

## )) Patientenzentrierte Evaluation des Gesundheitszustands in einem longitudinalen Qualitätsmanagementsystem im Krankenhaus (QMK)

S. Schneeweiss<sup>1</sup>, O. Sangha<sup>2</sup>, A. Manstetten<sup>2</sup> und die QMK Entwicklungsgruppe\*

<sup>1</sup> Harvard Medical School, Boston, USA

<sup>2</sup> Ludwig-Maximilians-Universität, München

**Zusammenfassung:** Zielsetzung: Ziel der vorliegenden Studie war, den Einsatz von patientenzentrierten Fragebogen zum Gesundheitszustand als Indikator der Ergebnisqualität in der stationären Versorgung zu prüfen, mögliche Schwachstellen zu identifizieren und Verbesserungsvorschläge zu erarbeiten.

**Methodik:** In 4 Krankenhäusern der Primärversorgung nahmen zwischen 1. August 1999 und 31. Oktober 1999 180 stationär aufgenommene internistische Patienten an einer umfassenden, longitudinalen Bewertung der Ergebnisqualität teil. SF-36-Fragebogen zum Gesundheitszustand wurden von Patienten sowohl bei Aufnahme als auch 4 Wochen nach Entlassung ausgefüllt. Pflegekräfte und Ärzte gaben ebenfalls Bewertungen zum Gesundheitszustand der Patienten ab. Lineare Zusammenhänge zwischen den Bewertungen des Gesundheitszustands wurden mit Korrelationskoeffizienten bestimmt. Odds Ratios (OR) und 95% Konfidenzintervalle aus multivariaten logistischen Regressionsmodellen quantifizierten die Assoziation zwischen ausgewählten Prädiktoren und dem Rücklauf von Bogen.

**Ergebnisse:** 87 von 115 SF-36-Bogen (75%), die bei Aufnahme ausgehändigt wurden, und 38 von 115 (33%), die bei Entlassung ausgehändigt wurden, wurden ausgefüllt zurückgesandt. Patienten mit Funktionsstörungen und Patienten, die verlegt wurden, antworteten signifikant seltener (OR = 1,3 [1,09; 1,66] bzw. OR = 3,7 [1,37; 9,87]). Der Zusammenhang zwischen SF-36 und SF-12 war  $r = 0,95$  [0,91; 0,97] für körperlich und  $r = 0,91$  [0,85; 0,94] für seelische Gesundheit. Ärzte und Pflegekräfte stimmten in ihrer Bewertung des Patienten mäßig überein ( $r = 0,38$  [0,22; 0,52]), beide waren jedoch weit von der Einschätzung der Patienten entfernt ( $r = 0,15$  [-0,08; 0,36] bzw.  $r = -0,01$  [-0,23; 0,21]).

**Schlussfolgerungen:** 1. Eine patientenzentrierte Bewertung des Gesundheitszustands sollte in die Beurteilung der Ergebnisqualität der stationären Versorgung eingehen, 2. kurze Instrumente wie der SF-12 können anstatt des SF-36 zum Einsatz kommen, 3. eine Patientenbefragung ca. 4 Wochen nach Entlassung ist praktikabel und 4. auf Patienten mit multiplen Funktionsstörungen und Patienten, die nicht nach Hause entlassen wurden, muss besonders eingegangen werden, um den Rücklauf zu erhöhen oder zumindest minimale Aussagen über diese Patientengruppe zu ermöglichen.

**Self-reported health status in a longitudinal hospital-based quality management system:** Objective: The objective was to assess the value of self-reported health status as an indicator of outcomes of acute hospital care, to identify potential practical limitations, and develop strategies for future applications.

**Methods:** 180 patients from 4 acute care hospitals in Germany participated in a comprehensive and longitudinal assessment of outcomes of care in general internal medicine between 8/1/99 and 10/31/99. Self-administered SF-36 surveys were completed at admission and 4 weeks after discharge. Additionally, nurses and physicians answered questions regarding the patients' health status. Linear relations between health status assessments were quantified as correlation coefficients. Odds ratios (OR) and 95% confidence limits from multivariate logistic regression models were reported for predictors of non-returned questionnaires.

**Results:** 33% of SF-36 surveys handed out at discharge were returned. Patients with impairments and referred patients were more likely not to return the survey (OR = 1,3 [1,09; 1,66] and OR = 3,7 [1,37; 9,87]). The linear relation of SF-36 and SF-12 scores in the same patients were  $r = 0,95$  [0,91; 0,97] for physical health and  $r = 0,91$  [0,85; 0,94] for mental health. Physicians and nurses moderately agreed in their assessment of patients' health ( $r = 0,38$  [0,22; 0,52]) but both professional groups showed poor agreement with self-reported health ( $r = 0,15$  [-0,08; 0,36] and  $r = -0,01$  [-0,23; 0,21]).

**Conclusions:** 1. Self-reported health status should be considered in the assessment of outcomes of acute care as a dimension that is to some extent independent of health status assessment by professionals, 2. shorter instruments, i.e., the SF-12, can be used instead of the SF-36, 3. a self-reported health status assessment is feasible 4 weeks after discharge, and 4. patients with multiple impairments or those who are transferred should get specific support in the completion of questionnaires to increase response or to receive at least minimal information about their health status.

