Dimensionen der gesundheitsbezogenen Lebensqualität, das körperliche und das psychische Wohlbefinden (Cella et al., 2005) im Selbst- und im Fremdrating zusammenhängen und ob die Variation in der Höhe des Zusammenhangs von den einzelnen Kliniken abhängig sein könnte. Dazu konnte an einer Stichprobe von 1.812 Patienten aus elf Fachkliniken für Patienten mit psychischen und psychosomatischen Störungen auf Einschätzungen zu Beginn und am Ende der Behandlung zurückgegriffen werden, die mittels des SF-8 vorgenommen wurden, wobei eine Selbst- und eine Fremdratingform eingesetzt wurden.

Die Ergebnisse zeigen, dass vor allem für den Zeitpunkt der Entlassung, aber auch bereits zur Aufnahme Zusammenhänge mit großer bzw. annähernd großer Effektstärke bestehen. Angesichts der im Konzept verankerten ausdrücklichen Betonung der subjektiven Einschätzung durch den Patienten mag dieser relativ hohe Zusammenhang überraschen. Jedoch bleibt trotz des hohen Zusammenhangs nur ein Anteil des festgelegten Unterschieds von 25% bis 36%, ein beträchtlicher Teil der Variation der Selbsteinschätzungen durch den Patienten kann demnach nicht durch den behandelnden Therapeuten erklärt werden. Zudem basieren die Fremdeinschätzungen der Therapeuten zumeist auf Informationen, die sie in ersten (anamnestischen) Gesprächen mit den Patienten erfahren haben, die

Fremdeinschätzungen enthalten also auch immer zu einem bestimmten Anteil Selbstauskünfte des Patienten.

Da der SF-8 so entwickelt worden ist, dass nicht nur die beiden Summenskalen interpretiert werden können, sondern jedes der acht Items eine der acht Skalen des umfassenderen SF-36 repräsentiert, können auch die Ergebnisse auf Ebene der einzelnen Items von Interesse sein: Hier finden sich von den vier Items der KSK insbesondere für das Item Schmerz (KS) hohe Übereinstimmungen, bei der PSK sind es die Items Psychisches Wohlbefinden (PW) und Emotionale Rollenfunktion (ER), die vergleichsweise höher übereinstimmend eingeschätzt werden. Diese Ergebnisse sind plausibel, weil Schmerz ein vergleichsweise leicht zu kommunizierendes Merkmal ist und zudem auch in der unmittelbaren Gesprächssituation mit dem Patienten gut beurteilt werden kann, was, mit Einschränkungen, auch für die globale Einschätzung psychischer Probleme gelten könnte. Einschränkungen in beruflichen und alltäglichen Rollenfunktionen wiederum weisen einen unmittelbaren Bezug zum rehabilitativen Auftrag der Behandlung in der hier untersuchten Indikationsgruppe auf. Denkbar ist also, dass diese Aspekte in den Gesprächen mit den Patienten besonders fokussiert wurden und deshalb die Übereinstimmung zwischen Selbst- und Fremdurteil etwas höher ausfällt als auf den anderen beiden Items der PSK.

Die vorgenommenen klinikvergleichenden Auswertungen machen deutlich, dass je nach Klinik bedeutsame Unterschiede in der Übereinstimmung zwischen Selbst- und Fremdeinschätzung bezüglich der gesundheitsbezogenen Lebensqualität der Patienten bestehen. Dieses kann sowohl auf Unterschiede in der Zusammensetzung der Patienten der Kliniken zurückzuführen sein, es kann aber auch unterschiedliche Struktur- und Prozessmerkmale der Einrichtungen widerspiegeln. Eine wichtige Variable in diesem Zusammenhang ist u.U. der Zeitpunkt des Ausfüllens der Fremdeinschätzung

und damit verbunden das unterschiedliche Ausmaß, in dem der behandelnde Arzt oder Therapeut Informationen vom Patienten selbst erhalten hat, die er zur Beurteilung der Lebensqualität heranziehen kann. Dies betrifft allerdings vor allem den Aufnahmezeitpunkt und kann die bestehenden Unterschiede zum Entlassungszeitpunkt nicht erklären. Zur Aufklärung dieser Ungleichheiten sind somit noch ausstehende Analysen im Zusammenhang mit den - ebenfalls im Rahmen dieser Studie - auch erhobenen Struktur- und Prozessmerkmalen der Kliniken erforderlich.

Die hier vorgestellten Ergebnisse unterliegen einigen Einschränkungen: Es wurden Patienten untersucht, die wegen psychischer Störungen eine stationäre, primär psychotherapeutische Behandlung erhalten haben: Inwieweit die Ergebnisse auch auf andere Indikationsgruppen und/oder Behandlungssettings übertragbar sind, bleibt offen. Da das Konzept der gesundheitsbezogenen Lebensqualität bei Patienten mit primär somatischen Erkrankungen seinen Ausgangspunkt genommen hat, ist eine Untersuchung dieser Patientengruppen von besonderer Bedeutung. Zu fragen ist auch, inwieweit die Zusammenhänge zwischen Selbst- und Fremdeinschätzung der Lebensqualität geringer ausfallen, wenn nicht ein 1-, sondern ein 4-Wochen Zeitfenster eingeschätzt wird. Ebenso fehlt es noch an Wissen zur Frage des Zusammenhangs der Einschätzung nicht mit Behandlern, sondern den Familienangehörigen der Patienten, da diese im Allgemeinen eine sehr viel vertraulichere und zeitlich umfassendere Grundlage zur Beurteilung besitzen. Weiterhin bleibt auch offen, inwieweit die vorliegenden Ergebnisse vom eingesetzten Fragebogen abhängig sind und ob bei anderen allgemeingültigen, aber vor allem auch krankheitsspezifischen Verfahren höhere Zusammenhänge zwischen Selbst- und Fremdeinschätzung bestehen. Als ein Vorteil des eingesetzten Verfahrens kann allerdings seine Kürze, dass heißt die Möglichkeit zur zeitökonomischen Nutzung und seine Anwendbarkeit angesehen werden:

Umfassende Analysen zu seinen psychometrischen Eigenschaften, die über die hier vorgestellten Auswertungen hinausgehen stehen jedoch noch aus und sind in Vorbereitung.

Als Fazit für die Qualitätssicherung von Behandlungen im Rahmen der medizinischen Rehabilitation kann festgehalten werden, dass die vorliegenden empirischen Befunde zur gesundheitsbezogenen Lebensqualität als einem zentralen Outcomekriterium es nicht erlauben, sich auf Fremdeinschätzungsverfahren zu stützen, sondern die Einschätzung der Patienten selbst der entscheidende Qualitätsindikator sein sollte.

## 5. Literatur

**Bullinger, M., & Kirchberger, I.** (1998): SF-36. Fragebogen zum Gesundheitszustand. Göttingen: Hogrefe-Verlag.

**Cella, D., Chang, C.-H., Wright, B. D., Von Roenn, J. H., & Skeel, R. T.** (2005): Defining Higher Order Dimensions of Self-Reported Health: Further Evidence for a Two-Dimensional Structure. *Evaluation & The Health Professions*, 28 (2), 122-141.

**Cohen, J.** (1988): *Statistical power analysis for the behavioral sciences* (2. ed.). Hillsdale, NJ: Lawrence Erlbaum.

**Engel, G. L.** (1977): The need for a new medical model: A challenge for biomedicine. *Science*, 196, 129-136.

**Kawski, S., & Koch, U.** (2004): Qualitätssicherung in der medizinischen Rehabilitation in Deutschland - Entwicklungsstand und Perspektiven. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, 47, 111-117.

**Schulz, H., Harfst, T., Dirmaier, J., Watzke, B., Andreas,**

**S., Kawski, S. et al.** (in Vorbereitung): Konstruktion und erste psychometrische Überprüfung einer Fremdeinschätzungsversion des SF-8-Fragebogens.

**Ware, J. E.** (2000): SF-36 Health survey update. *Spine*, 25 (24), 3130- 3139.

**Ware, J. E., Kosinski, M., Dewey, J. E., & Gandek, B.** (2000): *How to score and interpret single-item health status measures: A manual for users of the SF-8 Health Survey*. Lincoln: Quality Metric.

