

**Entwicklung eines
Qualitätssicherungsprogramms
im Bereich der Prozess- und Ergebnisqualität
sowie der Patientenzufriedenheit
für stationäre und ambulante Einrichtungen
der geriatrischen Rehabilitation**

Abschlussbericht

des

**BQS-Instituts für Qualität
und Patientensicherheit**

N. Birkner, L. Kwietniewski

Hinweis zur Gleichstellung:

Um die Texte in diesem Bericht verständlich und gut lesbar zu gestalten, verwenden wir überwiegend die männliche Sprachform. In solchen Fällen sind jedoch stets beide Geschlechter gemeint.

Düsseldorf, 05. September 2018

Inhalt

Abbildungsverzeichnis	I
Tabellenverzeichnis	II
Anlagenverzeichnis.....	III
1 Einleitung	1
2 Gesamtkonzeption.....	1
2.1 Projektziele	1
2.2 Herausforderungen.....	1
2.3 Eckpunkte des QS-Verfahrens.....	1
2.4 Erhebungsdesign der Qualitätsmessungen.....	2
2.5 Projektdurchführung	3
3 Exploration.....	4
3.1 Ergebnisqualität.....	4
3.1.1 Ausgangspunkt: individuelles Rehabilitations-Ziel.....	4
3.1.2 Systematik der Rehabilitations-Zielkategorien	5
3.1.3 Operationalisierung der Rehabilitations-Zielkategorien der Ebene II.....	8
3.1.4 Standardisierter Prozess der Rehabilitations-Zielkategorien-Auswahl	9
3.2 Prozessqualität und Zufriedenheit.....	9
3.3 Confounder.....	10
4 Erstentwürfe der Fragebögen und weitere Unterlagen	11
4.1 Patienten- und Bezugspersonenbögen	11
4.2 Behandlerbogen	12
4.3 Einwilligungserklärung	13
4.4 Handlungsanweisung für die Rehabilitations-Zielvereinbarung sowie die Auswahl und Dokumentation der Rehabilitations-Zielkategorien	14
4.5 Kognitive Pretests.....	14
5 Pretest	15
5.1 Ziele.....	16
5.2 Methode	16
5.3 Ablauf	17
5.3.1 Einrichtungsakquise.....	17
5.3.2 Schulung der Teilnehmer.....	17
5.3.3 Durchführung des Pretests	18
5.4 Ergebnisse.....	18
5.4.1 Stichprobe	18

5.4.2	Item-Analyse.....	19
5.4.3	Item-Revision.....	21
5.4.4	Auswahl der Rehabilitations-Zielkategorien	22
5.4.5	Nachträgliche Änderungen der ausgewählten Rehabilitations-Zielkategorien	23
5.4.6	Berechnung der Ergebnisqualitäts-Scores.....	24
5.4.7	Reliabilität der Ergebnisqualitäts-Scores.....	28
5.4.8	Validierung der Ergebnisqualitäts-Scores	29
5.4.9	Risikoadjustierung der Ergebnisqualitäts-Scores	29
5.4.10	Prüfung alternativer Abfragen bei Ausfall von Patienten.....	30
5.4.11	Feedback der Befragten zur Befragung.....	32
5.4.12	Änderungen für die Pilotphase	33
6	Pilotphase	35
6.1	Ziele.....	35
6.2	Methode	35
6.3	Ablauf	36
6.3.1	Einrichtungssakquise.....	36
6.3.2	Schulung der Teilnehmer.....	36
6.3.3	Durchführung der Pilotphase	36
6.4	Ergebnisse.....	36
6.4.1	Stichprobe	36
6.4.2	Item-Analyse.....	37
6.4.3	Auswahl der Rehabilitations-Zielkategorien	40
6.4.4	Reliabilität der Ergebnisqualitäts-Scores.....	43
6.4.5	Validität der Ergebnisqualitäts-Scores	45
6.4.6	Risikoadjustierung der Ergebnisqualitäts-Scores	46
6.4.7	Prüfung alternativer Abfragen bei Ausfall von Patienten	47
6.4.8	Feedback der Befragten zur Befragung	48
6.4.9	Empfehlungen und Vorgaben für die Routine	48
7	Ergebnisdarstellung für die Einrichtungen.....	49
8	Ausblick.....	51
8.1	Risikoadjustierung	51
8.2	Revision des Behandlerbogens	52
8.3	Überprüfung der RZK und der RZK-Systematik	52
8.4	Aufwände zur Realisierung des Verfahrens	52
8.4.1	Rehabilitations-Einrichtungen	52
8.4.2	Auswertungsstelle nach § 299 Abs. 3 SGB V	53