**Key words:** Self-reported Health Status – SF-36 – Outcomes Research – Quality Management – Health Services Research

## Einleitung

„Gesundheit ist ein Zustand von vollständigem körperlichen, mentalen und sozialen Wohlbefinden und nicht lediglich die Abwesenheit von Krankheit oder Gebrechen.“ Dies war die Definition von Gesundheit der WHO im Jahre 1947 [1] basierend auf Siegerests umfassendem Gesundheitskonzept. [2] Erst in den späten 80er und frühen 90er Jahren, mit der Entwicklung kurzgefasster und praktikabler Instrumente zur Erfassung der gesundheitsbezogenen Lebensqualität [3] wurden mentale und soziale Konsequenzen von Krankheit im Bereich des klinischen Qualitätsmanagement aufgegriffen.

In den USA sind heute Elemente des SF-36, einem standardisierten Fragebogen zur Erfassung der gesundheitsbezogenen Lebensqualität, in die Qualitätsdokumentation des National Council for Quality Assurance (NCQA) in Form von Querschnitterhebungen fest integriert [4]. Die NCQA-Akkreditierung ist unter Managed Care Organisationen in den USA weit verbreitet, da deren Berichte zur Versorgungsqualität Versicherern und Arbeitgebern einen direkten Vergleich unterschiedlicher Leistungsanbieter erlauben. Erste prospektive Längsschnittdaten zur Bewertung der Ergebnisqualität anhand des patientenbezogenen Gesundheitsstatus werden durch das in der Entwicklung befindliche ORYX-System der Joint Commission on Accreditation of Healthcare Organizations [5,6] und das Medicare Health Outcomes Project, ein Gemeinschaftsprojekt der Health Care Financing Administration (HCFA), der Anbieterorganisation Humana und QualityMetric, angestrebt [7].

In Deutschland wurde der Einsatz von Instrumenten zum patientenbezogenen Gesundheitszustand in der longitudinalen und prospektiven Bewertung von Ergebnisqualität bisher nur selten und meist für spezifische Indikationen untersucht [8–10]. Eine lediglich querschnittliche und retrospektive Bewertung des Gesundheitszustands nach stationären Aufnahmen [11] ist nicht ausreichend, da der Zustand der Patienten bei Aufnahme nicht charakterisiert werden kann und damit eine Quantifizierung der Veränderung des Gesundheitszustands und eine sinnvolle Risiko-Adjustierung nicht möglich sind [12].

Ziel der vorliegenden Studie war, den Einsatz von patientenzentrierten Fragebogen zum Gesundheitszustand als Indikator der Ergebnisqualität in einem longitudinalen und prospektiven Qualitätsmanagementsystem der stationären Versorgung im Rahmen des Projekts Qualitätsmodell Krankenhaus (QMK) zu prüfen, mögliche Schwachstellen zu identifizieren und Verbesserungsvorschläge zu erarbeiten.

## Methoden

### Das QMK-Instrument

Die Entwicklung des QMK-Instruments zur Bewertung von Ergebnisqualität wurde an anderer Stelle ausführlich beschrieben [13,14]. Generelles Prinzip des Instruments ist es, Ergebnisqualität anhand ausgewählter Indikatoren zu bewerten. Die Indikatoren werden bei Aufnahme und bei Entlassung bei jedem Patienten erfasst, wobei die Ergebnisqualität durch die Differenz zwischen Entlassungs- und Aufnahmezeitpunkt dargestellt wird. Eine patientenzentrierte Bewertung

des Gesundheitszustands wurde bei Aufnahme und ca. 4 Wochen nach Entlassung durchgeführt.

### Design der Pilotstudie

An der Pilotstudie nahmen vier Krankenhäuser der Primärversorgung teil. Die Auswahl erfolgte nach Kriterien der Durchführbarkeit und dem kurzfristigen Vorhandensein eines Projektverantwortlichen zur Betreuung der Studie. Daneben sollte in den beteiligten Abteilungen während der Pilotphase kein anderes Qualitätsprojekt (z.B. Zertifizierungsverfahren) durchgeführt werden.

Alle konsekutiven internistischen Patienten, die zwischen dem 1. August 1999 und 31. Oktober 1999 stationär aufgenommen wurden, wurden auf das Vorhandensein von definierten Tracer-Diagnosen aus den Bereichen (Module) Herz, Lunge, Bewegungsapparat, Gehirn oder Diabetes untersucht und diesen entsprechend zugeordnet. Ein Allgemein-Modul erfasst anthropometrische Variablen und den Funktions- und Pflegestatus der Patienten. Das Allgemein-Modul wurde durch Pflegekräfte ausgefüllt. Ärzte erfassten klinische Parameter und Begleiterkrankungen mit deren Schweregrad. Zusätzlich wurde jedem Patienten, der in die Studie aufgenommen wurde, bei Aufnahme ein Patientenfragebogen, der neben anderen Fragen den SF-36-Fragebogen zum Gesundheitszustand enthält, ausgeteilt. Bei Entlassung erhielt jeder Patient einen weiteren Patientenfragebogen mit dem SF-36, der 4 Wochen nach Entlassung ausgefüllt und an die Universität München geschickt werden sollte. Drei Wochen nach Entlassung erhielten alle Patienten ein Erinnerungsschreiben adressiert an ihren Wohnsitz. Für die Aufnahme in die Studie bestand keine Altersbeschränkung. Patienten wurden von der Patientenbefragung ausgenommen, falls deren kognitiver Zustand hinsichtlich der selbständigen Beantwortung des Fragebogens als fraglich bzw. nicht möglich eingestuft wurde.