## 6. Anhang

### SF-8 Fremdeinschätzungsbogen (Beispiel: Aufnahmezeitpunkt)

Fragen zum allgemeinen Gesundheitszustand zu Behandlungsbeginn (SF-8)					
1. Wie würden Sie den Gesundheitszustand des Patienten / der Patientin in der vergangenen Woche im Allgemeinen einschätzen					
ausgezeichnet <input type="checkbox"/>	sehr gut <input type="checkbox"/>	gut <input type="checkbox"/>	weniger gut <input type="checkbox"/>	schlecht <input type="checkbox"/>	sehr schlecht <input type="checkbox"/>
2. Wie sehr haben Probleme mit der körperlichen Gesundheit den Patienten / die Patientin in der vergangenen Woche bei normalen Tätigkeiten eingeschränkt (zu Fuß gehen, Treppen steigen)?					
überhaupt nicht <input type="checkbox"/>	sehr wenig <input type="checkbox"/>	mäßig <input type="checkbox"/>	ziemlich <input type="checkbox"/>	war zu körperlichen Tätigkeiten nicht in der Lage <input type="checkbox"/>	
3. Inwieweit hatte der Patient / die Patientin in der vergangenen Woche wegen seiner / ihrer körperlichen Gesundheit Schwierigkeiten bei der Ausübung seiner / ihrer täglichen Arbeit bzw. Tätigkeiten zu Hause und außer Haus?					
überhaupt nicht <input type="checkbox"/>	sehr wenig <input type="checkbox"/>	mäßig <input type="checkbox"/>	ziemlich <input type="checkbox"/>	war zu körperlichen Tätigkeiten nicht in der Lage <input type="checkbox"/>	
4. Wie stark waren die Schmerzen des Patienten / der Patientin in der vergangenen Woche?					
hatte keine Schmerzen <input type="checkbox"/>	sehr leicht <input type="checkbox"/>	leicht <input type="checkbox"/>	mäßig <input type="checkbox"/>	stark <input type="checkbox"/>	sehr stark <input type="checkbox"/>
5. Wie viel Energie hatte der Patient / die Patientin in der vergangenen Woche					
sehr viel <input type="checkbox"/>	ziemlich viel <input type="checkbox"/>	mäßig viel <input type="checkbox"/>	ein wenig <input type="checkbox"/>	gar keine <input type="checkbox"/>	
6. Wie sehr haben die körperlichen Gesundheit oder seelische Probleme des Patienten / der Patientin in der vergangenen Woche seine / ihre normalen Kontakte zu Familienangehörigen oder Freunden eingeschränkt?					
überhaupt nicht <input type="checkbox"/>	sehr wenig <input type="checkbox"/>	mäßig <input type="checkbox"/>	ziemlich <input type="checkbox"/>	war zu diesen Kontakten nicht in der Lage <input type="checkbox"/>	
7. Wie sehr haben dem Patienten / der Patientin seelische Probleme (z. B. Angst, Niedergeschlagenheit oder Reizbarkeit) in der vergangenen Woche zu schaffen gemacht?					
überhaupt nicht <input type="checkbox"/>	etwas <input type="checkbox"/>	mäßig <input type="checkbox"/>	ziemlich <input type="checkbox"/>	sehr <input type="checkbox"/>	
8. Wie sehr haben die persönlichen oder seelischen Probleme des Patienten / der Patientin ihn / sie in der vergangenen Woche daran gehindert, seine / ihre normalen Tätigkeiten im Beruf, in der Schule / im Studium oder andere alltägliche Tätigkeiten auszuüben?					
überhaupt nicht <input type="checkbox"/>	sehr wenig <input type="checkbox"/>	etwas <input type="checkbox"/>	ziemlich <input type="checkbox"/>	konnte alltägliche Aktivitäten nicht ausüben <input type="checkbox"/>	

## Ausblick zur Ergebnisqualität - Indikatoren einer erfolgreichen Rehabilitation sowie Einflussfaktoren auf das Rehabilitationsergebnis

Dr. Teresia Widera

Deutsche Rentenversicherung Bund

Sozialmedizin und Rehabilitation

Reha-Qualitätssicherung, Epidemiologie und Statistik

### Inhaltsverzeichnis

1. Reha-Ziele im Verhältnis zu Reha-Ergebnissen	169
2. Mögliche Erfolgskriterien	170
3. Stärken und Schwächen der einzelnen Erfolgsparameter	172
4. Einflussfaktoren auf den Reha-Erfolg und seine Bewertung	176
5. Implikationen für die Erfolgsmessung	183
6. Literatur	185

## 1. Reha-Ziele im Verhältnis zu Reha-Ergebnissen

Die Ziele medizinischer Rehabilitation sind außerordentlich weit gespannt und reichen von der Verbesserung bestehender Funktionseinschränkungen, der Entwicklung von Fähigkeiten und Fertigkeiten zur Bewältigung der Belastung durch chronische Erkrankungen bis zu der Vermeidung von Beeinträchtigungen der Teilhabe und der Aufrechterhaltung bzw. Steigerung der Leistungsfähigkeit.

Das spezielle Ziel medizinischer Rehabilitationsmaßnahmen durch die Rentenversicherung ist – entsprechend ihres gesetzlichen Auftrages zur Senkung des Risikos einer Minderung oder Aufhebung der Erwerbsfähigkeit aus krankheits- oder behinderungsbedingten Gründen – primär gerichtet auf die Eingliederung in das Erwerbsleben bzw. die Verhinderung des vorzeitigen Ausscheidens aus dem Erwerbsleben durch Erwerbsminderungsrente (vgl. dazu § 10 SGB VI).

Will man den Erfolg einer Rehabilitationsmaßnahme ermitteln, muss man Wirkungen signifikant nachweisen und die Wirkungen der Reha-Maßnahme sicher zuordnen können. Die Erfolgsfeststellung ist daran geknüpft, dass durch die Rehabilitations-Maßnahme (positive) Veränderungen im Sinne von Wirkungen gegenüber der Ausgangslage – als Zustand vor der Rehabilitation – verursacht worden sind. Hinzu kommt, dass die Erfolgsmessung nicht allein auf der Grundlage der Wirkung von Maßnahmen erfolgen kann, sondern auch Fragen der Wirksamkeit im Sinne der Erreichung von Reha-Zielen und Kosten-Nutzen-Optimierung berücksichtigen muss. Ferner sollte die Ergebnismessung so beschaffen sein, dass das Spektrum der Wirksamkeitsparameter der umfassenden bio-psycho-sozialen Zielsetzung der Rehabilitation gerecht wird (vgl. dazu z.B. DRV Bund, 2007) und mögliche Zielkonflikte zwischen Effizienz- und Effektivitätserwägungen gering gehalten werden.

Die zentrale Frage zur Ergebnisqualität: „Ist die Rehabilitation effektiv?“ lässt sich daher nur schwer mit einem eindeutigen „ja“ oder „nein“ beantworten (vgl. dazu z.B. Plath & Blaschke, 1999).

## 2. Mögliche Erfolgskriterien

Der Erfolg der Rehabilitation kann variieren, je nachdem an Hand welcher Erfolgskriterien man das Rehabilitationsergebnis misst (vgl. dazu die Aufstellung möglicher Erfolgskriterien für das Reha-Ergebnis in Tabelle 1).

Eine Messung des Rehabilitations-Erfolges über Kriterien des beruflichen Verbleibs wie Erwerbstätigkeit oder Eingliederungsraten in Arbeit kann zu anderen Ergebnissen führen als eine Outcomemessung über eine Prüfung der Zielerreichung auf gesundheitlicher Ebene (z.B. Symptomlinderung, Verbesserung gesundheitlicher Beschwerden), funktionaler Ebene (z.B. Schmerzreduktion, Gewichtsabnahme, verbesserte Vitalkapazität, erhöhte Mobilität und verstärkte Gelenkbeweglichkeit bei der Bewegungsprüfung, Leistung auf dem Ergometer etc.) oder edukativer Ebene (z.B. Verbesserung des Informationsstandes über die reha-begründende Grunderkrankung, Beherrschung von Coping-Strategien zur Krankheitsbewältigung, Erlernen von Techniken zum Abbau von gesundheitsrelevanten Risikoverhaltensweisen, Verbesserung der Anpassung an die Krankheit durch Entwicklung internaler, gesundheitsbezogener Kontrollüberzeugungen, Erlernen günstiger Krankheitsverarbeitungsprozesse).

Eine gute Prozessqualität und eine hohe Patientenzufriedenheit sind nicht notwendigerweise positiv korreliert; ein negatives Patientenurteil über eine Einrichtung geht nicht automatisch mit einer ablehnenden Peer-Bewertung der Einrichtung einher. Fragebogendaten zur

subjektiven Zufriedenheit mit dem Behandlungsergebnis (Patientenzufriedenheit) und dem subjektiven Reha-Outcome (Behandlungserfolg aus Patientensicht) führen ggf. zu anderen Einschätzungen des Reha-Ergebnisses als Verwaltungsdaten etwa zu Maßnahmeabbrüchen oder zur Anschlussarbeitslosigkeit (vgl. dazu z.B. Burcardi et al., 2002; Klein et al., 2002; Nischan et al., 2003; Schubert et al., 2007).

**Tabelle 1: Mögliche Erfolgskriterien für das Rehabilitationsergebnis**

Erfolgskriterium	Datenquelle
ZUFRIEDENHEIT des Rehabilitanden mit dem Behandlungsergebnis (Patientenzufriedenheit) und das subjektive Reha-Outcome (Behandlungserfolg aus Patientensicht)	Ergebnisberichte zu Rehabilitandenbefragungen (vgl. dazu z.B. Widera & Klosterhuis, 2007a)
VERBLEIBSKRITERIEN wie Erwerbstätigkeit, Berufsrückkehr bzw. (Wieder-) Eingliederungsraten in Arbeit oder „negativ“ formuliert Anschlussarbeitslosigkeit, AU-Zeiten, Krankmeldungen, Erwerbsminderungsrentenbezug	Sozialmedizinischer Verlauf auf der Grundlage der Reha-Statistik-Datenbasis der Rentenversicherung (vgl. dazu z.B. Egner & Grünbeck, 2003; Klosterhuis, Zollmann, Grünbeck, 2004; Rehfeld, 2008)
Erreichung von konkreten BEHANDLUNGSZIELEN, z.B. auf gesundheitlicher, funktionaler, edukativer, psychosozialer Ebene	Einsatz von Behandlerbögen, Erhebung des klinischen Status, Messung von Körperfunktionsparametern, Laborbefunde, Leistungstests/-beurteilungen, Erfassung medizinischer Basisparameter, Ergebnisse apparativer Diagnostik, ärztliche Entlassungsberichte (vgl. dazu z.B. DRV Bund, 2007b), sozialmedizinische Einschätzungen (vgl. dazu z.B. Ghafari, 2006), psychologische Testdaten auf der Grundlage von Assessments sowie Erhebung über gesundheitsbezogene Skalen (vgl. dazu z.B. Biefang et al., 1999)
Steigerung der krankheitsspezifischen und krankheitsübergreifenden gesundheitsbezogenen LEBENSQUALITÄT	Selbsteinschätzung der gesundheitlichen Befindlichkeit durch den Patienten in einem Fragebogen (vgl. dazu z.B. Bayerl, 2009; Bengel et al., 2008)
Aufrechterhaltung und Verbesserung der TEILHABE, Partizipation, Selbstwirksamkeit und Aktivität auch in Lebensbereichen jenseits der Arbeit und Erwerbsbiographie, z.B. im sozialen Leben oder bei der Selbstversorgung im häuslichen Kontext	Erfassung etwa über neuropsychologische Testbatterien zur Messung der Alltagskompetenz, Sozialfragebögen zur Überprüfung der sozialen Kontakte, Aktivitäten, Unterstützung; Wohnsituation; wirtschaftlichen Verhältnisse, MOSES-Fragebogen (vgl. dazu z.B. Farin, 2008)

Die Erfolgsevaluation ist also an die Verfügbarkeit von geeigneten Erfolgskriterien, ihre kombinierte Verwendbarkeit sowie die Bewertung der nachgewiesenen Wirkungen als ausreichend für die Erfolgsfeststellung geknüpft (vgl. dazu z.B. Linden, 2006).

