

---

*Carsten Maurischat, Matthias Morfeld, Thomas Kohlmann,  
Monika Bullinger (Hrsg.)*

---

**Lebensqualität: Nützlichkeit und Psychometrie  
des Health Survey SF-36/SF-12 in der medizinischen  
Rehabilitation**

---



PABST SCIENCE PUBLISHERS  
Lengerich, Berlin, Bremen, Miami,  
Riga, Viernheim, Wien, Zagreb

Bibliografische Information Der Deutschen Bibliothek  
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <<http://dnb.ddb.de>> abrufbar.

Das Werk, einschließlich aller seiner Teile, ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

© 2004 Pabst Science Publishers, D-49525 Lengerich  
Druck: KM Druck, D-64823 Groß Umstadt

ISBN 3-89967-142-2

## Inhaltsverzeichnis

<b>Vorwort</b>	Standardisierung der Outcomeerfassung <i>Carsten Maurischat, Matthias Morfeld, Thomas Kohlmann &amp; Monika Bullinger</i>	Seite 9
<b>Teil I: Der SF-36/SF-12 als Messinstrument</b>		
Kapitel 1	Der Health Survey SF-36/SF-12: Darstellung und aktuelle Entwicklungen <i>Monika Bullinger &amp; Matthias Morfeld</i>	Seite 15
Kapitel 2	Analysen zum Strukturmodell des SF-36/SF-12 – eine Übersicht <i>Carsten Maurischat &amp; Anja Krüger-Bödeker</i>	Seite 29
Kapitel 3	Zur Validität eines Transformationsalgorithmus' für die Fehlerwertersetzung bei psychosomatischen Patienten: Wer fehlende Werte im SF-12 ersetzt, wird mit mangelnder Vergleichbarkeit bestraft, aber mit höherer Stichprobenausschöpfung belohnt <i>Klaus Lang, Uwe Koch &amp; Holger Schulz</i>	Seite 49
<b>Teil II: Methodische Aspekte: Änderungssensitivität und Prädiktoren</b>		
Kapitel 4	Interpretation von Maßen zur Bestimmung von Effektgrößen <i>Rainer Leonhart</i>	Seite 65
Kapitel 5	Empirische Befunde zur Änderungssensitivität des SF-36/SF-12 <i>Wilmar Igl, Christian Zwingmann, Peter Schuck &amp; Hermann Faller</i>	Seite 79
Kapitel 6	SF-12 und SF-36 im Rahmen von Vorhersagen des Erfolges von Patientenschulungen – Eine Reanalyse gepoolter Daten <i>Marc Worbach, Andrea Reusch, Heiner Vogel &amp; Hermann Faller</i>	Seite 91

## Kapitel 5

### Empirische Befunde zur Änderungssensitivität des SF-36/SF-12

WILMAR IGL, CHRISTIAN ZWINGMANN, PETER SCHUCK &  
HERMANN FALLER

in Zusammenarbeit mit:

Prof. Dr. Manfred Beutel, Gießen; Prof. Dr. Wolfgang Beyer, Bad Füssing; Prof. Dr. Claus Bischoff, Bad Dürkheim; Prof. Dr. Emanuel Fritschka, Bad Brückenau; Dr. Dirk Gustson, Bad Neustadt/Saale; Dr. Alexander Kühn, Halle/Saale; Prof. Dr. Friedhelm Lamprecht, Hannover; Dr. Jürgen Schmidt, Karlsruhe; Dr. Roland Vauth, Basel; Prof. Dr. Claus-W. Wallesch, Magdeburg; Dr. Michael Wittmann, Bad Reichenhall

#### Theoretischer Hintergrund

In der Rehabilitation werden im Rahmen evaluativer Studien häufig Patientenfragebogen zu mehreren Messzeitpunkten eingesetzt, um auf der Grundlage von Prä-Post-Differenzen die Effekte von Rehabilitationsmaßnahmen aus Patientensicht zu erfassen. Dies kann nur gelingen, wenn die verwendeten Fragebogenskalen in der Lage sind, Veränderungen über die Zeit abzubilden, d.h. wenn es sich um änderungssensitive Instrumente handelt. Die kritische und vergleichende Überprüfung der Änderungssensitivität für die gegenwärtig eingesetzten deutschsprachigen generischen Assessmentinstrumente, zu denen auch der SF-36-Fragebogen und dessen Kurzform SF-12 [1] zählen, wurde jedoch bisher vernachlässigt. Im vorliegenden Beitrag sollen daher empirische Ergebnisse bezüglich der Änderungssensitivität berichtet werden, die das Ergebnis einer Reanalyse von SF-36- bzw. SF-12-Daten geeigneter Projekte aus der ersten Förderphase (1998-2001) des gemeinsamen Förderschwerpunkts „Rehabilitationswissenschaften“ des Bundesministeriums für Bildung

---

Igl, W., Zwingmann, C., Schuck, P. & Faller, H. (2004). Empirische Befunde zur Änderungssensitivität des SF-36/SF-12. In C. Maurischat, M. Morfeld, T. Kohlmann & M. Bullinger (Hrsg.), *Lebensqualität: Nützlichkeit und Psychometrie des Health Survey SF-36/SF-12 in der medizinischen Rehabilitation* (S. 79–90). Lengerich: Pabst.

und Forschung (BMBF) und der Deutschen Rentenversicherung sind. Die dargestellten Ergebnisse können als Entscheidungshilfe für die Auswahl von Skalen zum Einsatz in rehabilitationsbezogenen Evaluationsstudien verwendet werden [6, 7].