### Instrumente und Skalen

Der Short-Form-36 (SF-36) ist ein standardisiertes Instrument zur krankheitsübergreifenden patientenzentrierten Erfassung des Gesundheitszustands, der aus den Daten der Medical Outcome Study entwickelt und als Messinstrument der medizinischen Versorgungsqualität propagiert wurde [15,16]. Die Übersetzung ins Deutsche, die psychometrische Testung und Normierung dieses Instruments erfolgte durch die im International Quality of Life Assessment Project (IQOLA) zusammengesetzte Arbeitsgruppe [17]. Die deutsche Version [18] besteht aus 36 Items mit acht Subskalen: körperliche und soziale Funktionsfähigkeit, Rolleneinschränkung aufgrund körperlicher Probleme, Schmerz, seelisches Wohlbefinden, Rolleneinschränkung aufgrund emotionaler Probleme, Vitalität, allgemeine Gesundheitseinschätzung. Zusätzlich wurde eine Bewertung von nur 12 der 36 Fragen entsprechend des in der Literatur beschriebenen Auswertungsalgorithmus des SF-12 vorgenommen [19].

Um die Einschätzung des Gesundheitszustands durch unterschiedliche Personengruppen (Patienten, Pflegekräfte und Ärzte) bewerten zu können, wurden sowohl Pflegekräfte als auch Ärzte nach einer Bewertung des allgemeinen Gesundheitszustands der Patienten gefragt. Die Frage ist dem SF-36

Module	n	Beteiligte Kliniken			
		(1)	(2)	(3)	(4)
Allgemein-Modul: Rücklauf	180	58	57	28	37
Organ-Module: Rücklauf	148	48	55	27	18
Patienten-Module: Rücklauf					
Pat.befragung SF 36 bei Aufnahme	87	39 (67%)*	48 (84%)*	-	-
Pat.befragung SF 36 nach Entlassung	38	17 (29%)*	21 (37%)*	-	-

Tab. 1 Übersicht über den Rücklauf der einzelnen Module getrennt für die beteiligten 4 Kliniken

\* Prozentual bezogen auf Gesamtzahl der rückläufigen Allgemein-Module pro Klinik

entnommen (GH1) und lautet „Wie beurteilen Sie den Gesundheitszustand des Patienten im Allgemeinen?“ Die Antwortmöglichkeiten sind identisch zu denen des SF-36 und erlauben damit einen direkten Vergleich zu den Patientenangaben. Untersuchungen an einer repräsentativen deutschen Bevölkerungsstichprobe zeigten, dass Antwortverteilungen zu diesen Fragen alter- und morbiditätsunabhängig sind und weder Boden- noch Deckeneffekte aufweisen [20].

Funktionseinschränkungen und Begleiterkrankungen wurden anhand einer deutschen Übersetzung von Greenfields ICED-Score bewertet [21]. Der Score erfasst 11 Funktionsstörungen mit je 2 Schweregraden und 16 (Begleit-)Erkrankungen mit jeweils 4 Schweregraden. Der ICED wurde unabhängig von 2 Autoren (SS, AM) übersetzt und Unstimmigkeiten wurden gemeinsam mit dem dritten Autor geklärt. Eine validierte deutsche Version liegt derzeit nicht vor.

In zwei Kliniken wurden versuchsweise COOP-WONCA-Charts [22] zur grafisch unterstützten Erfassung des Gesundheitszustands eingesetzt. Der Rücklauf war jedoch so gering, um sinnvolle Aussagen treffen zu können.

Statistische Verfahren

Pearson-Korrelationskoeffizienten wurden berechnet, um lineare Zusammenhänge darzustellen [23]. Für multikategoriale Werte wurden Spearman-Korrelationskoeffizienten berechnet. Die entsprechenden 95%-Konfidenzintervalle wurden berechnet [24]. Zur Identifizierung von Prädiktoren für den erfolgreichen Rücklauf von Patientenbogen wurden multivariate logistische Regressionsmodelle den Daten angepasst. 95% Konfidenzintervalle vom Wald-Typ wurden berechnet [25]. Mittelwerte normalverteilter metrischer Variablen wurden mittels Varianzanalyse auf ihre Gleichheit zwischen mehreren Gruppen getestet. Der Scheffe-Test wurde verwendet, um den  $\alpha$ -Fehler für multiples Testen zu adjustieren [26]. Zweiseitige p-Werte werden berichtet. Alle Berechnungen wurden mit der SAS Software durchgeführt [27].

Ergebnisse

Rücklauf

Die 180 Studienteilnehmer waren durchschnittlich 68 Jahre alt ( $\pm 14,5$ ). 53% der Patienten waren weiblich, 9,7% wurden mit Bewusstseinsstörungen eingeliefert und 40% hatten mindestens eine schwere Funktionseinschränkung bei Einweisung. Die mittlere Verweildauer lag bei 12,1 Tagen ( $\pm 7,1$ ). Der

Rücklauf der einzelnen Fragebogen ist in Tab. 1 getrennt nach den vier beteiligten Test-Kliniken dargestellt. Insgesamt wurden 125 Fragebogen zum Gesundheitszustand (SF-36) ausgefüllt zurückgegeben. Dies entspricht einem Rücklauf von 75% bei Aufnahme und 33% vier Wochen nach Entlassung.