### 3. Stärken und Schwächen der einzelnen Erfolgsparameter

Jeder Qualitätsindikator für sich allein genommen vereint Stärken und Schwächen auf sich, welche nur durch eine kombinierte Betrachtungsweise überwunden werden können (vgl. dazu z.B. Glattacker & Jäckel, 2007; Reiter et al., 2007).

#### 3.1 Patientenurteil

Die subjektive Gesundheitsdefinition, die Zufriedenheit des Patienten mit der Rehabilitation und der patienten-seitig wahrgenommene Behandlungserfolg beziehen sich auf das Selbsterlebte des Individuums.

Das Patientenurteil wird häufig als „subjektiv“ gering geschätzt. Subjektivität wird oft mit Unsachlichkeit, Voreingenommenheit und Beeinflussung durch persönliche Gefühle, Interessen oder Vorurteile gleichgesetzt. Subjektive Eindrücke sollen Fehleinschätzungen unterliegen. Medizinische Prozesse, so heißt es vielfach, können von Patienten nicht richtig eingeschätzt werden; es mangelt an Beurteilungskompetenz des Patienten (vgl. dazu z.B. Meyer, 2003). Von rentenfixierten Patienten, die nur noch in der Rentenantragstellung und Rentengewährung den einzigen Ausweg aus ihrer gesundheitlich-beruflichen Problemlage sehen, sind positive Reha-Urteile und eine ausgeprägte Zufriedenheit mit der Reha selten zu erwarten, weswegen dieser Qualitätsindikator eine Schlagseite aufweist (vgl. dazu z.B. Gönner et al., 2008; Kulick et al., 2002).

Dem kann man entgegensetzen: Es kommt also nicht nur darauf an, ob man laut Arzturteil und Laborbefund gesund ist, sondern auch – in der Reha möglicherweise sogar wichtiger – ob man sich gesund fühlt. Die Qualität der Behandlung muss sich für den einzelnen Patienten nachhaltig auch in seinem eigenen Empfinden niederschlagen. Laborwerte sind Hilfsgrößen, welche nur eine sehr begrenzte Aussagekraft hinsichtlich des tatsächlichen Wohlbefindens eines Patienten aufweisen, da sie nicht von unmittelbarer Bedeutung für die Patienten sind, nur indirekt Relevanz für die Patienten haben und nicht immer mit einem sichtbaren Patientennutzen einhergehen. Wohingegen der so genannte „subjektive Gesundheitszustand“ von essentieller Bedeutung für Lebensqualität ist (vgl. dazu z.B. Bayerl, 2009).

Ferner ist die subjektive Einschätzung der Rehabilitation durch den Patienten ein guter Indikator für das Reha-Outcome (Beckmann et al., 2002; Mittag et al., 2006; Widera & Beckmann, 2007b). Die Patientenmeinung verfügt zudem über prognostische Validität für die Krankheitsbewältigung: Das subjektive Erleben der Behandlung und des Behandlungserfolges ist ein nicht zu unterschätzender Faktor bei der Bewältigung der Erkrankung und von entscheidender Bedeutung für den Rehabilitationsverlauf selbst. Die Zufriedenheit des Patienten kann demzufolge als Genesungsfaktor angesehen werden (vgl. dazu z.B. Ghafuri, 2006; Zwingmann et al., 2005).

### **3.2 Arzturteil**

Das Arzturteil wird häufig als dem Patientenurteil überlegen betrachtet: Dabei stellt das Arzturteil ebenso einen subjektiven Eindruck dar, weil die ärztliche Einschätzung in den meisten Fällen an eine persönliche Interpretationsleistung geknüpft ist (vgl. dazu z.B. Hansis & Hart, 2001).

Selbstverständlich weisen Patienten und Ärzte bzw. Kliniken zwei unterschiedliche Beurteilungsperspektiven

auf, welche nicht deckungsgleich sind, aber trotz der Asymmetrie zwischen der professionellen Sicht und der patientenseitigen Sicht besteht eine hohe Übereinstimmung zwischen dem Arzturteil und dem Patientenurteil hinsichtlich der Einschätzung verschiedener Aspekte der Rehabilitation und der Beurteilung des Reha-Erfolges. Die Erfolgskriterien des Arztes weichen nicht wesentlich von den Erfolgsmaßstäben des Patienten ab. Ärzte treffen im Großen und Ganzen mit ihrer Gesamtbeurteilung des Behandlungserfolges auch die Einschätzung der Patienten (vgl. dazu z.B. Evangelou, 2008).

### **3.3 Medizinische Parameter**

Vor einer Überbewertung medizinischer Faktoren in der Rehabilitationspraxis im Hinblick auf die Aufrechterhaltung oder Wiedererlangung der Arbeitsfähigkeit (im Kontext der Reha) ist zu warnen. Denn es kann am Ende der Reha oder zum Katamnesezeitpunkt eine subjektiv gute Gesundheitseinschätzung bei objektiv schlechter Gesundheit (Laborwerte, Körperfunktionsparameter, apparative Diagnostik) ebenso auftreten, wie eine subjektiv schlechte Gesundheitseinschätzung bei objektiv fehlendem organischen Krankheitsbefund (vgl. dazu z.B. Dörner, 2004).

### **3.4 Kriterien des beruflichen Verbleibs**

Wesentliches Ziel der Rehabilitation der Rentenversicherung ist es, möglichst lange die Erwerbsfähigkeit der Versicherten zu erhalten und so eine gesundheitlich bedingte Erwerbsminderungsrente zu verhindern oder zumindest zeitlich hinauszuschieben. Auf Grund dieser Vorgaben für die Rehabilitation der Rentenversicherung muss es folgerichtig von wesentlicher Bedeutung sein, den Erwerbsverlauf der Rehabilitanden, wie er sich nach einer medizinischen Rehabilitation empirisch darstellt, nachzuvollziehen.

Die Reha-Statistik-Datenbasis der Rentenversicherung erlaubt es, Rehabilitations- und Berentungsgeschehen im Zusammenhang zu analysieren. Dabei wird der

sozialmedizinische Verlauf des Versicherten nach einer Reha nachgezeichnet. Bei der Auswertung wird insbesondere der Stand der Beitragszahlungen zur Rentenversicherung als Indikator der beruflichen (Re-)Integration analysiert und Aussagen dazu getroffen, wie viele Rehabilitanden im Erwerbsleben verbleiben konnten, obwohl bei ihnen nach einem sozialmedizinischen Gutachten anlässlich der Reha-Antragsstellung bereits eine erheblich gefährdete oder sogar geminderte Erwerbsfähigkeit vorlag.

Streng genommen kann die weitere Erwerbsbiographie des Rehabilitanden nach der durchgeführten Reha-Maßnahme, wie sie aus den sozialmedizinischen Verlaufsergebnissen der RSD-Auswertung ersichtlich wird, auch nicht unkritisch als Erfolg einer Rehabilitation betrachtet werden, denn die weitere Erwerbstätigkeit von Rehabilitanden ist zu einem erheblichen Teil von Faktoren abhängig, die weder von der Rentenversicherung, noch von der Qualität der durchgeführten Reha-Maßnahme oder gar von dem Rehabilitanden beeinflusst werden können, wie z.B. der regionalen Arbeitsmarktlage vor Ort.

Ferner muss sichergestellt sein, dass Beitragsleistungen zur Rentenversicherung nach erfolgter Rehabilitation, die als Indikator einer erhaltenen oder wiederhergestellten Erwerbsfähigkeit fungieren, von Einflüssen der Arbeitslosigkeit bereinigt worden sind. So werden Zeiten der Arbeitslosigkeit bei der Rente berücksichtigt. Ein Teil der Beitragsleistungen nach einer Rehabilitation ist auch auf Arbeitslosigkeit zurückzuführen. Auch während der Zeit einer Arbeitslosigkeit werden Beiträge zur gesetzlichen Rentenversicherung gezahlt. Die Agentur für Arbeit trägt diese RV-Beiträge. Des Weiteren sollte beachtet werden, dass Arbeitssuchende ohne Leistungsanspruch nicht rentenpflichtversichert sind, und es werden auch keine Beiträge für die spätere Rente gezahlt; solche Zeiten sind höchstens unter bestimmten Bedingungen als Anrechnungszeiten ohne Bewer-

tung relevant. Die Versichertenstatistik enthält keine Informationen über diesen Personenkreis (vgl. dazu z.B. Egner & Grünbeck, 2003; Klosterhuis, Zollmann & Grünbeck, 2004; Rehfeld, 2008).

Vor dem Hintergrund der umfassenden Zielsetzung der Rehabilitation, die nicht nur die unmittelbare krankheitsbezogene Problematik, sondern auch die psychische, soziale und berufsbezogene Dimension einbezieht, ist zur Erfolgsevaluation ein Set an Qualitätsindikatoren notwendig. Erst die integrierte Betrachtung dieser Indikatoren im Zusammenspiel ermöglicht belastbare Aussagen zum Reha-Erfolg (vgl. dazu z.B. DRV Bund, 2007a).