Die Grundlage für das Konzept der Änderungssensitivität lieferten Kirshner und Guyatt [9] bereits Mitte der 1980er Jahre, als sie ein Rahmenmodell zur Kategorisierung von Messinstrumenten entsprechend ihres funktionalen Anwendungsziels vorstellten. Sie unterscheiden die Messziele:

1. Diskrimination, z.B. von klinischen Gruppen, um eine differentielle Zuweisung zu verschiedenen Behandlungsmodulen vornehmen zu können;
2. Prognose, z.B. des Behandlungserfolgs;
3. Evaluation, z.B. von Veränderungen durch Therapiemaßnahmen.

In Abhängigkeit vom Anwendungsziel des Messinstruments stehen unterschiedliche Gütekriterien bzw. Kennwerte im Vordergrund, um die Qualität eines Fragebogens zu beurteilen [9, 13, 14]. Im Fall der Evaluation kommt es vor allem darauf an, Veränderungen über die Zeit abzubilden. Psychometrische Kriterien, die bei Diskrimination und Prognose im Vordergrund stehen, nämlich Kriteriums- und Konstruktvalidität sowie Reliabilität, müssen im Fall der Evaluation um das Konzept der Änderungssensitivität ergänzt werden.

## Methodik

### Studiendesign und Kennwerte

Zur Bestimmung der Änderungssensitivität wurden verschiedene Methoden vorgeschlagen, welche unterschiedliche Studiendesigns voraussetzen (Messungen zu verschiedenen Messzeitpunkten an einer oder mehreren Stichproben) und aus den erhobenen Daten verschiedene Kennwerte bilden [10, 11, 12]. Im Folgenden werden nur dasjenige Studiendesign und diejenigen Kennwerte vorgestellt, die für die dargestellten Ergebnisse relevant sind.

Die vorliegenden Reanalysen basieren auf einem einfachen Prä-Post-Design, welches Messungen zu Beginn und am Ende einer Rehabilitationsmaßnahme erfordert. Es wird häufig zur Evaluation der Änderungssensitivität von Messinstrumenten angewendet, da es sich einfach implementieren lässt und die

Berechnung von Kennwerten ermöglicht, die sich gut mit anderen Studien vergleichen lassen, falls eine gleichartige Wirksamkeit unterstellt werden kann.

Eine wichtige Annahme dieses Designs ist, dass eine wirksame Intervention durchgeführt wird und reale Veränderungen bzw. Verbesserungen bei den Patienten auftreten. Sollte diese Voraussetzung nicht erfüllt sein, kann bei niedrigen Kennwerten für die Änderung bzw. nichtsignifikanten Ergebnissen nicht entschieden werden, ob dies auf mangelnde Wirksamkeit der Intervention oder unzureichende Änderungssensitivität des Instruments zurückzuführen ist.

Die erhaltenen Daten der Prä-Post-Messungen ermöglichen die Berechnung von Effektstärken zur Beurteilung der Änderungssensitivität [10, 11, 12]. Den berichteten Effektstärken liegt das Prinzip zu Grunde, dass eine Differenz zwischen zwei Mittelwerten  $M_{t1}$  und  $M_{t2}$  an einem Streuungsmaß  $SD$  standardisiert wird (vgl. Formel 1). Einzelne Effektstärkemaße unterscheiden sich dabei in der verwendeten Standardabweichung  $SD$  im Nenner.

$$ES = (M_{t2} - M_{t1}) / SD \quad (\text{Formel 1})$$

Die standardized effect size (SES, in angloamerikanischen Publikationen oft nur als „the“ effect size benannt) verwendet als Streuungsmaß die Standardabweichung  $SD_{t1}$  der Messwerte zum Prä-Messzeitpunkt. Dies hat zur Folge, dass bei Stichproben, deren Messwerte zum Prä-Messzeitpunkt homogen sind, größere Effektstärken resultieren. Der standardized response mean (SRM) hingegen setzt die Mittelwertdifferenz in Bezug zur Standardabweichung der Messwertdifferenzen  $SD_{t2-t1}$ . Deshalb wird diese Effektstärke zwar nicht von der Homogenität der Stichprobe zum Prä-Messzeitpunkt beeinflusst, aber von der Homogenität der Veränderungswerte. So ergeben sich höhere Effektstärken, wenn sich die Messwerte der Patienten in ähnlicher Weise zwischen den beiden Messzeitpunkten verändern. Hinsichtlich der Frage, welche der beiden Effektstärken zur Beurteilung von Änderungssensitivität zu bevorzugen ist, existieren unterschiedliche Meinungen [5, 8]. Auf dem gegenwärtigen Stand der Diskussion erscheint es jedoch empfehlenswert, stets mehrere Kennwerte zu berichten [7, 10].