Patienten, die nicht nach Hause entlassen wurden, sondern in andere Krankenhäuser, Rehabilitationseinrichtungen oder Pflegeheime verlegt wurden, hatten eine 3,7fach höhere Wahrscheinlichkeit (95%-Konfidenzintervall: 1,37–9,87), den patientenzentrierten Fragebogen nicht zurückzusenden, nachdem für die Einflüsse von Alter, Anzahl und Schweregrad der Funktionsstörungen und Begleiterkrankungen adjustiert wurde. Mit einer Zunahme des Funktionsstörungs-Scores um eine Einheit nahm das Risiko, den Bogen nicht zurückzusenden, um das 1,3fache zu (1,09–1,66).

Entsprechend den Anleitungen zum Ausfüllen der Fragebogen zum Gesundheitszustand nach Entlassung wurden diese im Mittel nach vier Wochen ausgefüllt und an die Universität München gesandt (Abb. 1).

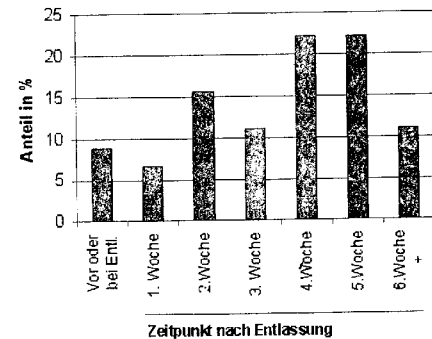


Abb. 1 Zeitliche Verteilung des Ausfüllungsdatums des zweiten Patientenbogens in Bezug zum Entlassungszeitpunkt. Vorgabe war, den Bogen vier Wochen nach Entlassung auszufüllen.

Skala	Anzahl Fragen pro Teil-Skala (SF36)*	mittlerer Anteil fehlender Werte
1: körperliche Funktionsfähigkeit (PF)	10	4,0%
2: körperliche Rolleneinschränkung (RP)	4	9,6%
3: Schmerz (BP)	2	18,4%
4: allgemeiner Gesundheitseinschätzung (GH)	5	12,8%
5: Vitalität (VT)	4	22,4%
6: soziale Funktionsfähigkeit (SF)	2	6,4%
7: emotionale Rolleneinschränkung (RE)	3	10,4%
8: seelisches Wohlbefinden (MH)	5	12,0%

\* 35 Fragen plus eine Frage zur Veränderung des Gesundheitszustands

**Tab. 2** Fehlende Angaben in 125 zurückgegangenen SF-36-Instrumenten bei Einsatz in der stationären internistischen Versorgung. Alle Fragebogen (Aufnahme und vier Wochen nach Entlassung) wurden berücksichtigt.

**Tab. 3** Lineare Korrelation der Skalen zur körperlichen und seelischen Gesundheit zwischen SF-36 und SF-12 innerhalb 125 zurückgegangener SF-36-Instrumente bei Einsatz in der stationären Versorgung in der inneren Medizin (Aufnahme und nach Entlassung).

Skala und Zeitpunkt	n	Pearson-Korrelationskoeffizient (SF-36 vs. SF-12)	95%-Konfidenz-Intervall	R <sup>2</sup>
körperliche Gesundheit bei Aufnahme	42	0,949	0,907–0,972	0,90
körperliche Gesundheit nach Entlassung	15	0,946	0,842–0,982	0,89
körperliche Gesundheit gesamt	57	0,946	0,910–0,968	0,89
seelische Gesundheit bei Aufnahme	42	0,915	0,846–0,954	0,84
seelische Gesundheit nach Entlassung	15	0,884	0,679–0,961	0,78
seelische Gesundheit gesamt	57	0,907	0,846–0,944	0,82

#### Fehlende Angaben innerhalb des SF-36

Aus insgesamt 125 zurückgesandten Bogen konnten 62 Gesamtskalen für die körperliche und seelische Gesundheit berechnet werden (49,6%). In den übrigen Fällen war dies nicht möglich, da jeweils zu mehr als einer Frage in einer der 8 Subskalen keine Angaben gemacht wurden. Tab. 2 stellt im Einzelnen dar, welche Subskalen welchen Anteil an fehlenden Angaben aufwies. Fragen zur Vitalität blieben mit 22,4% aller zurückgegangenen Bogen am häufigsten unbeantwortet.

Einzelne Fragen mit besonders hohem Anteil fehlender Angaben waren: „Wie stark waren Ihre Schmerzen in den vergangenen 4 Wochen?“ (BP1; 26,4%); „Wie oft waren Sie in den vergangenen 4 Wochen voller Schwung?“ und „... voller Energie?“ (VT1 und VT2; 30,4% und 29,6%); „Ich bin genauso gesund wie alle anderen, die ich kenne“ (GH3; 27,2%); „Ich erfreue mich ausgezeichneter Gesundheit“ (GH5; 28%); „Wie oft waren Sie in den vergangenen 4 Wochen ruhig und gelassen?“ und „... glücklich?“ (MH3 und MH5; 27,2% und 31,2%).

Im SF-12 werden die Fragen BP1, MH5, VT1, GH3 sowie GH5 nicht gefragt. Entsprechend konnten in 76 von 125 Fällen (60,8%) Gesamtskalen für die körperliche und seelische Gesundheit berechnet werden, 23% mehr als für den SF-36 für dieselben Patienten.

#### Zusammenhang zwischen SF-36 und SF-12

Die Gesamtskalen für die körperliche und seelische Gesundheit zeigten einen ausgeprägten linearen Zusammenhang zwischen SF-36 und SF-12 mit Korrelationskoeffizienten von

\* Bewertung auf einer 5-stufigen Likert-Skala von schlecht bis ausgezeichnet.