## 4. Einflussfaktoren auf den Reha-Erfolg und seine Bewertung

Neben der Auswahl einer geeigneten Kriteriumsvariablen für die Erfolgsmessung hängt die Erfolgsfeststellung von einer Reihe von Einflussfaktoren auf den Reha-Erfolg ab. Die Einflussfaktoren stellen mögliche Determinanten der Bewertung von Prozess- und Ergebnismerkmalen einer Rehabilitations-Maßnahme dar und/oder bestimmen das Ausmaß des gemessenen Reha-Erfolges. Die bedeutsamen Einflussfaktoren können den Einflussgruppen in Tabelle 2 zugeordnet werden:

**Tabelle 2: Reha-Erfolg – Einflussfaktoren**

### **Einflussfaktoren auf den Reha-Erfolg bzw. die Bewertung des Reha-Erfolgs**

1. Umgebungs-, Kontext- oder Situationsmerkmale der Reha-Maßnahme
2. (Reha-)Einrichtungsspezifische Faktoren
3. Bedingungen der Zielerreichung während der Reha-Maßnahme
4. Personale Voraussetzungen der Rehabilitanden
5. Gesundheitsbezogene, berufliche und soziale Ausgangslage des Rehabilitanden vor der Reha-Maßnahme
6. Merkmale der Situation des Rehabilitanden nach der Reha-Maßnahme

Mit den Kontextbedingungen der Reha-Maßnahme sind die gesetzlichen Vorgaben zur Ausgestaltung und Durchführung der Reha-Leistung, die Vergütung der Reha-Maßnahme durch die Reha-Träger sowie das Wirtschaftlichkeitsgebot für rehabilitative Leistungen gemeint. Die Bedingungen der Zielerreichung beinhalten Merkmale wie die Dauer der Maßnahme, die Organisation, spezifische Zielsetzung und inhaltliche Ausrichtung der Reha-Maßnahme, die Indikation, die während der Reha behandelt wird, die Erfolgsprognose der Behandler und der Behandelten, den Zeitabstand zwischen Akutereignis und Beginn der Rehabilitation sowie Anzahl der Rehabilitations-Maßnahmen vor der letzten Rehabilitation. Unter einrichtungsspezifischen Faktoren versteht man beispielsweise das Behandlungs- und Betreuungsangebot der Klinik sowie Qualitätsaspekte der Reha-Einrichtung bzw. der Reha-Maßnahme.

Diese drei vorgenannten Einflüsse sind zweifelsohne bedeutsame Prädiktoren für den Reha-Erfolg; sie sollen jedoch kein Gegenstand der Betrachtungen an dieser Stelle sein, sondern sind ausführlich in Kosten-Nutzen-Analysen zur Rehabilitation behandelt (vgl. dazu z.B. DRV Bund, 2006; Hansmeyer et al., 1997; Koch & Dintsios, 2008).

Vor dem Hintergrund messtheoretischer Forderungen an die Erfassung der Ergebnisqualität wichtiger sind mögliche Determinanten der Bewertung von Prozess- und Ergebnismerkmalen einer Rehabilitationsmaßnahme, welche häufig unberücksichtigt bleiben, die Ergebnismessung jedoch verfälschen. Dazu zählen die personalen Voraussetzungen der Rehabilitanden, davon abgegrenzt die gesundheitsbezogene, berufliche und soziale Ausgangslage der Rehabilitanden vor der Reha-Maßnahme sowie zum dritten die Merkmale der Situation des Rehabilitanden nach der Reha-Maßnahme.

#### **4.1 Personale Voraussetzungen der Rehabilitanden mit Einfluss auf die Reha-Bewertung /-Zufriedenheit**

Zu den personalen Voraussetzungen der Rehabili-

tanden mit Einfluss auf die Reha-Bewertung/-Zufriedenheit gehören im engeren Sinne die soziodemographischen Merkmale, die gesundheitsrelevanten Einstellungen und Verhaltensweisen sowie psychologische Faktoren des Rehabilitanden:

##### **Alter**

Jüngere Patienten tendieren zu kritischeren Bewertungen als ältere Patienten. Mit zunehmendem Alter nimmt die Kritikbereitschaft ab bzw. die Zufriedenheit zu (vgl. dazu z.B. Bijani, 2005; Deck et al., 2000; Farin, 2005; Klotz et al., 1996; Mitschele & Klosterhuis, 2007; Modestin et al., 2003; Nischan & Klosterhuis, 2002; Olandt & Krentz, 1998; Rundel et al., 2004; Raspe et al., 2006; Steffen & Pfaff, 2004).

##### **Geschlecht**

Es besteht eine gemischte Befundlage dahingehend, ob es überhaupt einen Geschlechtseffekt gibt, und wer kritischer urteilt, Männer oder Frauen. Ein Teil der Studien zeigt, dass das Patientenmerkmal Geschlecht einen Einfluss auf die Kontakthäufigkeit, die Arzt-Patienten-Interaktion, auf diagnostische Aktivitäten, auf die Festlegung von Therapiezielen und die Auswahl, Durchführung und Erfolgskontrolle von therapeutischen Maßnahmen und damit auch auf die subjektive Ergebnisqualität der Behandlung hat. Beide Geschlechter weisen verschiedene Rehabilitations-Bedürfnisse auf. So wünschen sich Frauen häufig mehr psychische und soziale Unterstützung in der Rehabilitation, die psychologische Betreuung tritt aber bisher im Rehabilitations-Alltag gegenüber körperlichen Maßnahmen deutlich in den Hintergrund, was sich auf die subjektive Beurteilung des Reha-Erfolges negativ auswirken kann. Andere Studien zeigen, dass Frauen tendenziell positiver, weil milder, urteilen als Männer (vgl. dazu z.B. Farin, 2004; Jäckel, 2002; Modestin et al., 2003; Mohr et al., 2008; Olandt & Krentz, 1998; Rundel et al., 2004; Raspe et al., 2006; Tesmer et al., 2006).

### **(Schul-)Bildung / Bildungsstand**

Die Befunde zur Ausprägung dieses Effektes stellen sich uneinheitlich dar. Auf der anderen Seite existieren Studien, die man folgendermaßen zusammenfassen könnte: Je höher der Bildungsstand, desto zufriedener. Je länger die Schulausbildung, desto größer ist die Zufriedenheit mit der rehabilitativen Versorgung. Auf der anderen Seite liegen Ergebnisse vor, denen zufolge in der Gruppe der vollkommen Zufriedenen die meisten einen Hauptschulabschluss aufweisen (vgl. dazu z.B. Bijani, 2005; Bleich et al., 2006; Bührlen-Armstrong et al., 1998; Mielck et al., 1993; Modestin et al., 2003; Rabung, 2005; Raspe et al., 2006).

### **Wohnort der Rehabilitanden**

Versicherte der neuen Bundesländer geben in vielen Bereichen der Rehabilitandenbefragung bessere Bewertungen ab als die Versicherten aus den alten Bundesländern. Rehabilitanden der „Ost“-Kliniken beurteilen die Rehabilitations-Maßnahme positiver als Rehabilitanden der „West“-Kliniken. Die Bewertungsunterschiede nivellieren sich jedoch mit zunehmender Dauer der Wiedervereinigung. Ggf. ergeben sich auch Zusammenhänge zur Einwohnerzahl des Wohnortes: Sowohl die vollkommen als auch die überwiegend Zufriedenen stammen mehrheitlich aus ländlicheren Gegenden mit einer Ortsgröße < 30.000 Einwohnern. Rund die Hälfte der gar nicht Zufriedenen kommt aus größeren Städten mit einer Einwohnerzahl von mehr als 100.000 (vgl. dazu z.B. Bijani, 2005; Jäckel, 2002; Nübling et al., 2003; Raspe et al., 2006).

### **Psychologische Faktoren**

Für die Differenzierungen bei den Ergebnissen von Patienten- oder Rehabilitandenbefragungen spielen auch unterschiedliche psychologische Konstrukte der Betroffenen, wie z.B. Selbstwirksamkeitserwartungen, Kontrollüberzeugungen, günstige/ungünstige Strategien zur Verarbeitung von Belastungen, angemessene/unangemessene Krankheitsverarbeitungsprozesse,

rehamotivationale Aspekte oder erfüllte/nicht erfüllte subjektive Erwartungen an die Rehabilitation, eine große Rolle. Ungünstige Ausprägungen auf diesen Parametern beeinflussen ihrerseits die Reha-Ergebnisse nach der Rehabilitation. Nicht erfüllte Erwartungen von Reha-Patienten korrelieren mit ihrer Rehabilitationszufriedenheit. Diese können aber durch andere Variablen, wie beispielsweise Geschlecht, Alter, Bildungsstand etc. vermittelt sein (vgl. dazu z.B. Deck, 2000, 2004, 2006; Dörner, 2004; Dohnke, 2004; Huhn, 2000).

### **Gesundheitsrelevante Einstellungen und Verhaltensweisen**

Gesundheitsrelevante Einstellungen (Risikokonnotative Einstellungen) und Verhaltensweisen (Alkohol- und Drogenkonsum, sportliche Aktivität, Essverhalten etc.) des Rehabilitanden sind sowohl für die patientenseitige Bewertung der Reha-Maßnahme zu einem Katamnesezeitpunkt nach der Reha (Einfluss der Persönlichkeit auf die subjektiv wahrgenommene Gesundheit) als auch für die Verstetigung des Reha-Erfolgs von Bedeutung (langfristige Aufrechterhaltung von gesundheitsrelevantem Verhalten). Konfundierende Zusammenhänge zwischen Reha-Bewertung, gesundheitsrelevanten Einstellungen und Verhaltensmustern sowie Geschlecht, Migrantenstatus, Bildungsniveau sowie Einkommen sind jedoch gegeben (vgl. dazu z.B. Lampert & Ziese, 2005; Schmidt et al., 1996; Weber, 2005).

### **Staatsangehörigkeit / Migrantenstatus bzw. Deutschkenntnisse**

Die Patientenzufriedenheit variiert im interkulturellen Vergleich. Migranten erweisen sich insgesamt als unzufriedener als deutsche Rehabilitanden mit der Versorgung im Krankenhaus und in Reha-Einrichtungen. Allerdings bestehen Korrelationen zwischen besseren deutschen Sprachkenntnissen und höherer Zufriedenheit. Nichtsdestotrotz ist die Zufriedenheit der befragten deutschen Patienten auch nach Berücksichtigung verschiedener Strukturunterschiede, wie z.B. Alter, Bildungsgrad und deutsche Sprachkenntnisse, deutlich

größer als die der Migranten. Des Weiteren fällt der Behandlungserfolg für Patienten mit Migrationshintergrund geringer aus und erzielt in gesundheitsbezogenen Skalen keine nachweisbaren Verbesserungen. Eine Vermittlung über eine höhere psychophysische Eingangsbelastung ist dabei jedoch nicht ausgeschlossen (vgl. dazu z.B. Borde, 2002 und 2003; Brandes & Walter, 2007; Mösko et al., 2008).