### Stichproben und Auswertung

In die Reanalysen gingen Daten aus zehn Forschungsprojekten aus verschiedenen Verbänden des Förderschwerpunkts „Rehabilitationswissenschaften“ ein (s. Tab. 1). Da sich die Projekte hinsichtlich der eingesetzten Fragebogen unterscheiden, ergeben sich unterschiedliche Patientenzahlen für die Auswertung des SF-36 (N=2.827) und des SF-12 (N=794). Die Patienten weisen ein breites Spektrum an Störungen auf, von denen nur die Bereiche Orthopädie/Rheumatologie, Kardiologie und Psychosomatik (soweit vorhanden) dargestellt werden. Mangels spezifischer SF-12-Daten können für die Bereiche Orthopädie/Rheumatologie und Kardiologie streng genommen keine entsprechenden Ergebnisse zur Änderungssensitivität berichtet werden. Um trotzdem nicht auf Aussagen für diese wichtigen Indikationsbereiche verzichten zu müssen, wurden entsprechende Item-Daten aus dem SF-36-Datensatz extrahiert und eine Berechnung der SF-12-Summenskalen vorgenommen.

**Tabelle 1: Beschreibung der reanalysierten Stichproben**

MERKMAL	DATENSATZ SF-36	DATENSATZ SF-12
Geschlecht	1.639 (58%) Frauen 1.186 (42%) Männer	299 (38%) Frauen 495 (62%) Männer
Range (Alter)	18 bis 77 Jahre	18 bis 68 Jahre
M (Alter)	44 Jahre	45 Jahre
SD (Alter)	11 Jahre	10 Jahre
n (Orthopädie/Rheumatologie)	1.134 (40%)	(0%) extr.
n (Kardiologie)	70 (3%)	(0%) extr.
n (Psychosomatik)	1.369 (48%)	329 (41%)
n (Nephrologie)	(0%)	281 (35%) n.v.
n (Pneumologie)	(0%)	184 (23%) n.v.
n (Psychiatrie)	201 (7%) n.v.	(0%)
n (Neurologie)	53 (2%) n.v.	(0%)
N (gesamt)	2.827 (100%)	794 (100%)

(n.v.= Daten werden für den vorliegenden Beitrag nicht verwendet, extr.= Daten wurden aus dem SF-36-Datensatz extrahiert, da keine spezifischen SF-12-Daten vorliegen)

Die Indikationen werden getrennt analysiert, da Veränderungen indikations-spezifisch sein können. In die Berechnung der Kennwerte gehen aus methodischen Gründen nur diejenigen Patienten ein, die eine bereits als wirksam unterstellte Rehabilitationsmaßnahme (Standardbehandlung/usual care) erhalten haben, d.h. keine experimentellen Therapien. Weiterhin können von diesen Patienten nur jene berücksichtigt werden, bei denen mindestens bei Reha-Beginn und bei Reha-Ende Daten vorliegen. Es werden sowohl SES- als auch SRM-Kennwerte berichtet, die sich – wenn nicht anders angegeben – auf die Veränderungen zwischen Reha-Beginn und Reha-Ende beziehen. Wenn möglich, werden darüber hinaus auch SES- und SRM-Kennwerte für den Messzeitraum zwischen Reha-Beginn und der 3-Monats-Katamnese angegeben.

Längere Zeit vor Reha-Beginn erhobene Baseline-Messwerte oder Messwerte von Kontrollgruppen, die die methodischen Anforderungen zur Berechnung von weiteren Kennwerten ermöglicht hätten, lagen in den Reanalyse-Daten nicht vor. Die Daten wurden vor der Berechnung der Effektstärken umgepolt, so dass positive Werte günstigen und negative Werte ungünstigen Veränderungen entsprechen. Effektstärken, die im Text als statistisch bedeutsam bezeichnet werden, unterscheiden sich entsprechend ihres 95%-Konfidenzintervalls signifikant von Null.

Obwohl die Faustregel von Cohen [2], wonach Effektstärken mit einem Wert von ca. 0.20 als „klein“, Werte von ca. 0.50 als „mittel“ und Werte von ca. 0.80 als „groß“ bezeichnet werden, nicht ohne Bedenken auf Kennwerte der Änderungssensitivität angewandt werden können, wurden diese Bezeichnungen im Text dennoch verwendet. Dies ermöglicht es, den Wertebereich der Effektstärken darstellen zu können, ohne auf Zahlenwerte zurückgreifen zu müssen. Dies impliziert jedoch noch nicht unbedingt eine inhaltliche oder klinische Bedeutung.

Falls Cohens Faustregel dennoch als Anhaltspunkt zur Beurteilung der inhaltlichen Bedeutung von Effektstärkewerten herangezogen werden sollte, ist zu beachten, dass diese für Effektstärken, in welche die Streuung von Differenzwerten eingeht, wie z.B. in den SRM, geringere Aussagekraft besitzt als für Effektstärken, die Streuungen von Messwerten zu einem Messzeitpunkt verwenden, wie z.B. die SES. Dies ist darin begründet, dass die SES konzept-

tionell eher den von Cohen vorgeschlagenen Effektstärken für Wirksamkeitsstudien entspricht.