$r = 0,946$  und  $r = 0,907$  (Tab. 3). Damit lassen sich jeweils 89% bzw. 82% der Varianz um die Regressionsgerade erklären.

Die Unterschiede in individuellen Patienten zwischen den von 0 bis 100 reichenden standardisierten Skalen des SF-36 und SF-12 für körperliche und seelische Gesundheit waren im Mittel 0,62 und 0,83 und damit nicht signifikant unterschiedlich von 0 ( $p = 0,23$  und  $p = 0,13$ ). Diese hohen Übereinstimmungen zwischen SF-36 und SF-12 waren unabhängig von Hauptdiagnose, Alter, Geschlecht, Bewusstsein bei Aufnahme, Funktionseinschränkungen, Begleiterkrankungen oder Pflegestufen nach Korrektur des  $\alpha$ -Fehlers (alle  $p > 0,10$ ).

#### Übereinstimmung in der Bewertung des Gesundheitszustands

Abb. 2 stellt den linearen Zusammenhang in der Bewertung des Gesundheitszustands durch Patienten, Pflegekräfte und Ärzte dar. Die Übereinstimmung in der Bewertung des Gesundheitszustands durch Pflegekräfte und Ärzte erschien mäßig ( $r = 0,38$ ). Kein statistisch signifikanter Zusammenhang bestand hingegen zwischen Patienten und deren ärztlichen ( $r = 0,15$ ) und pflegerischen Betreuern ( $r = -0,01$ ). Tendenziell scheinen in der Stichprobe Ärzte der Patienteneinschätzung ihrer Gesundheit näher zu kommen als Pflegekräfte.

Entsprechend ist auch die Übereinstimmungen in der Beurteilung der Verbesserung des Gesundheitszustands im Rahmen des stationären Aufenthalts schlecht (Abb. 3). Innerhalb von Patienten und Ärzten ist die Reproduzierbarkeit der Veränderung durch zwei unterschiedliche Messverfahren, der numerischen Differenz aus der Bestimmungen bei Aufnahme und Entlassung (= prospektive Messung) bzw. eine direkte Frage nach der Veränderung der Gesundheit (= retrospektive Messung) mäßig hoch ( $r = 0,49$  und  $r = 0,57$ ). Der Ausgangswert des Gesundheitszustands war in der ärztlichen Bewertung negativ mit der Veränderung korreliert ( $r = -0,64$ ;

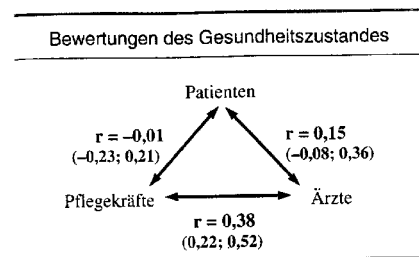


Abb. 2 Spearman-Korrelationen zwischen Bewertungen des Gesundheitszustandes\* durch Patienten, Pflegepersonal und Ärzte.

[-0,53; -0,73]), d.h. war der Gesundheitszustand bei Aufnahme schlechter, so wurden größere Verbesserungen berichtet.

**Diskussion**

Vielfach ist bereits die Forderung erhoben worden, dass Patienten bei der Beurteilung des medizinischen Behandlungserfolgs mit eingebunden werden müssen [28,29]. Der longitudinale Einsatz von patientenzentrierten Instrumenten in der Bewertung der Behandlungsqualität im stationären Bereich mit Messungen bei Aufnahme und nach Entlassung wurde bisher im deutschen Sprachraum nur selten demonstriert [9,10,30]. Wir haben den praktischen Einsatz eines patientenzentrierten Instruments zur Bewertung des Gesundheitszustandes, dem SF-36, in deutschen Kliniken erprobt und konnten zeigen, dass die Bewertung des Behandlungserfolgs durch Patienten erheblich von der ärztlichen Einschätzung abweicht. Daraus resultiert die Notwendigkeit einer patientenzentrierten Evaluation von Gesundheitsparametern neben einer Einschätzung durch Ärzte und Pflegekräfte durchzuführen. Bei dem Einsatz des Instruments in der klinischen Routineversorgung sollten verschiedene Aspekte berücksichtigt werden:

Der Rücklauf aller ausgehändigten SF-36-Bogen nach Entlassung der Patienten mit 33% möglicherweise nicht ausreichend für repräsentative Aussagen. Wichtigste Prädik-

toren für nicht zurückgegangene Patientenbogen waren erwartungsgemäß das Vorhandensein von schweren körperlichen und seelischen Funktionsstörungen oder falls Patienten nicht nach Hause entlassen wurden.

Mit lediglich einem schriftlichen Erinnerungsschreiben an die Patienten wurde erreicht, dass der Zeitpunkt des Ausfüllens – mit den unvermeidlichen Schwankungen – mäßig gut eingehalten wurde. Nachdem der unvollständige Rücklauf hauptsächlich auf andere Faktoren wie Funktionseinschränkungen zurückzuführen ist, scheint die Wahl des Zeitpunkts geeignet sowie lediglich ein Erinnerungsschreiben ausreichend.