#### **4.2 Ausgangslage des Rehabilitanden vor der Reha-Maßnahme**

Betrachtet man die Einflussfaktoren auf den Reha-Erfolg, so ist die Ausgangssituation des Rehabilitanden vor der Reha-Maßnahme von den personalen Voraussetzungen der Rehabilitanden im eigentlichen Sinne abzugrenzen. Insbesondere die individuelle gesundheitsbezogene Ausgangslage des Rehabilitanden bei Aufnahme sowie sein Erwerbs- und Arbeitsfähigkeitsstatus zu Reha-Beginn sind als Funktion von Person und Situation anzusehen. D.h. z.B. die vorherige Arbeitssituation als Erwerbsstatus incl. Stellung im Betrieb bzw. Beruf des Rehabilitanden ist das Resultat von Personenmerkmalen wie Ausbildungsgüte und Umgebungsmerkmalen wie regionale Arbeitsmarktlage.

##### **Individuelle gesundheitsbezogene Ausgangslage**

Die gesundheitsbezogene Eingangsbelastung, indiziert über die Art der Erkrankung, den Schweregrad der Erkrankung, Ko- oder Multimorbidität, Vorliegen zusätzlicher psychischer Diagnosen bei somatischer Erstdiagnose, Chronifizierungs- bzw. Krankheitsdauer, Grad der Selbständigkeit oder Aktivitätseinschränkungen durch die Erkrankung, gesundheitsrelevante Ressourcen und Belastungen des Rehabilitanden bei Aufnahme etc., stellt einen bedeutsamen Prädiktor für den Reha-Erfolg bzw. für eine Wiederherstellung der Arbeitsfähigkeit im Anschluss an eine Reha-Behandlung dar. So bewerten beispielsweise Rehabilitanden, welche medizinische Leistungen wegen maligner Erkrankungen erhalten, die Rehabilitation positiver als Rehabilitanden mit einer

günstigeren Krankheitsprognose. Die allgemeine Patientenzufriedenheit in der psychosomatischen Rehabilitation ist geringer ausgeprägt als in der somatischen Rehabilitation: Möglicherweise treffen hier symptomorientierte Behandlungserwartungen der Patienten und problemorientierte Therapie aufeinander. Ferner unterscheidet sich die Patientenzufriedenheit in Abhängigkeit von der jeweiligen Indikation: In der neurologischen Rehabilitation ist die Zufriedenheit durchschnittlich höher als in anderen Indikationsbereichen. Des Weiteren lassen sich ungünstigere Patientenurteile zur durchgeführten Rehabilitation bei Vorliegen einer Komorbidität zu Reha-Beginn konstatieren. Überdies finden sich in der Gruppe der Unzufriedenen häufiger schwere Nebendiagnosen und Begleiterkrankungen. Außerdem sollte die psychische Befindlichkeit des Rehabilitanden berücksichtigt werden, denn die Zufriedenheit fällt niedriger aus, wenn die Befragten psychisch belastet sind (z.B. depressiv). Untersucht man den Einfluss der Verweildauer als Indikator für Krankheitsschwere und Krankheitsdauer lässt sich ebenso feststellen, dass Kurzzeitpatienten positiver antworten als Langzeitpatienten (vgl. dazu z.B. Farin, 2005; Irle et al., 2002; Mans, 1995; Mohr et al., 2008; Nübling, et al., 2003; Olandt & Krentz, 1998; Rabung, 2005; Raspe et al., 2006; Rundel et al., 2004; Wettach, 2001).

##### **Erwerbsstatus / Arbeitsfähigkeit des Rehabilitanden zu Reha-Beginn (bei Aufnahme)**

Arbeitslose Rehabilitanden, Rentenantragsteller unter den Rehabilitanden und Rehabilitanden mit langen Arbeitsunfähigkeit-Zeiten (AU) vor der Rehabilitation tendieren zu kritischeren Reha-Bewertungen als berufstätige Rehabilitanden sowie Rehabilitanden, die bei Reha-Aufnahme arbeitsfähig waren, oder Rehabilitanden, die keine Gedanken an Frühberentung vor, während und nach der Rehabilitation hatten/haben. Als mögliche Begründungen werden diskutiert: Höhere gesundheitliche Eingangsbelastung, Rentenfixierung sowie niedrigere Reha-Motivation (vgl. dazu z.B. Bleich

et al., 2006; Gönner et al., 2008; Huhn, 2000; Kulick et al., 2002; Rabung, 2005; Raspe et al., 2006).

#### 4.3 Merkmale der Situation des Rehabilitanden nach der Reha-Maßnahme

Die patientenseitige Zufriedenheit und die subjektive Bewertung des Reha-Erfolges hängen auch von gewissen Situationsmerkmalen des Rehabilitanden nach Ende der Reha-Maßnahme ab: War der Rehabilitand nach der Durchführung der Reha kritischen Lebensereignissen ausgesetzt? Nahm er Selbsthilfeangebote in Anspruch? War er in Sachen Nachsorge aktiv? Welches soziale Umfeld fand er nach seiner Rückkehr aus der Reha vor? Wie sieht seine allgemeine Zukunftsperspektive aus? Die Ergebnisse dazu können folgendermaßen zusammengefasst werden: Wenn das soziale Umfeld stabilisierend und unterstützend wirkt, wenn der Rehabilitand ausreichend beruflich integriert ist, und wenn eine stabile Wohnsituation gegeben ist, dann ist die Zufriedenheit mit der Reha als Indikator des Reha-Erfolges zum Katamnesezeitpunkt größer als wenn der Rehabilitand nach Ende der Reha belastenden Lebenssituationen, z.B. in Bezug auf die Wohnsituation, die Arbeitsintegration sowie die psychosoziale Einbindung, ausgesetzt wäre. Selbsthilfeaktivitäten und Nachsorgeaktivitäten des Rehabilitanden sind dabei insbesondere in ihrer Funktion für die Aufrechterhaltung des Reha-Erfolges nach der Entlassung aus der Einrichtung bedeutsam (vgl. dazu z.B. Weber et al., 2007).

## 5. Implikationen für die Erfolgsmessung

Die Reha-Qualitätssicherung der Rentenversicherung hat noch keine Königslösung zur Messung der Ergebnisqualität gefunden, weder bei der Auswahl der geeigneten Erfolgskriterien noch bei der Selektion der zu kontrollierenden Einflussfaktoren des Reha-Erfolges. Für die Reha-Qualitätssicherung sind primär Indika-

toren zu betrachten, die von den Reha-Einrichtungen auch zu beeinflussen sind. Allgemeine Indikatoren können für die Bewertung der Rehabilitation genutzt werden, für Einrichtungsvergleiche nach § 20 SGB IX sind sie eher nicht anwendbar.

**Tabelle 3: Messung der Ergebnisqualität in der Reha-Qualitätssicherung**

- Reha-Ziele mit Reha-Ergebnissen abgleichen
- Reha-Ziele und -ergebnisse mit Teilhabe-Relevanz formulieren
- Bio-psycho-soziale Dimensionen von Reha-Ergebnissen beachten
- Sets an differentiellen (≠ singulären) Ergebnisqualitätsindikatoren bestimmen
- Unterschiedliche Datenquellen der Erfolgsbewertung wählen
- Sowohl formale wie auch inhaltliche Kriterien der Erfolgsbewertung definieren
- Bei der Frage der Anzahl der Messzeitpunkte zwischen Ökonomie und methodischer Güte abwägen
- Patienteneigenschaften (risiko)adjustieren
- Sowohl generische als auch indikationsspezifische Instrumente nutzen
- Den Objektivitätsanspruch auch an die Behandler, die Laborwerte und die Körperfunktionsparameter stellen
- Bei Mehrpunkterhebungen zur Veränderungsmessung die Ausgangslage nicht vergessen
- Neue Möglichkeiten der Datenzugangs sowie der Datenauswertung ausschöpfen
- Randomisierungen zum Wirksamkeitsnachweis möglich machen

Der Workshop zur Ergebnisqualität ergab zentrale Impulse für die Weiterentwicklung der Reha-Qualitätssicherung im Bereich der Ergebnisqualität. Es wurden – kurz gefasst – die in Tabelle 3 aufgelisteten Forderungen an die Messung der Ergebnisqualität in der Reha-Qualitätssicherung der Rentenversicherung gestellt.

Die Forderungen an die Messung der Ergebnisqualität können folgendermaßen zusammengefasst werden: Eine Aussage über die Zielerreichung in der Rehabilitation kann nur auf der Grundlage zuvor festgelegter Rehabilitations-Ziele getroffen werden. Bei der Zieldefinition und der Zielüberprüfung muss das bio-psycho-soziale

Modell unter besonderer Beachtung der Teilhabe-Relevanz angemessen berücksichtigt werden. Die Erfolgsbewertung darf nicht nur formale Dimensionen betreffen (Beschäftigungsverhältnis ja oder nein), sondern muss sich auch an inhaltlichen Kriterien der Rehabilitation und der beruflichen sowie sozialen Integration des Rehabilitanden orientieren (z.B. Entwicklung von Bewältigungsmechanismen von Arbeitsanforderungen und Kompensationsmöglichkeiten des Rehabilitanden). Bei der Ausgestaltung des Instrumentariums zur Erhebung der Ergebnisqualität ist ein kombiniertes Indikatorenset unter Berücksichtigung unterschiedlicher Datenquellen zu verwenden. Dabei haben sich alle Zugangswege zur Ergebnisqualität dem Objektivitätskriterium zu stellen. Bei der Entscheidung über die Anzahl der Messzeitpunkte zur Veränderungsmessung ist eine Abwägung zwischen ökonomischen Gesichtspunkten und methodischen Gütekriterien vorzunehmen. Risikoadjustierende Verfahren zur statistischen Bereinigung des Effektes intervenierender Einflüsse auf die Patientenzufriedenheit und die Bewertung des Reha-Erfolges müssen angewendet werden. Das wichtigste zu berücksichtigende Patientenmerkmal für die Fallmix-Adjustierung ist die Eingangsbelastung des Rehabilitanden. Die Möglichkeiten moderner statistischer und testtheoretischer Verfahren sollten bei der Messung der Ergebnisqualität in einem höheren Ausmaß genutzt werden. Ein randomisierter Vergleich von unterschiedlichen Reha-Konzepten ist bei der Messung der Ergebnisqualität anzustreben. Für die Erfolgsdefinition ist sowohl der Reha-Erfolg als auch die Nachhaltigkeit des Reha-Erfolgs bedeutsam.