8.5	Nutzung von Daten nach § 299 SGB V für das beschlossene Verfahren	55
8.6	Mobile geriatrische Rehabilitation	56
8.6.1	Regelmäßige Erhebung von Basis- und Strukturdaten	56
8.6.2	Jahresbericht	57
8.6.3	Qualitätssicherung in der mobilen geriatrischen Rehabilitation	58
8.6.4	Fazit	59
9	Bewertung des beschlossenen QS-Verfahrens	59
10	Danksagung	61
11	Literatur/Quellen	62

Abbildungsverzeichnis

Abbildung 1: Häufigkeitsverteilung der ausgewählten Rehabilitations-Zielkategorien nach der RZK-Systematik sortiert (Legende der Rehabilitations-Zielkategorien s. Anlage 3.2)	7
Abbildung 2: Häufigkeitsverteilung der ausgewählten Rehabilitations-Zielkategorien absteigend nach Häufigkeit sortiert (Legende der Rehabilitations-Zielkategorien s. Anlage 3.2)	7
Abbildung 3: Häufigkeitsverteilung der Anzahl ausgewählter Rehabilitations-Zielkategorien der Ebene II.....	22
Abbildung 4: Häufigkeitsverteilung der ausgewählten RZK absteigend nach Häufigkeit sortiert	23
Abbildung 5: RZK 1.2 Körperfunktion: Nahrungsaufnahme, Schlucken; Antwortbeispiel	24
Abbildung 6: Häufigkeitsverteilung für ein 3-stufiges Item und Vorher-Nachher-Messung	25
Abbildung 7: z-Werte-Verteilungen für jede der drei t0-Stufen und den t2-Ergebnissen.....	26
Abbildung 8: Häufigkeitsverteilung der Anzahl ausgewählter Rehabilitations-Zielkategorien der Ebene II.....	40
Abbildung 9: Häufigkeitsverteilung der ausgewählten RZK absteigend nach Häufigkeit sortiert	40

Tabellenverzeichnis

Tabelle 1.a: Erhebungsdesign, wenn der Patient und ggf. die Bezugsperson an der Befragung teilnimmt	3
Tabelle 2.b: Erhebungsdesign, wenn anstatt des Patienten nur die Bezugsperson an der Befragung teilnimmt	3
Tabelle 3: Finale Liste der Rehabilitations-Zielkategorien der Ebenen I und II.....	6
Tabelle 4: Zur Exploration der Prozesse und zur Durchführung von kognitiven Pretests besuchte geriatrische Rehabilitationseinrichtungen	15
Tabelle 5: Erhebungsdesign des Pretests	16
Tabelle 6: Durchgeführte Schulungen	17
Tabelle 7: Korrelationen zwischen Patienten- und Bezugspersonenangaben, Alpha-Fehler (ρ) und Fallzahl (N).....	30
Tabelle 8: Häufigkeitsverteilung der Schmerzmitteleinnahme zu Rehabilitations-Beginn (t_0) und zwei Wochen nach Rehabilitations-Ende (t_2)	43
Tabelle 9: Cronbachs Alpha für den Ergebnisqualitäts-Score der RZK „Schmerzen“ in Abhängigkeit von der Schmerzmitteleinnahme zu Rehabilitations-Beginn (t_0) und zwei Wochen nach Rehabilitations-Ende (t_2)	43
Tabelle 10: Erhebungsdesign in der Routine, wenn der Patient an der Befragung teilnimmt	47
Tabelle 11: Erhebungsdesign in der Routine, wenn anstatt des Patienten die Bezugsperson an der Befragung teilnimmt.....	47