Nachdem aber gerade die Bewertung des Gesundheitszustands nach Entlassung aus der stationären Behandlung für die Bewertung des Behandlungserfolgs wichtig ist, müssen Alternativen zum SF-36 und dessen Handhabung in Erwägung gezogen werden. Aus diesem Grund explorierten wir die Möglichkeit, die abgekürzte Version des SF-36, den SF-12, einzusetzen. Unsere Studie zeigte, dass der sehr starke Zusammenhang zwischen SF-36 und SF-12 mit den Ergebnissen der Entwickler des Instruments vergleichbar ist. Für die Skala zur körperlichen Gesundheit gaben Ware et al. [19,31] einen Korrelationskoeffizienten von 0,951 an (unsere Studie: 0,910) und für die seelische Gesundheit von 0,969 (unsere Studie: 0,907). Die hohe Übereinstimmung war unabhängig von Faktoren, die den Krankheitsverlauf im Krankenhaus beeinflussen können, wie Alter, Funktionseinschränkungen, Hauptdiagnose oder Begleiterkrankungen. Durch den Einsatz des SF-12 werden Fragen umgangen, die besonders häufig nicht beantwortet wurden, so dass bei dem Einsatz des SF-12 die Anzahl der interpretierbaren Gesamtskalen für die körperliche und seelische Gesundheit um 23% gesteigert werden konnte. Es besteht die Möglichkeit, dass auch der Rücklauf durch ein kürzeres Instrument gesteigert werden kann. Die Verkürzung des SF-36 auf den SF-12 ist damit aus metrischer Sicht möglich, auch wenn die mit einer leichten Reduktion der Reliabilität einhergehen kann [32]. Aus Gründen der Praktikabilität und Effizienz ist damit der Einsatz des SF-12 zu empfehlen.

Nachdem Funktionsstörungen und Verlegungen wichtigste Prädiktoren für das Nichtausfüllen waren, sollte den Bogen ein Deckblatt angeheftet werden, das durch Angehörige oder betreuendes Personal ausgefüllt werden kann, falls Patienten

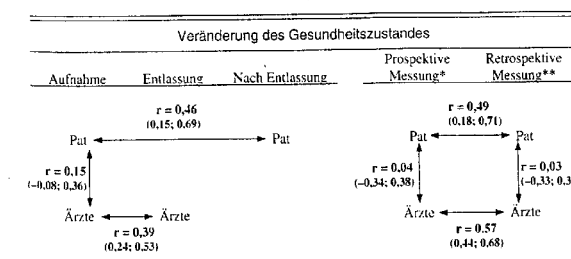


Abb. 3 Spearman-Korrelationen zwischen Bewertungen der Veränderung des Gesundheitszustandes durch Patienten (Pat) und Ärzte. (\* Prospektive Messung der Veränderung; Es handelt sich um die numerischen Differenzen von den 5-stufigen Aufnahme- und Entlassungswerten (Ärzte) bzw. vier Wochen danach (Patienten) innerhalb eines Patienten. \*\* Retrospektive Messung der Veränderung; Gefragt wurde nach der Veränderung des Gesamtzustandes bei Entlassung (Ärzte) bzw. vier Wochen danach (Patienten). Die Veränderung wurde auf einer 5-stufigen Skala von „deutlich verschlechtert“ bis „deutlich gebessert“ bewertet. „Unverändert“ wurde als 0 kodiert, Verbesserungen positiv, Verschlechterungen negativ.)

dazu nicht mehr in der Lage sind. Dieses Deckblatt sollte mit wenigen Fragen nach den Gründen forschen, warum Patienten den Bogen nicht ausfüllen können. Dies würde Schlüsse auf die Repräsentativität zukünftiger Ergebnisse ermöglichen. Ferner sollte mittels ein oder zwei Fragen der Gesundheitszustand beschrieben werden.

Im Rahmen der QMK-Studie wurde der Gesundheitszustand in den vergangenen vier Wochen erfragt. Einzelnen Rückfragen von Patienten zufolge gab es Unklarheiten, ob dieser Zeitraum das Ereignis, das letztlich zur Krankenhauseinweisung geführt hatte, mit einschließt. Um Unklarheiten zu vermeiden, ist daher zu überlegen, das Zeitfenster von vier Wochen auf sieben Tage einschließlich des Aufnahmetages zu reduzieren. Damit würden gleichzeitig die Unterschiede zwischen Aufnahme- und Entlassungswerten trennschärfer dargestellt. Die Entwickler des SF-36 haben zu diesem Zweck eine „Akutversion“ des SF-36 entwickelt, die den Zeithorizont der einzelnen Fragen auf eine Woche eingrenzt (z. B.: „Wie würden Sie Ihren Gesundheitszustand im Allgemeinen während der vergangenen Woche einschätzen?“) [33]. Mit einer prospektiven Erfassung des Gesundheitszustands bei Aufnahme einschließlich des Aufnahmetages und erneut ca. 4 Wochen nach Entlassung aus der stationären Versorgung lässt sich damit die Ergebnisqualität während des stationären Aufenthalts sinnvoll darstellen.

Obwohl die inhaltliche Interpretation der Ergebnisse nicht im Vordergrund dieser Studie stand, so zeigte sich doch eine Notwendigkeit, patientenzentrierte Fragebogen zum Gesundheitszustand bei der Bewertung von Ergebnisqualität einzusetzen, da sich eine diskrepante Bewertung durch Patienten, Ärzte und Pflegepersonal deutlich zeigte. Proxi-Bewertungen des Gesundheitszustands anhand des SF-36 wurden als sehr gut beschrieben, wenn sie durch nahe Verwandte oder im Haushalt lebende Personen beantwortet werden [34]. Die vorliegenden Daten können jedoch nicht vollständig klären, ob die von uns beobachteten unterschiedlichen Bewertungen durch die unterschiedlichen Befragungszeitpunkte verursacht wurden (med. Personal bei Entlassung, Patienten 4 Wochen später) oder ob es sich tatsächlich um eine unterschiedliche Bewertung desselben Konstrukts durch unterschiedliche Personenkreise handelt. Zunächst sollte sowohl die professionelle wie auch die patientenbezogene Bewertung in die Erfassung von Ergebnisqualität der stationären Behandlung eingehen, bevor weitere Untersuchungen mit größeren Fallzahlen mehr Klarheit schaffen.