## Ergebnisse

### SES und SRM im Vergleich

Die Kennwerte SES und SRM, bezogen auf den Zeitraum zwischen Reha-Beginn und Reha-Ende, korrelieren in den untersuchten SF-36-Daten zu  $r(\text{SES}, \text{SRM})=0.94$  ( $n=141$  Wertpaare,  $p<0.01$ ) miteinander. Dabei beträgt die durchschnittliche absolute Abweichung der Kennwerte  $M(|\text{SES}-\text{SRM}|)=0.049$  bei einer Standardabweichung der Differenzen  $SD(|\text{SES}-\text{SRM}|)=0.048$ . Die Abweichung zwischen SES und SRM nimmt maximal einen Wert von  $\text{Max}(|\text{SES}-\text{SRM}|)=0.24$  an. Für die SF-12-Daten (einschließlich der extrahierten SF-36-Daten) ergeben sich ähnliche Ergebnisse mit  $r(\text{SES}, \text{SRM})=0.95$  ( $n=84$  Wertpaare,  $p<0.01$ ). Die mittlere Differenz  $M(|\text{SES}-\text{SRM}|)=0.039$  ist bei einer Standardabweichung der Differenzwerte  $SD(|\text{SES}-\text{SRM}|)=0.041$  wie für die SF-36-Daten als gering zu beurteilen. Der maximale Unterschied zwischen SES- und SRM-Kennwerten beträgt  $M(|\text{SES}-\text{SRM}|)=0.19$ . Auf eine differenzierte Darstellung der SES- und SRM-Kennwerte im Text wird daher verzichtet und stattdessen auf die Tabellen 2 und 3 verwiesen.

### SF-36

Während der Behandlung von orthopädischen/rheumatologischen Patienten ( $n_{\text{Min}}=523$ ,  $n_{\text{Max}}=867$ ) erweisen sich sowohl die „Körperliche Summenskala“ als auch die „Psychische Summenskala“ des SF-36 als schwach änderungssensitiv (s. Tab. 2). Kleine bis mittlere Effektstärken sind bei den Skalen „Schmerz“, „Vitalität“ und „Psychisches Wohlbefinden“ festzustellen. Das Item „Veränderung des Gesundheitszustandes“ übertrifft diese Werte jedoch bei weitem. Die Effektstärken der übrigen Skalen bewegen sich in einem niedrigen Bereich.

Hinsichtlich der Änderungssensitivität für Veränderungen vom Beginn der Rehabilitation bis zur 3-Monats-Katamnese liegen bei einer Subgruppe von orthopädischen/rheumatologischen Patienten ( $n_{\text{Min}}=61$ ,  $n_{\text{Max}}=63$ ) Hinweise dafür vor, dass die Skalen ähnlich geeignet sind wie für den Zeitraum von

Reha-Beginn bis Reha-Ende. Die „Körperliche Summenskala“ zeigt dabei tendenziell höhere, die „Psychische Summenskala“ tendenziell niedrigere, nicht signifikant von Null verschiedene Werte. Das Item „Veränderung des Gesundheitszustandes“ scheint in diesem Messzeitraum Veränderungen sogar noch besser zu erfassen als zwischen Reha-Beginn und Reha-Ende.

Für kardiologische Patienten ( $n_{\text{Min}}=27$ ,  $n_{\text{Max}}=59$ ) können nur ungefähre Aussagen vorgenommen werden, da in den Daten nur wenige Messungen zu Reha-Beginn und Reha-Ende vorliegen. Die „Körperliche Summenskala“ erscheint hier ebenso wie die „Psychische Summenskala“ mittelgradig änderungssensitiv. Auch die Skalen „Körperliche Rollenfunktion“, „Emotionale Rollenfunktion“, „Vitalität“ sowie das Item „Veränderung des Gesundheitszustandes“ liegen in einem Bereich mittlerer Änderungssensitivität. „Psychisches Wohlbefinden“ und „Schmerz“ scheinen hingegen weniger geeignet, um bei kardiologischen Patienten Effekte am Ende der Rehabilitationsmaßnahme abzubilden. Die übrigen Skalen variieren in einem niedrigen Bereich und weichen statistisch nicht signifikant von Null ab.

Während der Rehabilitation von psychosomatischen Patienten ( $n_{\text{Min}}=75$ ,  $n_{\text{Max}}=166$ ) zeigen sich deutliche Unterschiede zwischen der „Körperlichen Summenskala“ und der „Psychischen Summenskala“, wobei die erste deutlich weniger änderungssensitiv ist. Die Skala „Allgemeine Gesundheitswahrnehmung“ scheint bei psychosomatischen Rehabilitanden von größerer Bedeutung zu sein als bei den beiden anderen Indikationen. „Vitalität“ und „Psychisches Wohlbefinden“ liegen in einem Bereich mittlerer Änderungssensitivität. Das Item „Veränderung des Gesundheitszustandes“ erweist sich hier wie bereits bei orthopädischen/rheumatologischen Patienten am änderungssensitivsten, während die übrigen Skalen nur schwach änderungssensitiv sind.

Für Veränderungen zwischen Reha-Beginn und dem 3-Monats-Follow-Up liegen bei einer Subgruppe von psychosomatischen Patienten ( $n_{\text{Min}}=72$ ,  $n_{\text{Max}}=137$ ) nahezu identische Effektstärken vor wie zwischen Reha-Beginn und Reha-Ende. Darauf weisen die Kennwerte sowohl der Summenskalen als auch mehrerer Einzelskalen hin.