Anlagenverzeichnis

Anlage 1	Gesamtkonzeption für ein Qualitätsmessverfahren, Stand: 03.02.2015
Anlage 2	Liste der Interviewten
Anlage 2.1	Leitfaden für explorative Interviews mit Experten
Anlage 2.2	Leitfaden für explorative Interviews mit Patienten- und Angehörigenvertretern
Anlage 3.1	Rehabilitations-Zielkategorien-Raster zur Patienteneinordnung
Anlage 3.2	Rehabilitations-Zielkategorien-Nummerierung alt, Stand: Frühjahr 2016
Anlage 4	Arbeitsmaterial für Klinik-Workshops
Anlage 5	Operationalisierungen aller 33 Rehabilitations-Zielkategorien für den Pretest
Anlage 6.1	Vollständiger Item-Pool
Anlage 6.2	Systematik Entwicklungsprojekt Geriatrie
Anlage 6.3	Confounder-Pool Bewertung
Anlage 7.1	Einwilligungserklärung-Patient
Anlage 7.2	Einwilligungserklärung-Betreuer
Anlage 8	Handlungsanweisung Rehabilitations-Zielvereinbarung
Anlage 9	Fragebogen zur Bewertung der Schulungen
Anlage 10	Prozessqualität und Patientenzufriedenheit: Item-Kennwerte, Pretest
Anlage 11	Faktorenanalysen: Ergebnisse Pretest
Anlage 12	Übereinstimmung der RZK-Bewertungen, Patienten und Bezugspersonen
Anlage 13	Confounder Prüfung auf Signifikanz Pretest
Anlage 14	Drop-Out-Gründe in der Pilotphase
Anlage 15.1	Itemkennwerte, Pilotphase, PQ/Z, Patient stationär
Anlage 15.2	Itemkennwerte, Pilotphase, PQ/Z, Patient ambulant
Anlage 15.3	Itemkennwerte, Pilotphase, PQ/Z, Bezugsperson
Anlage 16	Faktorenanalysen: Ergebnisse Pilotphase
Anlage 17	Muster-Ergebnisbericht
Anlage 18	Leitfaden für Feedback der Einrichtungen
Anlage 19	Feedback der Einrichtungen zum Pretest
Anlage 20	Definition der Qualitäts-Scores
Anlage 21	Patienten- und Bezugspersonenfragebögen
Anlage 22	Behandlerbögen

1 Einleitung

Der GKV-Spitzenverband betreibt seit 2012 gemäß § 135a Abs. 2 Nr. 1 SGB V ein bundesweit verbindliches, einheitliches und routinemäßig angewandtes Qualitätssicherungsverfahren im Bereich der ambulanten und stationären Rehabilitation, das „QS-Reha[®]-Verfahren“. Mit diesem Verfahren werden die Qualitätsdimensionen Struktur-, Prozess- und Ergebnisqualität sowie Patientenzufriedenheit erfasst. Für die externe Qualitätssicherung unterschiedlicher Indikationsbereiche und Patientengruppen werden spezifische Verfahren angewendet, die die Qualität der Leistungserbringung objektiv erfassen und einen Vergleich zwischen den Einrichtungen erlauben, um auf diese Weise zu einer kontinuierlichen Weiterentwicklung der Rehabilitationsleistungen beizutragen.