Die Schlussfolgerungen aus der vorliegenden Studie sind 1., dass eine patientenzentrierte Bewertung des Gesundheitszustands in die Beurteilung der Ergebnisqualität der stationären Versorgung eingehen sollte, da sie eine weitere Dimension der Beurteilung beisteuert, 2., dass kurze Instrumente wie der SF-12 eingesetzt werden können, 3., dass eine Patientenbefragung ca. 30 Tage nach Entlassung praktikabel ist und 4., dass auf Patienten mit multiplen Funktionsstörungen und auf Patienten, die nicht nach Hause entlassen wurden, besonders eingegangen werden muss, um den Rücklauf zu erhöhen oder minimale Aussagen über diese Patientengruppe zu ermöglichen.

### Danksagung

Das Qualitätsmodell Krankenhaus (QMK) ist ein gemeinsames Projekt des AOK-Bundesverbandes sowie der ASKLEPIOS Kliniken GmbH und der HELIOS Kliniken GmbH. Die wissenschaftliche Begleitung erfolgt durch die Ludwig-Maximilians-Universität München in Kooperation mit der Harvard Medical School sowie dem Wissenschaftlichen Institut der Ortskrankenkassen. Wir danken allen Pflegekräften und Ärzten der beteiligten Testkliniken für ihre Unterstützung dieser Studie und bei den Mitarbeitern der wissenschaftlichen Institute für ihr Engagement in diesem Projekt.

### Anhang

#### *Beteiligte der QMK-Pilotstudie in alphabetischer Reihenfolge*

**AOK Bundesverband:**  
Dr. Wulf-Dietrich Leber  
Dr. Nicole Schlottmann

**ASKLEPIOS KLINIKEN GMBH:**  
PD Dr. Wolfgang Hartmann  
Stefanie Hewig  
Dr. Axel Paeger  
Dr. Wolfram Schulz

**HELIOS Kliniken GmbH:**  
Petra Euler  
Michael Knupfer  
Dr. Michael Liebetrau  
Dr. Bernhard Walter

**Ludwig-Maximilians-Universität München:**  
Astrid Manstetten  
Prof. Dr. Günter Rauh  
Dr. Oliver Sangha  
Dr. Sebastian Schneeweiss (Harvard Medical School)  
Dr. Manfred Wildner

**Wissenschaftliches Institut der AOK:**  
Astrid Eichenlaub  
Jürgen Klauber  
Joachim Klose

**Konsultierte internationale Experten:**  
Prof. S. Greenfield, Prof. S. Kaplan (New England Medical Center, Tufts University)  
Prof. H. Palmer (Harvard School of Public Health)  
Prof. M. Liang (Harvard Medical School), alle in Boston