**Tabelle 2: Änderungssensitivität des SF-36 für verschiedene Indikationen im Veränderungszeitraum zwischen Reha-Beginn und Reha-Ende**

SF-36 Skalen	ORTHOPÄDIE/RHEUMATOLOGIE n=[523;867]						KARDIOLOGIE n=[27;59]						PSYCHOSOMATIK n=[75;166]					
	SES			SRM			SES			SRM			SES			SRM		
	MW	UG	OG	MW	UG	OG	MW	UG	OG	MW	UG	OG	MW	UG	OG	MW	UG	OG
KSK	0.23*	0.14	0.31	0.30*	0.21	0.39	0.41*	0.01	0.80	0.47*	0.07	0.87	0.29*	0.05	0.52	0.32*	0.09	0.55
PSK	0.22*	0.13	0.30	0.29*	0.21	0.38	0.34	-0.05	0.72	0.50*	0.09	0.89	0.59*	0.34	0.83	0.53*	0.29	0.78
KoFu	0.15*	0.07	0.23	0.23*	0.15	0.31	0.16	-0.19	0.52	0.33	-0.04	0.69	0.20*	0.03	0.37	0.30*	0.12	0.47
KoRo	0.29*	0.20	0.37	0.28*	0.20	0.37	0.64*	0.25	1.03	0.61*	0.22	1.00	0.27*	0.07	0.47	0.28*	0.08	0.48
Schm	0.32*	0.25	0.39	0.43*	0.36	0.50	0.21	-0.05	0.47	0.32*	0.05	0.58	0.16*	0.01	0.32	0.20*	0.04	0.35
AGes	0.14*	0.06	0.22	0.17*	0.09	0.25	0.22	-0.15	0.58	0.23	-0.13	0.59	0.35*	0.17	0.52	0.37*	0.20	0.54
Vita	0.36*	0.28	0.43	0.43*	0.35	0.51	0.40*	0.04	0.76	0.59*	0.21	0.96	0.56*	0.38	0.73	0.60*	0.42	0.78
SoFu	0.16*	0.09	0.24	0.20*	0.13	0.28	0.14	-0.21	0.49	0.18	-0.17	0.53	0.31*	0.11	0.50	0.32*	0.12	0.51
EmRo	0.14*	0.05	0.22	0.15*	0.07	0.24	0.51*	0.10	0.90	0.66*	0.24	1.07	0.26*	0.07	0.45	0.29*	0.10	0.47
Psych	0.31*	0.23	0.39	0.39*	0.31	0.47	0.24	-0.12	0.60	0.38*	0.01	0.74	0.53*	0.36	0.71	0.55*	0.37	0.72
VerGes	0.82*	0.73	0.91	0.74*	0.65	0.83	0.39*	0.02	0.75	0.50*	0.12	0.87	0.83*	0.64	1.02	0.69*	0.50	0.87

(MW=Mittelwert des Konfidenzintervalls [95%]; Werte, deren Konfidenzintervall Null nicht einschließt, sind mit einem Stern \* markiert; UG=Untergrenze; OG=Obergrenze; KSK=Körperliche Summenskala, PSK=Psychische Summenskala; KoFu=körperliche Funktionsfähigkeit, KoRo=körperliche Rollenfunktion, Schm=Schmerz, AGes=allgemeine Gesundheitswahrnehmung, Vita=Vitalität, SoFu=soziale Funktionsfähigkeit, EmRo=emotionale Rollenfunktion, Psych=psych. Wohlbefinden, VerGes=Gesundheitsveränderung)

**SF-12**

Bei orthopädischen/rheumatologischen Patienten (n<sub>Min</sub>=438, n<sub>Max</sub>=646) ergibt sich für die „Körperliche Summenskala“ und die „Psychische Summenskala“ des SF-12 eine niedrige Änderungssensitivität zwischen Reha-Beginn und Reha-Ende (s. Tab. 3). Tendenziell ändersensitiver reagieren lediglich die Items 1.) „Allgemeine Gesundheit“, 4.) „weniger geschafft (körperliche Gesundheit)“ und 8.) „behindert wegen Schmerzen“. Für die übrigen Items resultieren Effektstärken im unteren Bereich, die sich jedoch überzufällig von Null unterscheiden.

Der SF-12 scheint sich bei kardiologischen Patienten während der Rehabilitation sowohl hinsichtlich der „Körperlichen Summenskala“ als auch der „Psychischen Summenskala“ ändersensitiver als bei orthopädischen/rheumatologischen Patienten zu verhalten, obwohl die Ergebnisse für die „Körperliche Summenskala“ keine statistische Bedeutsamkeit erlangen. Für eine Reihe von Items treten Effektstärken in einem mittleren Bereich auf, wozu Item 1.) (s.o.), Item 4.) (s.o.), Item 6.) „weniger geschafft (emotionale Gesundheit)“, Item 7.) „nicht so sorgfältig“, Item 8.) (s.o.) und Item 10.) „voller Energie sein“ zählen. Die übrigen Effektstärken befinden sich in einem niedrigen Bereich, ohne sich überzufällig von Null zu unterscheiden.