## 6. Literatur

**Bayerl, B.** (2009): Individuelle Rangfolgen – Kosten einsparen oder Qualität verbessern? Deutsches Ärzteblatt, 106, 17, 696-698.

**Beckmann, U., Grünbeck, P. & Mitschele, A.** (2002): Sozialmedizinische Leistungsbeurteilung und ihre Aussagekraft im Zusammenhang mit Ergebnissen aus der Patientenbefragung. DRV-Schriften, Band 33, 304-305.

**Bengel, J., Wirtz, M. & Zwingmann, C.** (2008). Diagnostische Verfahren in der Rehabilitation. Göttingen: Hogrefe-Verlag.

**Biefang, S., Potthoff, P. & Schliehe, F.** (1999): Assessmentverfahren für die Rehabilitation. Göttingen: Hogrefe-Verlag.

**Bijani, J.** (2005): Erleben der Diagnoseeröffnung bei einer Krebserkrankung. Dissertationsschrift an der Medizinischen Fakultät der Rheinischen Friedrich-Wilhelms-Universität zu Bonn.

**Bleich, C., Pawils, S., Koch, U.** (2006): Patientenzufriedenheit: Abhängigkeit von Alter und Geschlecht. In: Hinz, A. & Decker, O. (Ed.): Gesundheit im gesellschaftlichen Wandel – Altersspezifik und Geschlechterrollen. Gießen: Psychosozial-Verlag, 179-199.

**Bleich, C., Harfst, T., Andreas, S., Kawski, S., Koch, U. & Schulz, H.** (2006): Ist für den Klinikvergleich der Patientenzufriedenheit in der stationären psychischen/psychosomatischen Rehabilitation eine Risikoadjustierung notwendig? DRV-Schriften, Band 64, 210-211.

**Brandes, I. & Walter, U.** (2007): Rehabilitative Versorgung von Personen mit Migrationshintergrund. Selbstmanagementprozesse in der Rehabilitation – Hemmende und fördernde Faktoren. Herbsttagung der Korefo am 24.11.2007 in Hannover.

**Brode, T.** (2002): PatientInnenerwartungen und PatientInnenzufriedenheit im interkulturellen Vergleich: Ergebnisse der Berliner MigrantInnenstudie. Das Gesundheitswesen, 64, 476-485.

**Brode, T.** (2003): PatientInnenerwartungen und PatientInnenzufriedenheit im interkulturellen Vergleich: Dokumentation der Fachtagung „Frauen, Migration und Gesundheit“ am 22. Mai 2003 in Heidelberg, 21-35.

**Bührlen-Armstrong, B., de Jager, U., Schochat, T. & Jäckel, W. H.** (1998): Patientenzufriedenheit in der Rehabilitation muskuloskelettaler Erkrankungen – Einfluss von Merkmalen der Patienten, der Behandlung, des Messzeitpunkts und Zusammenhang mit dem Behandlungsergebnis. *Die Rehabilitation*, 37 (1), 38-46.

**Burchardi-Grabic, N., Farin, E., Follert, P. & Jäckel, W.H.** (2002): Bewertung des Behandlungsergebnisses in der orthopädischen Rehabilitation aus Arzt- und Patientensicht. *DRV-Schriften*, Band 33, 415-417.

**Deck, R., Zimmermann, M. & Raspe, H.** (2000): Zufriedene oder unzufriedene Patienten: Welche Merkmale bestimmen die Zufriedenheit mit einer Rehabilitations-Maßnahme? *DRV-Schriften*, Band 20, 39-40.

**Deck, R.** (2004): Welchen Einfluss haben Reha-Informationen auf Reha-Erwartungen und Reha-Ergebnisse? In: Petermann, F. & Ehlebracht-König, I. (Ed.): *Motivierung, Compliance und Krankheitsbewältigung*. Regensburg: Roderer-Verlag, 63-76.

**Deck, R.** (2006): Erwartungen und Motivationen von Patienten in der medizinischen Rehabilitation. Die Motivationseinschätzung von Patienten und ihren behandelnden Ärzten und ihr jeweiliger Einfluss auf die Rehabilitationsergebnisse. In: Nübling, R., Muthny, F. A. & Bengel J. (Ed.): *Reha-Motivation und Behandlungserwartung*. Bern: Verlag Hans Huber, Hogrefe-Verlag, 76-95.

**Deutsche Rentenversicherung Bund** (2007a): Rahmenkonzept zur medizinischen Rehabilitation in der gesetzlichen Rentenversicherung. Deutsche Rentenversicherung Bund,

Geschäftsbereich Sozialmedizin und Rehabilitation (Ed.).

**Deutsche Rentenversicherung Bund** (2007b): Der ärztliche Reha-Entlassungsbericht – Leitfaden zum einheitlichen Entlassungsbericht in der medizinischen Rehabilitation der gesetzlichen Rentenversicherung. Deutsche Rentenversicherung Bund, Geschäftsbereich Sozialmedizin und Rehabilitation (Ed.).

**Deutschen Rentenversicherung Bund** (2006): Effektivität und Effizienz der Rehabilitation – Aspekte, Entwicklungen, Erfolge. Tagungsband zum Reha-Forum der Deutschen Rentenversicherung Bund (Ed.) am 2. und 3. November 2006 in Berlin.

**Dörner, U.** (2004): Die Bedeutung von Krankheitsverarbeitungsprozessen für den Rehabilitations-Erfolg kardiologischer Patienten. Dissertationsschrift des Instituts für Medizinische Psychologie am Universitätsklinikum Münster.

**Dohnke, B.** (2004): Emotionale und motivationale Effekte von Erwartungen und erwartungsbezogenen Erfahrungen im Rehabilitations-Prozess. Eine Untersuchung von Effekten erwartungsgemäßer und erwartungsdiskrepanter Behandlungsergebnisse. Dissertationsschrift des Instituts für Rehabilitationswissenschaften an der Humboldt-Universität zu Berlin.

**Dorenburg, U., Huck-Langer, K., Nischan, P. & Winnefeld, M.** (2001): Kontinuierliche, klinikvergleichende Patientenbefragung im Reha-Qualitätssicherungsprogramm der Rentenversicherung: Konzept, Methodik, Erfahrungen. In: Satzinger, W., Trojan, A., Kellermann-Mühlhoff, P. (Ed.): *Patientenbefragungen in Krankenhäusern. Konzepte, Methoden, Erfahrungen*. Sankt Augustin: Asgard-Verlag, 361-369.

**Egner, U. & Grünbeck, P.** (2003): Analyse der Reha-Verlaufs-Daten zu den Folgen der Sucht und zum Verlauf nach der Rehabilitation. *Sucht aktuell*, 2, 29-32.

**Evangelou, E., Tsianos, G. & Loannidis, J.P.A.** (2008): Doctors' versus patients' global assessments of treatment effectiveness: Empirical survey of diverse treatments in clinical trials. *BMJ*, 336, 1287-1290.

**Farin, E.** (2008): Patientenorientierung und ICF-Bezug als Herausforderungen für die Ergebnismessung in der Rehabilitation. *Die Rehabilitation*, 47, 67-76.

**Farin, E.** (2005): Die Anwendung hierarchischer linearer Modelle für Einrichtungsvergleiche in der Qualitätssicherung und Rehabilitationsforschung. *Die Rehabilitation*, 44, 3, 157-164.

**Farin, E., Glattacker, M., Follert, P., Kuhl, C. & Jäckel, W.H.** (2004): Effekte und Prädiktoren des Outcomes in der Rehabilitation muskuloskelettaler Erkrankungen. *Phys Med Rehab Kuror*, 14, 123-133.

**Glattacker, M. & Jäckel, W.H.** (2007): Evaluation der Qualitätssicherung – aktuelle Datenlage und Konsequenzen für die Forschung. *Gesundheitswesen*, 69, 277-283.

**Gönner, S., Bischoff, C. & Limbacher, K.** (2008): Rentenantragsteller und Langzeitarbeitsunfähige – eine nicht behandelbare Klientel? Differenzielle Ergebnisqualität psychosomatischer Rehabilitation bei Patienten mit sozialmedizinischer Fragestellung. *Klinische Verhaltensmedizin und Rehabilitation*, 79, 61-69.

**Grichting, E.** (2001): Die Therapie im Rückblick – ihre Wirkung aus Sicht der ehemaligen Klientinnen und Klienten vor dem Hintergrund der Lebenssituation nach Austritt. In: Tagungsdokumentation zum XI. Symposium „Was uns hilft und zufrieden stellt – Therapieerfolg und -zufriedenheit aus Klientensicht“ des Forschungsverbundes Stationäre Suchttherapie (FOS) am 26. Oktober 2001 in Zürich, 7-13.

**Ghafuri, S.** (2006): Evaluation der Sozialmedizinischen

Leistungsbeurteilung. Dissertationsschrift des Fachbereichs Medizin an der Universität Hamburg.

**Hansis, M.L. & Hart, D.** (2001): Medizinische Behandlungsfehler in Deutschland. Reihe: Gesundheitsberichterstattung des Bundes, 4, Berlin: Robert-Koch-Institut (Ed.).