## Literatur

- <sup>1</sup> WHO. World Health Organization Constitution. Geneva: World Health Organization, 1947
- <sup>2</sup> Siegerest HE. Medicine and Human Welfare. New Haven: Yale University Press, 1941
- <sup>3</sup> Ware JE, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30: 473-483
- <sup>4</sup> National Committee for Quality Assurance. URL: <http://www.ncqa.org>, 24.1.2001
- <sup>5</sup> Joint Commission on Accreditation of Healthcare Organizations. ORYX: The Next Evolution in Accreditation. URL: [http://www.jcaho.org/oryx\\_frm.html](http://www.jcaho.org/oryx_frm.html), 24.1.2001
- <sup>6</sup> Lee KY, Loeb JM, Nadzam DM, Hanold LS. An overview of the Joint Commission's ORYX initiative and proposed statistical methods. *Health Services & Outcomes Research Methodology* 2000; 1: 63-73
- <sup>7</sup> Medicare Health Outcomes Project. URL: <http://www.amihealth.com>, 24.1.2001
- <sup>8</sup> Küchler TH, Rappat S, Holst K, Craul J, Wood-Dauphinee S, Henne-Bruns D, Schreiber HW. Zum Einfluss psychosozialer Betreuung auf Lebensqualität und Überlebenszeit von Patienten mit gastrointestinalen Tumoren. *Forum Deutsche Krebsgesellschaft* 1996; 11: 452-466
- <sup>9</sup> Eypasch E, Wood-Dauphinee S et al. Der Gastrointestinale Lebensqualitätsindex (GLQI): Ein klinimetrischer Index zur Befindlichkeitsmessung in der gastroenterologischen Chirurgie. *Chirurg* 1993; 64: 254-274
- <sup>10</sup> Greiner W, Uber A. Gesundheitsökonomische Studien und der Einsatz von Lebensqualitätsindizes am Beispiel des I-Q-Indexes EQ-5D (Euro-QoL). In: Ravens-Sieberer U, Cieza A, v. Steinbüchel N, Bullinger M (Hrsg). *Lebensqualität und Gesundheitsökonomie in der Medizin*. Landsberg: Ecomed, 2000
- <sup>11</sup> Bitzer EM, Dörning H, Busse R, Schwartz FW. Hospital outcomes research in Germany. Results from a retrospective survey among sickness fund beneficiaries. *Med Care* 1997; 35: OS112-OS122
- <sup>12</sup> Schneeweiss S, Sangha O. Weiterentwicklung von Krankenhausbetriebsvergleichen: Wie wichtig ist Risiko-Adjustierung für den Krankenhausvergleich? In: Sieben G, Litsch M (Hrsg). *Berlin: Springer Krankenhausbetriebsvergleiche*, 2000
- <sup>13</sup> Schneeweiss S, Sangha O, Manstetten A et al. Identifikation von Indikatoren der Ergebnisqualität in der internistischen Krankenhausversorgung: Ergebnisse der QMK Pilotstudie. *Gesundheitsökonomie und Qualitätsmanagement*, 2000; 6: 173-182
- <sup>14</sup> Qualitätsmodell Krankenhaus. URL: <http://www.qmk-online.de/>, 24.1.2001
- <sup>15</sup> Tarlov AR, Ware JE, Greenfield S, Nelson EC, Perrin E, Zubkoff M. The Medical Outcome Study: An application of methods for monitoring the results of medical care. *JAMA* 1989; 262: 925-930
- <sup>16</sup> Rogers WH, Wagner AK, Shkiennik A, Kulich R, Barrett AM, Wittink H, Latsch KE, Ware JE, Wurm WH, Carr DB. Outcomes assessment in chronic pain treatment using the SF-36. *Drug Information Journal* 1996; 30: 865
- <sup>17</sup> Aaronson NK, Acquadro C, Alonso J et al. International quality of life assessment (IQOLA) project. *Quality of Life Res* 1992; 1: 349-351
- <sup>18</sup> Bullinger M. German translation and psychometric testing of the SF-36 Health Survey: Preliminary results from the IQOLA Project. *Social Science and Medicine* 1995; 41: 1359-1366
- <sup>19</sup> Ware JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey (SF-12): construction of scales and preliminary tests of reliability and validity. *Med Care* 1996; 32: 220-233
- <sup>20</sup> Bellach BM, Radoschewski M. Gesundheitsbezogene Lebensqualität als Parameter der Gesundheit von Bevölkerungen. In: Ravens-Sieberer U, Cieza A, v. Steinbüchel N, Bullinger M (Hrsg). *Lebensqualität und Gesundheitsökonomie in der Medizin*. Landsberg: Ecomed, 2000
- <sup>21</sup> Greenfield S, Sullivan L, Dukes KA, Silliman R, D'Agostino R, Kaplan SH. Development and testing of a new measure of case mix for use in office practice. *Medical Care* 1995; 33: AS47-55
- <sup>22</sup> Landgraf JM, Nelson EC. Summary of the WONCA/COOP International Health Assessment Field Trial. *Aust Fam Physician* 1992; 21: 255-257, 260-262, 266-269 (3)
- <sup>23</sup> Armitage P, Berry G. *Statistical Methods in Medical Research*. London: Blackwell, 1994: 163 ff.
- <sup>24</sup> Gardner MJ, Altman DG. *Statistics with confidence*. London: BMJ Publishing, 1989
- <sup>25</sup> Hosmer DW, Lemeshow S. *Applied logistic regression*. New York: Wiley, 1989
- <sup>26</sup> Daniel WW. *Biostatistics: A foundation for analysis in the health sciences*. New York: Wiley, 1991: 274 ff.
- <sup>27</sup> SAS Institute Inc. *SAS/STAT. User's Guide*. Version 6. Fourth Edition, Vol.1&2. Cary: SAS Institute, 1989
- <sup>28</sup> Kane RL. *Understanding health care outcomes research*. Gaitersburg: Aspen, 1997
- <sup>29</sup> Sangha O, Schneeweiss S. Lebensqualität und Qualitätsmanagement im Krankenhaus. In: Ravens-Sieberer U, Cieza A, v. Steinbüchel N, Bullinger M (Hrsg). *Lebensqualität und Gesundheitsökonomie in der Medizin*. Landsberg: Ecomed, 2000
- <sup>30</sup> Ware JE, Brook RH, Rogers WH, Keeler EB, Davies AR, Sherbourne CD, Goldberg GA, Camp P, Newhouse JP. Comparison of health outcomes at a health maintenance organization with those of fee-for-service care. *Lancet* 1986; 1017-1022
- <sup>31</sup> Ware JE, Kosinski M, Keller SD. SF-12: How to score the SF-12 physical and mental health summary scales. Lincoln: Quality-Metric Inc, 1998
- <sup>32</sup> Streiner DL, Norman GR. *Health Measurement Scales*. Oxford: Oxford University Press, 1994
- <sup>33</sup> Kirchberger I. Der SF-36-Fragebogen zum Gesundheitszustand: Anwendung, Auswertung und Interpretation. In: Ravens-Sieberer U, Cieza A, v. Steinbüchel N, Bullinger M (Hrsg). *Lebensqualität und Gesundheitsökonomie in der Medizin*. Landsberg: Ecomed, 2000
- <sup>34</sup> Wagner AK, Keller SD, Kosinski M, Ware JE. Are SF-36 data provided by proxy reliable and valid substitutes for self-reported data? The case of epilepsy patients. *Drug Information Journal* 1996; 30: 864-865

Dr. med. Sebastian Schneeweiss, ScD

Department of Medicine  
Brigham and Women's Hospital (BL1341) and Harvard Medical School  
221 Longwood Ave  
Boston, MA 02115  
USA

E-mail: [schneeweiss@post.harvard.edu](mailto:schneeweiss@post.harvard.edu)