Im Bereich Psychosomatik (n<sub>Min</sub>=60, n<sub>Max</sub>=79) erfassen die SF-12-Summenskalen aufgetretene Veränderungen während des Rehabilitationsaufenthalts eher schwach sensitiv. Die „Psychische Summenskala“ zeigt hier wie beim SF-36 tendenziell bessere Ergebnisse als die „Körperliche Summenskala“. Die Items 1.) (s.o.) und 10.) (s.o.) erweisen sich mit mittelgroßen Effektstärken als deutlich ändersensitiver als die übrigen Items. Zwar erlangen die Items 4.) (s.o.), 9.) „ruhig und gelassen“, 11.) „entmutigt und traurig“ und 12.) „Kontakte beeinträchtigt (Häufigkeit)“ ebenfalls statistische Bedeutsamkeit, liegen aber vergleichsweise niedriger. Die Kennwerte der übrigen Items schwanken in einem unteren Bereich, ohne sich statistisch bedeutsam von Null zu unterscheiden. Hinsichtlich der Änderungssensitivität bei der 3-Monats-Katamnese (n<sub>Min</sub>=41, n<sub>Max</sub>=52) liegen bei psychosomatischen Patienten vom Profil her ähnliche Ergebnisse wie für den SF-36 vor. Jedoch erhöhen sich die SES- bzw. SRM-Kennwerte für den Zeitraum zwischen Reha-Beginn und der 3-Monats-Katamnese durchschnittlich im Vergleich zu den SES- bzw. SRM-Kennwerten bei Reha-Ende.

**Tabelle 2: Änderungssensitivität des SF-36 für verschiedene Indikationen im Veränderungszeitraum zwischen Reha-Beginn und Reha-Ende**

SF-12	ORTHOPÄDIE/RHEUMATOLOGIE (extrahiert) n=[438;646]						KARDIOLOGIE (extrahiert) n=[22;31]						PSYCHOSOMATIK n=[60;79]					
	SES			SRM			SES			SRM			SES			SRM		
	MW	UG	OG	MW	UG	OG	MW	UG	OG	MW	UG	OG	MW	UG	OG	MW	UG	OG
KSK	0.26*	0.16	0.35	0.33*	0.24	0.43	0.40	-0.04	0.83	0.43	-0.01	0.86	0.24	-0.01	0.50	0.31*	0.05	0.57
PSK	0.21*	0.12	0.31	0.28*	0.19	0.38	0.34	-0.10	0.77	0.47*	0.02	0.91	0.36*	0.10	0.62	0.39*	0.13	0.66
Item 1	0.34*	0.26	0.42	0.35*	0.27	0.43	0.35	-0.01	0.71	0.45*	0.08	0.82	0.59*	0.35	0.83	0.47*	0.23	0.70
Item 2	0.08	0.00	0.16	0.09*	0.01	0.17	-0.12	-0.48	0.25	-0.12	-0.48	0.25	0.16	-0.06	0.39	0.18	-0.05	0.40
Item 3	0.08*	0.01	0.16	0.10*	0.02	0.18	0.24	-0.12	0.60	0.36	-0.01	0.72	0.12	-0.10	0.35	0.13	-0.10	0.35
Item 4	0.38*	0.29	0.46	0.34*	0.26	0.43	0.72*	0.32	1.11	0.55*	0.17	0.92	0.31*	0.08	0.54	0.30*	0.07	0.53
Item 5	0.15*	0.07	0.24	0.15*	0.06	0.23	0.35	-0.02	0.71	0.31	-0.06	0.68	0.17	-0.06	0.39	0.17	-0.05	0.40
Item 6	0.13*	0.05	0.22	0.14*	0.06	0.22	0.43*	0.04	0.80	0.50*	0.11	0.88	0.06	-0.16	0.29	0.06	-0.16	0.29
Item 7	0.13*	0.04	0.21	0.14*	0.05	0.22	0.44*	0.04	0.83	0.52*	0.12	0.92	0.18	-0.05	0.40	0.14	-0.09	0.36
Item 8	0.39*	0.30	0.47	0.47*	0.38	0.56	0.42*	0.05	0.78	0.47*	0.10	0.84	0.19	-0.04	0.42	0.23	0.00	0.45
Item 9	0.24*	0.16	0.32	0.24*	0.16	0.32	0.24	-0.13	0.60	0.30	-0.06	0.67	0.31*	0.08	0.55	0.33*	0.09	0.56
Item 10	0.26*	0.18	0.34	0.30*	0.22	0.38	0.27	-0.09	0.62	0.46*	0.08	0.83	0.57*	0.32	0.82	0.46*	0.22	0.70
Item 11	0.24*	0.16	0.32	0.26*	0.18	0.33	0.14	-0.22	0.51	0.23	-0.14	0.59	0.35*	0.12	0.58	0.37*	0.14	0.60
Item 12	0.17*	0.09	0.25	0.20*	0.12	0.28	0.13	-0.23	0.48	0.15	-0.21	0.50	0.42*	0.19	0.65	0.34*	0.11	0.57