**Hansmeyer, T., Müller-Fahrnow, W. & Klosterhuis, H.** (1997): Kosten-Nutzen-Bewertung der medizinischen Rehabilitation der RV und die Auswirkungen des WFG. *DAngVers*, 5, 226-233.

**Huhn, B.** (2000): Untersuchung zu medizinpsychologischen Aspekten bei arbeitslosen Patienten. Dissertationsschrift der medizinischen Fakultät der Universität zu Halle.

**Irle, H., Korsukéwitz, C., Klosterhuis, H. & Grünbeck, P.** (2002): Psychische Belastung: Prävalenz und zielorientierte Beeinflussung in der medizinischen Rehabilitation. *DRV-Schriften*, Band 33, 277-279.

**Jäckel, W. H.** (2002): Qualitätssicherung in der medizinischen Rehabilitation. In: Tagungsband zum Kongress „Qualitätsmanagement in der Rehabilitation“ am 14. Juni 2002 in Bad Herrenalb: Schriftenreihe der Hans-Ruland-Stiftung für Präventions- und Rehabilitationsforschung, 13, Rotunde/Bad Herrenalb: Selbstverlag, 26-55.

**Klein, K., Englert, N., Farin, E., Gerdes, N. & Jäckel, W.H.** (2002): Patientenzufriedenheit und Ergebnisqualität: Der Zusammenhang zwischen Zufriedenheit mit der Rehabilitation und dem Rehabilitations-Erfolg in kardiologischen und orthopädischen Rehabilitations-Kliniken. *DRV-Schriften*, Band 33, 117-119.

**Klosterhuis, H., Zollmann, P. & Grünbeck, P.** (2004): Verlauforientierte Auswertungen zur Rehabilitation – Aktuelle Ergebnisse aus der Reha-Statistik-Datenbasis. *Deutsche Rentenversicherung*, 59, 5, 287-296.

**Klotz, T., Zumbé, R., Velmans, R. & Engelmann, U.** (1996): Die Bestimmung der Patientenzufriedenheit als Teil des Qualitätsmanagements im Krankenhaus. Deutsche medizinische Wochenschrift, 121, 889-895.

**Koch, K. & Dinstios, Ch.-M.** (2008): Die Bewertung von Kosten und Nutzen in der Medizin: Ein Vorschlag. Die Ersatzkasse, 3, 114-117.

**Kulick, B., Florain, L. & Enge, D.** (2002): Rentenantragsteller als Problemgruppe in der Rehabilitation. DRV-Schriften, Band 33, 296-298.

**Lampert, T. & Ziese, T.** (2005): Armut, soziale Ungleichheit und Gesundheit. Expertise des Robert-Koch-Instituts zum 2. Armuts- und Reichtumsbericht der Bundesregierung. Schriftenreihe: Lebenslagen in Deutschland, Bonn: BMGS.

**Lezzoni, L.** (2003): Risk adjustment for measuring health care outcomes. Chicago: Health Administration Press.

**Linden, M.** (2006): Effektivität und Effizienz der Rehabilitation aus der Sicht von Patient und Arzt. In: Effektivität und Effizienz der Rehabilitation – Aspekte, Entwicklungen, Erfolge. Reha-Forum der Deutschen Rentenversicherung Bund (Ed.) am 2. und 3. November 2006 in Berlin, 110-133.

**Mans, E. J.** (1995): Die Patientenzufriedenheit als Kriterium der Qualitätssicherung in der stationären psychosomatischen Rehabilitation. Gesundheitswesen, 57, 63-68.

**Meyer, S.** (2003): Zählt die Zufriedenheit des Patienten? Frankfurt am Main: Verlag Peter Lang.

**Mielck, A., Satzinger, W. & Apelt, P.** (1993): Zufriedenheit mit der ambulant-ärztlichen Versorgung: Unterschiede nach Schulbildung in Görlitz. Journal für Sozial- und Präventivmedizin, 38, 3, 142-147.

**Mitschele, A. & Klosterhuis, H.** (2007): Hat das Alter einen Einfluss auf die Rehabilitandenzufriedenheit? Indikationsspezifische Auswertungen. DRV-Schriften, Band 72, 526-528.

**Mittag, O., Meyer, T., Glaser-Möller, N., Matthis, C. & Raspe, H.** (2006): Vorhersage der Erwerbstätigkeit in einer Bevölkerungsstichprobe von 4225 Versicherten der LVA über einen Prognosezeitraum von 5 Jahren mittels einer kurzen Skala (SPE-Skala). Gesundheitswesen, 68, 294-302.

**Modestin, J., Hanselmann, F., Rüesch, P., Grünwald, H. & Meyer, P.** (2003): Der Züricher Fragebogen zur Patientenzufriedenheit in der Psychiatrie: Entwicklung und Charakteristika. Schweizer Archiv für Neurologie und Psychiatrie, 154/3, 127-138.

**Möske, M., Schneider, J., Koch, U. & Schulz, H.** (2008): Beeinflusst der türkische Migrationshintergrund das Behandlungsergebnis? Ergebnisse einer prospektiven Versorgungsstudie in der stationären Rehabilitation von Patienten mit psychischen / psychosomatischen Störungen. Psychother Psych Med, 58, 176-182.

**Mohr, B., Gräf, T., Forster, M., Krohn-Grimberghe, B., Kurzeja, R., Mantel, F., Thomsen, M. & Hampel, P.** (2008): Der Einfluss von Depressivität und Geschlecht auf den Rehabilitationserfolg bei chronischen Rückenschmerz: Eine Pilotstudie. Die Rehabilitation, 47, 284-298.

**Nischan, P. & Klosterhuis, H.** (2002): Ergebnisse der Patientenbefragung im Qualitätssicherungsprogramm der RV im Verlauf. DRV-Schriften, Band 33, 194-195.

**Nischan, P., Klosterhuis, H. & Rütten-Köppel, E.** (2003): Führt eine gute Prozessqualität auch zu einer hohen Patientenzufriedenheit? Analyse des Zusammenhanges zwischen klinikbezogenen Ergebnissen des Peer-Review-Verfahrens und der Patientenbefragung. DRV-

Schriften, Band 40, 156-157.

**Nübling, R., Löschmann, C., Kohl, C.F.R., Rundel, M., Wille, J., Witzke, M. & Schmidt, J.** (2003): Benchmarking in der medizinischen Rehabilitation – Möglichkeiten im Rahmen eines trägerinternen QM. DRV-Schriften, Band 40, 173-175.

**Olandt, H. & Krentz, H.** (1998): Patientenzufriedenheit: Subjektive Qualitätswahrnehmung der Patienten und Erfolgsfaktor für Kliniken. Gesundheitswesen, 60, 721-728.

**Plath, H.-E. & Blaschke, D.** (1999): Probleme der Erfolgsfeststellung in der beruflichen Rehabilitation. Sonderdruck aus: Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, 32. Jg., 61-69.

**Rabung, S.** (2005): Outcome-Kriterien am Beispiel der (stationären) medizinischen Rehabilitation bei Abhängigkeitserkrankungen. Dokumentation des Workshops „Wirkungsorientierte Steuerung in der ambulanten Suchtkrankenhilfe“ am 7. Dezember 2005 in der Hamburgischen Landesstelle für Suchtfragen e. V., Zentrum für Interdisziplinäre Suchtforschung der Universität Hamburg (ZIS), 29-38.

**Schmidt, J.** (1996): Kurz- und langfristige Veränderungen gesundheitsrelevanter Merkmalsbereiche (Risikofaktoren, Verhaltensweisen, Einstellungen) durch psychosomatische Rehabilitationsbehandlungen. In: DRV-Schriften, Band 6, 160-162.

**Steffen, P., & Pfaff, H.** (2004): Prädiktoren von Patientenzufriedenheit im Krankenhaus. Gesundheitswesen, 66, 8/9, 644.

**Raspe, Bühner, M. & Ziegler, M.** (2006): Abschlussbericht zum Projekt Weiterentwicklung der Patientenbefragung: Psychometrische Ergebnisse des Pretests zur Weiterentwicklung des Fragebogens zur Patientenzufriedenheit und gesundheitlichen Effekten im Qualitätssicherungs-

programm der Deutschen Rentenversicherung Bund.

**Raspe, Bühner, M., Ziegler, M. & Wegscheider,** (2006): Abschlussbericht zum Projekt Weiterentwicklung der Rückmeldung zur Patientenbefragung im Qualitätssicherungsprogramm der Deutschen Rentenversicherung Bund – Vorschlag für eine modifizierte Rückmeldekonzepktion.

**Raspe, H., Weber, U., Voigt, S., Kosinski, A. & Petras, H.** (1997): Qualitätssicherung durch Patientenbefragungen in der medizinischen Rehabilitation: Wahrnehmungen und Bewertungen von Rehastrukturen und -prozessen (Rehabilitandenzufriedenheit). Die Rehabilitation; 36, 31-42.

**Rehfeld, U.** (2008): Zu Stand und Fortentwicklung des statistischen Berichtswesens der gesetzlichen Rentenversicherung. DRV-Schriften, Band 6, 507-526.

**Reiter, A., Fischer, B., Kötting, J., Geraedts, M., Jäckel, W.H., Berlag, H. & Döbler, K.** (2007): QUALIFY: Ein Instrument zur Bewertung von Qualitätsindikatoren. Düsseldorf: BQS gGmbH.

**Rundel, M., Koh, C. F. R., Nübling, R., Staffanowski, A., Körner, M., Bishop, A. & Löschmann, C.** (2004): Benchmarking der Patientenbefragung als Bestandteil des QM. DRV-Schriften, Band 52, 212-213.

**Schubert, M., Behrens, J., Höhne, A., Schaepe, C. & Zimmermann, M.** (2007): Den Arbeitsmarkt als Frühberentungsgrund – zur Spezifik arbeitsmarktbedingter ErwerbsminderungsrentnerInnen. DRV-Schriften, Band 72, 202-204.