Bis zum Jahr 2017 wurden im Rahmen des QS-Reha[®]-Verfahrens für geriatrische Rehabilitationseinrichtungen nur Merkmale der Strukturqualität erfasst. Der GKV-Spitzenverband beschloss gemeinsam mit den Spitzenverbänden der Leistungserbringer im Jahr 2013 die Entwicklung der fehlenden Verfahren zur Messung der Prozess- und Ergebnisqualität sowie der Patientenzufriedenheit. Für die Ausschreibung sowie die Begleitung und Unterstützung der Entwicklung der Instrumente wurde ein Steuerungskreis aus Vertretern von Spitzenverbänden der Leistungserbringer, Leistungsträger und Fachgesellschaften gebildet. Unter Leitung des GKV-Spitzenverbandes beteiligte sich der Steuerungskreis intensiv an jeder Phase der Instrumentenentwicklung, die die Konzeption der Qualitätsmessung, die Exploration relevanter Qualitätsaspekte sowie die Fragebogenkonstruktion selbst, verbunden mit der Bewertung der empirischen Befunde aus dem Pretest und der Pilotphase, umfasst.

2 Gesamtkonzeption

2.1 Projektziele

Ziel des Projektes war die Entwicklung eines Qualitätssicherungsverfahrens für die Versorgung von Patienten in der geriatrischen Rehabilitation, das sich in das QS-Reha[®]-Verfahren einfügt und die Lücke in den bisher verfügbaren indikationsspezifischen QS-Modulen des QS-Reha[®]-Verfahrens schließt. Das QS-Verfahren soll dazu die Prozess- und Ergebnisqualität sowie die Patientenzufriedenheit für stationäre und ambulante Einrichtungen der geriatrischen Rehabilitation valide und reliabel erfassen.

2.2 Herausforderungen

Die besonderen Herausforderungen für dieses QS-Verfahren erwachsen aus den spezifischen Eigenschaften der geriatrischen Patienten und lagen dann in den daraus resultierenden konzeptionellen Besonderheiten des Messverfahrens.

Geriatrische Patienten zeichnen sich durch Multimorbidität in Verbindung mit einem höheren Lebensalter (meist über 70 Jahre) aus. Das bedingt ein inhomogenes Beschwerdebild und ein Patientenkollektiv mit einem hohen Anteil kognitiver Einschränkungen. Der erste Punkt macht es erforderlich, ein Verfahren zu entwickeln, das konzeptionell über eine auf alle unterschiedlichen Fälle anwendbare standardisierte Methodik verfügt. Der zweite Punkt impliziert, die zu erwartenden Ausfälle möglichst klein zu halten, um dennoch belastbare Einrichtungsergebnisse zu gewährleisten.

Der das Patientenkollektiv kennzeichnenden Inhomogenität der Beschwerdebilder wird mit einem standardisierten Verfahren zur individualisierten Messung der Ergebnisqualität begegnet (s. Kapitel 3.1). Dem kognitiv bedingten Ausfall von Patienten für die Befragung wird durch die Einbeziehung der Bezugspersonen in das Verfahren entgegengetreten (s. Tabelle 1.b).

2.3 Eckpunkte des QS-Verfahrens

Nach intensiven Erörterungen mit dem Steuerungskreis wurden folgende Eckpunkte für das Entwicklungsprojekt Geriatrie verabschiedet (s. Gesamtkonzeption, Anlage 1):

Prozessqualität, Ergebnisqualität und Patientenzufriedenheit werden durch Befragungen erhoben.

Die Prozessqualität wird ereignisorientiert mit Reporting-Fragen abgefragt und die Patientenzufriedenheit von Teilaspekten der stationären bzw. ambulanten Versorgung mit Rating-Fragen.

Für stationär bzw. ambulant behandelte Patienten, die teilweise unterschiedliche Prozesse während ihrer Rehabilitations-Behandlung erleben, werden spezifische Befragungsinstrumente entwickelt.

Die Messung der Ergebnisqualität erfolgt auf Basis von Patientenbefragungen. Es wird ein Vorher-Nachher-Vergleich durchgeführt, wobei zu den Messzeitpunkten, Rehabilitations-Beginn (t0) und zwei Wochen nach Entlassung aus der Rehabilitation (t2) jeweils der aktuelle Status abgefragt wird und die errechnete Differenz beider Messungen die Ergebnisqualität beschreibt (indirekte Veränderungsmessung).