(MW=Mittelwert des Konfidenzintervalls [95%]; Werte, deren Konfidenzintervall Null nicht einschließt, sind mit einem Stern \* markiert; UG=Untergrenze; OG=Obergrenze; KSK=Körperliche Summenskala, PSK=Psychische Summenskala; KoFu=körperliche Funktionsfähigkeit, KoRo=körperliche Rollenfunktion, Schm=Schmerz, AGes=allgemeine Gesundheitswahrnehmung, Vita=Vitalität, SoFu=soziale Funktionsfähigkeit, EmRo=emotionale Rollenfunktion, Psych=psych. Wohlbefinden, VerGes=Gesundheitsveränderung)



## Diskussion und Fazit

Die Ergebnisse der Reanalysen von Daten aus zehn Forschungsprojekten, die aus mehreren rehabilitationswissenschaftlichen Forschungsverbänden stammen, liefern einige aufschlussreiche Aussagen zur Änderungssensitivität des SF-36 und des SF-12 im Kontext der Rehabilitation. In methodischer Hinsicht ist zu beobachten, dass beide Effektstärkevarianten, SES und SRM, ähnliche Ergebnisse liefern und daher gleichermaßen geeignet erscheinen zur Beurteilung der Änderungssensitivität unter den hier vorliegenden Rahmenbedingungen. In inhaltlicher Hinsicht erweisen sich der SF-36 bzw. SF-12 und ihre Skalen – abhängig von Indikation und Veränderungszeitraum – als unterschiedlich ändersensitiv. Im Bereich Orthopädie/Rheumatologie scheinen der SF-36 und der SF-12 zur Messung von Veränderungen weniger geeignet zu sein, wenngleich einzelne Skalen und Items durchaus in einem Bereich mittlerer bis hoher Änderungssensitivität liegen. Bei kardiologischen Patienten verhalten sich beide Instrumente mittelgradig ändersensitiv. In der Psychosomatik scheinen der SF-36/SF-12 ähnlich schwach ändersensitiv auf körperliche Veränderungen zu reagieren, wobei der SF-36 im Unterschied zum SF-12 hinsichtlich psychischer Veränderungen eine deutlich höhere, mittlere Änderungssensitivität aufweist.

Wie bereits betont wurde, haben die verwendeten Kennwerte zur Änderungssensitivität nur dann Aussagekraft, wenn das Studiendesign eine nachweisbar wirksame Intervention enthält. Selbst wenn diese Voraussetzung erfüllt ist, kann bei den hier ausgewerteten Eingruppen-Prä-Post-Designs Änderungssensitivität und Wirksamkeit der Therapie nicht wirklich getrennt werden. Das vorliegende Datenmaterial erlaubt also streng genommen keine Aussage darüber, ob z.B. die Rehabilitationsmaßnahmen bei kardiologischen Patienten effektiver oder ob der SF-36/SF-12 bei diesen Kollektiven ändersensitiver sind. Um zu einer aussagekräftigen Beurteilung der Änderungssensitivität dieser Fragebogen zu gelangen, erscheint auch ein kritischer Vergleich mit anderen, ähnlichen Instrumenten erforderlich, auf welchen an dieser Stelle nicht eingegangen werden konnte. Dieser Ansatz wird zur Zeit im Forschungsprojekt A7 des Rehabilitationswissenschaftlichen Forschungsverbundes Bayern (RFB) in Form einer prospektiven Studie an orthopädischen/rheumatologischen und kardiologischen Rehabilitanden verfolgt, in

welcher der SF-36-Fragebogen, der IRES-3-Fragebogen [4] und einige Skalen der SCL-90-R [3] parallel eingesetzt werden.

Zusammenfassend kann die Schlussfolgerung gezogen werden, dass die vorliegenden Ergebnisse erste Anhaltspunkte zur Beurteilung der Änderungssensitivität der Skalen des SF-36 und des SF-12 liefern. Diese können bei der Planung von evaluativen Studien in der Rehabilitation zur Auswahl des geeigneten Instruments oder geeigneter Skalen verwendet werden, aber auch bei der Festlegung von Outcome-Kriterien für eine erfolgreiche Rehabilitation nützlich sein. Schließlich liefern sie auch Ansatzpunkte für Verbesserungen der im Augenblick vorliegenden Versionen der untersuchten Assessmentinstrumente.

## Danksagung

Dieser Beitrag basiert auf Ergebnissen des Forschungsprojekts A7 des Rehabilitationswissenschaftlichen Forschungsverbundes Bayern (RFB), welches im Rahmen der zweiten Förderphase des gemeinsamen Förderschwerpunkts „Rehabilitationswissenschaften“ vom Bundesministerium für Bildung und Forschung (BMBF) und der Deutschen Rentenversicherung finanziert wird. Die Durchführung der Reanalysen wäre nicht möglich gewesen ohne die Zusammenarbeit mit zu Beginn genannten Personen, die freundlicherweise Daten aus ihren Projekten zur Verfügung stellten. Gedankt sei des Weiteren allen Projekt- und Klinikmitarbeitern sowie allen Patienten in den Rehabilitationseinrichtungen, die zum Gelingen jedes einzelnen Projekts beigetragen haben.