**Tesmer, I., Thode, N. & Radoschewski, F.M.** (2006): Geschlechterunterschiede in der Bewertung von Therapie- und Behandlungsmaßnahmen in der orthopädischen Rehabilitation. In: DRV-Schriften, Band 64, 212-213.

**Weber, A., Hörmann, G. & Heipertz, W.** (2007): Arbeitslosigkeit und Gesundheit aus sozialmedizinischer Sicht. Dtsch Arztebl, 104 (43), A, 2957-2962.

**Weber, H.** (2005): Persönlichkeit und Gesundheit. In: Schwarzer, R. (Ed.): Gesundheitspsychologie, Göttingen: Hogrefe-Verlag.

**Wettach, R.** (2001): Therapiebeurteilung und -zufriedenheit von Klienten stationärer Suchteinrichtungen im Lichte ausgewählter Merkmale ihrer Persönlichkeit. In: Tagungsdokumentation zum XI. Symposium „Was uns hilft und zufrieden stellt – Therapieerfolg und -zufriedenheit aus Klientensicht“ des Forschungsverbundes Stationäre Suchttherapie (FOS) am 26. Oktober 2001 in Zürich, 14-17.

**Widera, T. & Klosterhuis, H.** (2007a): Patientenorientierung in der Praxis – 10 Jahre Rehabilitandenbefragung im Rahmen der Reha-Qualitätssicherung der Rentenversicherung. RVaktuell, 54, 6, 177-182.

**Widera, T. & Beckmann, U.** (2007b): Zusammenhänge zwischen subjektivem und objektivem Rehabilitations-Erfolg: Zur Vorhersagbarkeit der Rentenantragstellung über Ergebnisse der Rehabilitandenbefragung und ärztliche Leistungsbeurteilungen. In: DRV-Schriften, Band 72, 200-202.

**Zielke, M.** (2006): Kosten-Nutzen-Analysen und Return of Investment der psychosomatischen Rehabilitation. In: Effektivität und Effizienz der Rehabilitation – Aspekte, Entwicklungen, Erfolge. Reha-Forum der Deutschen Rentenversicherung Bund (Ed.) am 2. und 3. November 2006 in Berlin, 110-133.

**Zwingmann, Ch., Moock, J. & Kohlmann, T.** (2005): Patientennahe Assessmentinstrumente in der deutschsprachigen Rehabilitationsforschung – Aktuelle Entwicklung aus dem Förderschwerpunkt „Rehabilitationswissenschaften“. Die Rehabilitation, 44, 307-315.

## Referenten und Ansprechpartner (alphabetisch)

### **Dr. med. Ulrike Beckmann**

Deutsche Rentenversicherung Bund  
Geschäftsbereich Sozialmedizin und Rehabilitation  
Bereich Reha-Qualitätssicherung, Epidemiologie  
und Statistik  
Leiterin des Referates  
Grundsatzfragen und Weiterentwicklung der  
Qualitätssicherung  
Ruhrstraße 2  
10704 Berlin  
Telefon: 030 - 865-31836  
E-Mail: [ulrike.beckmann@drv-bund.de](mailto:ulrike.beckmann@drv-bund.de)  
Internet: [www.deutsche-rentenversicherung-bund.de](http://www.deutsche-rentenversicherung-bund.de)

### **Dipl.-Soz. Wiss. Hans Dörning**

IQEM – Institut für Qualitätsmanagement und  
Qualitätssicherung in der Medizinischen  
Rehabilitation e.V.  
Geschäftsführer des Institutes  
Lavesstr. 80  
30159 Hannover  
Telefon: 0511 - 53091-10  
E-Mail: [info@iqem.de](mailto:info@iqem.de), [doerning@iqem.de](mailto:doerning@iqem.de)  
Internet: [www.iqem.de](http://www.iqem.de)

### **Uwe Egner**

Deutsche Rentenversicherung Bund  
Abteilung Rehabilitation  
Fachbereichsleiter Zentrale und Grundsatzaufgaben  
Koordinator der Reha-Qualitätssicherung der  
Rentenversicherung  
Hohenzollerndamm 46/47  
10704 Berlin  
Telefon: 030 - 865 82750  
E-Mail: [uwe.egner@drv-bund.de](mailto:uwe.egner@drv-bund.de)  
Internet: [www.deutsche-rentenversicherung-bund.de](http://www.deutsche-rentenversicherung-bund.de)

**PD Dr. phil. Dipl.-Psych. Erik Farin**

Universitätsklinikum Freiburg  
Abteilung Qualitätsmanagement und Sozialmedizin  
Engelbergerstr. 21  
79106 Freiburg  
Telefon: 0761 - 270-7443 /-7447  
E-Mail: erik.farin@uniklinik-freiburg.de  
Internet: www.aqms.de

**Dr. med. Eberhard Virtus Grosch**

Deutsche Rentenversicherung Braunschweig-Hannover  
Ärztlicher Dienst (Leit. Arzt)  
Koordinator der Reha-Qualitätssicherung der  
Rentenversicherung  
Lange Weihe 2  
30880 Laatzen  
Telefon: 0511 - 8293201  
E-Mail: eberhard.grosch@drv-bsh.de  
Internet: www.drv-bsh.de

**Dr. rer. physiol. Dipl.-Psych. Hans-Günter Haaf**

Deutsche Rentenversicherung Bund  
Geschäftsbereich Sozialmedizin und Rehabilitation  
Bereich Rehabilitationswissenschaften  
Leiter des Referats Gesundheitsökonomie  
und Systemfragen  
Ruhrstraße 2  
10704 Berlin  
Telefon: 030 - 865-39333  
E-Mail: dr.hans-guenter.haaf@drv-bund.de  
Internet: www.deutsche-rentenversicherung-bund.de

**Dr. med. Here Klosterhuis**

Deutsche Rentenversicherung Bund  
Geschäftsbereich Sozialmedizin und Rehabilitation  
Leiter des Bereichs Reha-Qualitätssicherung,  
Epidemiologie und Statistik  
Koordinator der Reha-Qualitätssicherung der  
Rentenversicherung  
Ruhrstraße 2  
10704 Berlin  
Telefon: 030 - 865 31234  
E-Mail: Dr.med.Here.Klosterhuis@drv-bund.de  
Internet: www.deutsche-rentenversicherung-bund.de

**Berthold Lindow**

Deutsche Rentenversicherung Bund  
Geschäftsbereich Sozialmedizin und Rehabilitation  
Bereich Reha-Qualitätssicherung, Epidemiologie  
und Statistik  
Leiter des Referates  
Durchführung der Qualitätssicherung  
Ruhrstraße 2  
10704 Berlin  
Telefon: 030 - 865-35958  
E-Mail: berthold.lindow@drv-bund.de  
Internet: www.deutsche-rentenversicherung-bund.de

**Dr. phil. Dipl.-Psych. Thorsten Meyer**

Universitätsklinikum Schleswig-Holstein  
Institut für Sozialmedizin  
Campus Lübeck  
Beckergrube 43-47  
23552 Lübeck  
Telefon: 0451 - 799-2526  
E-Mail: thorsten.meyer@uk-sh.de  
Internet: www.sozmed.uni-luebeck.de

**Prof. Dr. PH Dipl.-Soz. Matthias Morfeld**

University of Applied Sciences Magdeburg and Stendal  
Osterburger Str. 25  
39576 Stendal  
Telefon: 03931 - 2187-4847  
E-Mail: matthias.morfeld@hs-magdeburg.de  
Internet: www.hs-magdeburg.de

**PD Dr. phil. Dipl.-Psych. Holger Schulz**

Universitätsklinikum Hamburg-Eppendorf,  
Zentrum für Psychosoziale Medizin  
Institut und Poliklinik für Medizinische Psychologie  
Martinistraße 52, S35  
20246 Hamburg  
Telefon: 040 - 42803 6806  
E-Mail: schulz@uke.uni-hamburg.de  
Internet: www.uke.de

**Heinz-Gert Verhorst**

Deutsche Rentenversicherung Westfalen  
Abteilung für Rehabilitation – Grundsatzreferat  
Koordinator der Reha-Qualitätssicherung der  
Rentenversicherung  
Gartenstraße 194  
48125 Münster  
Telefon 0251 - 238-2606  
E-Mail: heinz-gert.verhorst@drv-westfalen.de  
Internet: www.deutsche-rentenversicherung-westfalen.de

**Dr. phil. Dipl.-Psych. Teresia Widera**

Deutsche Rentenversicherung Bund  
Geschäftsbereich Sozialmedizin und Rehabilitation  
Bereich Reha-Qualitätssicherung, Epidemiologie  
und Statistik  
Ruhrstraße 2  
10704 Berlin  
Telefon: 030 - 865-39343  
E-Mail: teresia.widera@drv-bund.de  
Internet: www.deutsche-rentenversicherung-bund.de

**Prof. Dr. phil. Dipl.-Psych. Markus Wirtz**

Pädagogische Hochschule Freiburg  
Abteilung für Forschungsmethoden  
Kartäuserstr. 61b  
79117 Freiburg  
Telefon: 049 - 761 682-905  
E-Mail: markus.wirtz@ph-freiburg.de  
Internet: www.ph-freiburg.de

**Michael Zellner**

Deutsche Rentenversicherung Bayern Süd  
Leiter der Abteilung Kliniken  
Koordinator der Reha-Qualitätssicherung  
der Rentenversicherung  
Thomas-Dehler-Str. 3  
81729 München  
Telefon 089 - 6781-2713  
E-Mail: michael.zellner@drv-bayernsued.de  
Internet: drv-bayernsued.de

## Impressum

Herausgeber: Deutsche Rentenversicherung Bund  
Geschäftsbereich Sozialmedizin und  
Rehabilitation  
Bereich Reha-Qualitätssicherung,  
Epidemiologie und Statistik  
Ruhrstraße 2, 10709 Berlin

Satz und  
Gestaltung: Baxter & Baxter Werbeagentur GmbH  
Alt-Bornheim 26, 60385 Frankfurt am Main

1. Auflage (08/2009)