Anders als in den indikationsspezifischen QS-Reha[®]-Modulen ist für die Patienten der geriatrischen Rehabilitation das hohe Alter und deren Multimorbidität charakteristisch. Zudem liegt bei ihnen kein einheitliches Krankheitsbild vor. Ein standardisiertes Messverfahren müsste alle Variationen, die für das Patientenkollektiv relevant sind, berücksichtigen können. Im vorliegenden Fall ist es jedoch ausgeschlossen, alle Patienten alle möglichen Erfolgsparameter abzufragen, zumal immer nur ein Ausschnitt davon für einen Patienten von Relevanz ist. Das Gebot der Datensparsamkeit steht diesem Ansatz entgegen. Zur Lösung dieser praktischen Problematik wird die Befragung individualisiert. D.h., jeder Patient wird in dem Verfahren nur das gefragt, was für ihn relevant ist. Bezogen auf das QS-Verfahren heißt das, dass die Ergebnisqualität sich nur auf die therapeutischen Leistungen bzw. auf Therapieeffekte solcher Aspekte bezieht, die während der Rehabilitation im Fokus der therapeutischen Intervention standen. Die Standardisierung der individualisierten Befragung besteht zum einen darin, dass der Auswahlprozess, also die Festlegung dessen, was für den einzelnen Patienten relevante Aspekte der Ergebnisqualität darstellen, standardisiert erfolgt. Zum anderen erfolgen die Messungen der einzelnen Leistungsaspekte ebenfalls standardisiert. Details der Ergebnisqualitätsmessung werden in Kapitel 3.1 ausgeführt.

Im Gesamtkonzept, das zu Beginn der Entwicklung des QS-Verfahrens beschlossen wurde, war man übereingekommen: Zur Messung der unterschiedlichen Qualitätsdimensionen werden Patienten und deren Bezugspersonen¹ befragt, sofern diese verfügbar und willens sind. Die Einbeziehung von Bezugspersonen in das QS-Verfahren wurde als notwendig erachtet, um den Ausfall von Angaben kognitiv eingeschränkter Patienten zu kompensieren. Außerdem wird damit der Wichtigkeit der Bezugsperson für die Behandlung der geriatrischen Patienten Rechnung getragen. Daher gibt es in dem Verfahren für die Bezugspersonen eine eigene spezifische Befragung zur Prozessqualität und zur Zufriedenheit, die sich an dem orientiert, was die Bezugspersonen selbst bei der Begleitung ihrer Angehörigen und deren Rehabilitation erlebt haben².

Wenn Patienten an der Befragung nicht teilnehmen wollen oder können, dafür aber Bezugspersonen für die Befragung gewonnen werden konnten, dann werden diese zur Messung der Ergebnisqualität herangezogen und damit ins QS-Verfahren einbezogen. In diesen Fällen bekommen sie zu Rehabilitations-Beginn (t0) und zwei Wochen nach Rehabilitations-Ende (t2) exakt die Fragen vorgelegt, die sonst die Patienten zu diesen Zeitpunkten zur Bestimmung der Ergebnisqualität hätten beantworten sollen³.

Die zur Ergebnismessung abzufragenden Merkmale und Stufen der Leistungsfähigkeit sollen für Laien durch Selbst- und Fremdbeurteilung, also für Patienten und Bezugspersonen gleichermaßen gut beobachtbar und damit reliabel und valide erfassbar sein.

2.4 Erhebungsdesign der Qualitätsmessungen

Bei diesem QS-Verfahren gibt es drei Befragtengruppen (Patienten, Bezugspersonen, Behandler), drei Messzeitpunkte (Rehabilitations-Beginn (t0), Rehabilitations-Ende (t1), zwei Wochen nach Rehabilitations-Ende (t2) sowie die Messung dreier Qualitätsdimensionen (Prozessqualität: PQ,

¹ Im Verlauf der Entwicklung wurde das Verfahren dahin geändert, dass in der Routine in keinem Fall Patienten und Bezugspersonen parallel befragt werden, sondern nur noch Patienten oder alternativ Bezugspersonen.