## Literatur

- 1 Bullinger, M. & Kirchberger, I. (1998). SF-36 Fragebogen zum Gesundheitszustand (SF-36). Handbuch für die deutschsprachige Fragebogenversion. Göttingen: Hogrefe.
- 2 Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- 3 Franke, G. H. (1995). SCL-90-R: Die Symptom-Checkliste von Derogatis – Deutsche Version: Manual. Göttingen: Beltz.
- 4 Gerdes, N., Bühlren, B. & Jäckel, W. H. (2003). Indikatoren des Reha-Status – IRES-Fragebogen (Version 3). <http://www.hri.de> (01/11/03).
- 5 Hevey, D. & McGee H. M. (1998). The effect size statistic: useful in health outcomes research? *Journal of Health Psychology*, 3, 163-170.



- 6 Igl, W., Schuck, P., Zwingmann, Ch. & Faller, H. (2003). Änderungssensitivität von generischen Patientenfragebogen in der Rehabilitation – Erste Ergebnisse einer verbundübergreifenden Reanalyse. In Verband Deutscher Rentenversicherungsträger (Hrsg.), 12. Rehabilitationswissenschaftliches Kolloquium – Rehabilitation im Gesundheitswesen – vom 10. bis 12. März 2003 in Bad Kreuznach (DRV-Schriften, Bd. 40, S. 80-81). Frankfurt am Main: VDR.
- 7 Igl, W., Zwingmann, Ch., Faller, H. & Schuck, P. (in Vorb.). Änderungssensitivität. Serie Methoden. Rehabilitation.
- 8 Kazis, L. E., Anderson, J. J. & Meenan, R. F. (1989). Effect sizes for interpreting changes in health status. *Medical Care*, 27, 178-189.
- 9 Kirshner, B. & Guyatt, G. (1985). A methodological framework for assessing health indices. *Journal of Chronic Disease*, 38, 27-36.
- 10 Schuck, P. (2000). Designs und Kennziffern zur Ermittlung der Änderungssensitivität von Fragebogen in der gesundheitsbezogenen Lebensqualitätsforschung. *Zeitschrift für Medizinische Psychologie*, 9, 125-130.
- 11 Schuck, P. & Zwingmann, Ch. (2003). The 'smallest real difference' as a measure of sensitivity to change: a critical analysis. *International Journal of Rehabilitation Research*, 26, 85-91.
- 12 Stratford, P. W., Binkley, J. M. & Riddle, D. L. (1996). Health status measures: strategies and analytic methods for assessing change scores. *Physical Therapy*, 76, 1109-1123.
- 13 Juniper, E. F., Guyatt, G. H. & Jaeschke, R. (1996). How to develop and validate a new quality of life instrument. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (pp. 49-56). Philadelphia, PA: Lippincott-Raven.
- 14 Zwingmann, Ch. (2002). Der IRES-Patientenfragebogen. Psychometrische Reanalysen an einem rehabilitationsspezifischen Assessmentinstrument. Regensburg: Roderer.

## Kapitel 6

### SF-12 und SF-36 im Rahmen von Vorhersagen des Erfolges von Patientenschulungen – Eine Reanalyse gepoolter Daten

MARC WORBACH, ANDREA REUSCH, HEINER VOGEL &  
HERMANN FALLER

#### Einleitung

Dieser Beitrag konzentriert sich auf Fragen, die die Verwendung des SF-12 und SF-36 im Rahmen von Erfolgsvorhersagen von Patientenschulungen in der stationären medizinischen Rehabilitation betreffen. Die Strategie ist, (1) die Korrelationen des SF-12 und SF-36 mit psychologischen Variablen zum Rehabilitationsbeginn  $t_1$  zu untersuchen, (2) aus SF-12 und SF-36 gebildete Erfolgskriterien untereinander und mit parallel erhobenen Veränderungs- und Statuswerten zum Katamnesezeitpunkt  $t_3$  zu vergleichen und (3) den Einsatz des SF-12 und SF-36 in einer multiplen Regression zur Vorhersage des Patientenschulungserfolges zu analysieren.

#### Das Projekt Patientenschulung und die Vorhersage des Patientenschulungserfolges

SF-12 und SF-36 im Rahmen prädiktiver Fragen zu diskutieren, ist eng verwoben mit den Zielen des Projektes „Verbundübergreifende Reanalysen zu Prädiktoren des Patientenschulungserfolges und Entwicklung von allgemeinen Empfehlungen zur Evaluation von Patientenschulungen“ des Rehabilitationswissenschaftlichen Forschungsverbundes Bayern. Ein Projektziel ist es, allgemeine Prädiktoren des Patientenschulungserfolges zu untersuchen. Hierfür wurden Daten aus 11 Patientenschulungsprojekten unabhängig von Indi-

Worbach, M., Reusch, A., Vogel, H. & Faller, H. (2004). SF-12 und SF-36 im Rahmen von Vorhersagen des Erfolges von Patientenschulungen – Eine Reanalyse gepoolter Daten. In C. Maurischat, M. Morfeld, T. Kohlmann & M. Bullinger (Hrsg.), *Lebensqualität: Nützlichkeit und Psychometrie des Health Survey SF-36/SF-12 in der medizinischen Rehabilitation* (S. 91-102). Lengerich: Pabst.