² Die spezifische Befragung zur Prozessqualität und zur Zufriedenheit der Bezugspersonen wurde entwickelt und im Pretest und in der Pilotphase empirisch überprüft. Auf Grund der geringen Teilnehmerzahlen in der Pilotphase wurde auf deren Einsatz im Routinebetrieb verzichtet.

³ Dieser Ansatz wurde nach dem Pretest insofern modifiziert, als die Formulierung der Einleitung der Items für die Befragung der Bezugspersonen geändert wurde, um der anderen Fragenperspektive (Selbstbeobachtung vs. Fremdbeobachtung) Rechnung zu tragen. Der abgefragte Inhalt sowie das Antwortformat blieben dabei unverändert (s. Kapitel 3.1.3).

Ergebnisqualität: EQ, Patientenzufriedenheit: Z). Zudem werden diverse Confounder⁴ für eine spätere Risikoadjustierung erhoben. Wann welche Daten von welchen Personen abgefragt werden, stellt das Erhebungsdesign dar.

Grundsätzlich werden dabei zwei Designs unterschieden, je nach dem, ob der Patient oder die Bezugsperson an der Befragung teilnimmt. Wenn der Patient willens und kognitiv dazu in der Lage ist, einen Fragebogen selbständig auszufüllen, dann erfolgt die Befragung zur Messung der Ergebnisqualität zu zwei Messzeitpunkten über den Patienten (s. Tabelle 1.a). Die Prozessqualität und die Zufriedenheit werden spezifisch für ambulant und stationär behandelte Patienten sowie für Bezugspersonen zwei Wochen nach Rehabilitations-Ende abgefragt.

Tabelle 1.a: Erhebungsdesign, wenn der Patient oder ggf. die Bezugsperson an der Befragung teilnimmt

Befragte	Messzeitpunkt		
	Rehabilitations-Beginn	Rehabilitations-Ende	2 Wochen nach Rehabilitations-Ende
Patient, ambulant	EQ		EQ, PQ _a , Z _a
Patient, stationär	EQ		EQ, PQ _s , Z _s
Bezugsperson			PQ _{BP} , Z _{BP}
Behandler	Confounder	Confounder	

Tabelle 1.b: Erhebungsdesign, wenn anstatt des Patienten nur die Bezugsperson an der Befragung teilnimmt

Befragte	Messzeitpunkt		
	Rehabilitations-Beginn	Rehabilitations-Ende	2 Wochen nach Rehabilitations-Ende
Bezugsperson	EQ*		EQ*, PQ _{BP} , Z _{BP}
Behandler	Confounder	Confounder	

*) Parallelversionen für Bezugspersonen zur EQ-Messung (s. Kapitel 3.1.3).

Wenn der Patient an der Befragung nicht teilnimmt, dann wird die Bezugsperson zur Messung der Ergebnisqualität herangezogen, sofern verfügbar und willens, und exakt die Fragen gestellt, die sonst an den Patienten gerichtet worden wären (s. Tabelle 1.b). Zudem werden Prozessqualität und Zufriedenheit spezifisch für Bezugspersonen abgefragt, sofern sie an der Befragung teilnehmen.

2.5 Projektdurchführung

Der Projektstart erfolgte mit der Erteilung des Auftrages durch den GKV-Spitzenverband (GKV-SV), in Abstimmung mit den maßgeblichen Spitzenorganisationen der Rehabilitationseinrichtungen, am 02.10.2013, Vertragsunterzeichnung war am 11.10.2013.

Die Durchführung des Projektes gliedert sich grob in vier Abschnitte. Im 1. Abschnitt wurde in Abstimmung mit dem Auftraggeber die Gesamtkonzeption für das Qualitätssicherungsverfahren entwickelt, die am 03.02.2015 konsentiert vorgelegt wurde (s. Anlage 1).

Im 2. Abschnitt erfolgt die Exploration des Themenfeldes. In Experteninterviews und Gesprächen mit den Behandlern in den Rehabilitations-Einrichtungen vor Ort wurden unterschiedliche Bestandteile des QS-Verfahrens untersucht. Mit den Experteninterviews wurden Items zur Messung der Prozessqualität und Patientenzufriedenheit sowie die zu berücksichtigenden Confounder gewonnen. Die Gespräche mit den Behandlern vor Ort dienten dazu, den allgemeinen Prozess zur Findung der individuellen Rehabilitations-Ziele aus der Praxis abzuleiten, sowie die Systematik der Rehabilitations-Zielkategorien zu diskutieren und den Prozess der Rehabilitations-Zielkategorien-Auswahl sowie die Stufen innerhalb der einzelnen Rehabilitations-Zielkategorien zu definieren.

Im 3. Abschnitt erfolgt die Erstellung der Fragebögen zusammen mit dem Aufbau der erforderlichen IT-Strukturen, die für das Verfahren genutzt werden sollten.

⁴ Damit werden Faktoren bezeichnet, die einen Einfluss auf ein Messergebnis ausüben. Im Rahmen der Qualitätssicherung sind das Bedingungen, die die Leistungserbringer selbst nicht beeinflussen können, z.B. das Geschlecht und das Alter der Patienten.

Im 4. Abschnitt folgt die Empirie. Hier wurden die entwickelten Fragebögen hinsichtlich ihrer Messeigenschaften an zwei unabhängigen Patientenstichproben überprüft. Zudem wurden die Abläufe, wie sie im Routinebetrieb geplant sind, erprobt.

3 Exploration

Im Rahmen der Exploration wurden alle für das Qualitätssicherungsverfahren relevanten Themenfelder untersucht und vielfältige Informationen zum Untersuchungsgegenstand zusammengetragen, um sie dann zu systematisieren und auf dieser Grundlage das Qualitätssicherungsverfahren systematisch aufzubauen.

Zum einen geht es dabei um die Systematik für die individualisierte Ergebnisqualitätsmessung, zum anderen um die Erstellung der Erstentwürfe der einzusetzenden Fragebögen.

Um das Qualitätssicherungsverfahren so nahe wie möglich an die Verhältnisse, wie sie in der täglichen Praxis der ambulanten und stationären geriatrischen Rehabilitation gelebt werden, anzulehnen, wurden diverse Rehabilitations-Einrichtungen besucht (s. Tabelle 3) und vor Ort mit den Mitarbeitern der unterschiedlichen Professionen im Plenum und in homogenen Kleingruppen vertiefende Gespräche geführt.

Zusätzlich wurden 19 halbstrukturierte Experteninterviews mit Vertretern unterschiedlicher Organisationen, Rehabilitations-Einrichtungen und Professionen durchgeführt (s. Anlage 2), um möglichst viele unterschiedliche Aspekte der Prozessqualität und Patienten- und Bezugspersonenzufriedenheit zu erfassen. Außerdem wurden mögliche Confounder erfragt, die nach Ansicht der Experten einen Einfluss auf die Ergebnisqualität und die Zufriedenheit der Patienten und Bezugspersonen haben könnten, ohne von den Rehabilitations-Einrichtungen selbst beeinflusst werden zu können.

3.1 Ergebnisqualität

Bei der Entwicklung der Instrumente zur Messung der Ergebnisqualität wurden Experten aus der Rehabilitation hinzugezogen, auf Therapeuten, Ärzte, Mitarbeiter der Pflege und der Sozialdienste. Zum einen wurden sie zu den internen Abläufen befragt, die zwischen der Aufnahme der Patienten in die Einrichtung und der Festlegung des Therapieplans erfolgen, um auf dieser Basis ein an der Praxis orientiertes allgemeines Prozessmodell für die Therapiezielfindung abzuleiten (s. Kapitel 3.1.1). Zum anderen wurden sie zu konkreten Tätigkeiten und Aktivitäten der Patienten befragt, mit denen Laien (Patienten und Bezugspersonen) den Status des Leistungsvermögens der Patienten durch Selbst- und Fremdbeobachtung valide und reliabel graduell beschreiben können, um bei der Qualitätsmessung als Marker substanzialer Leistungsstufen verwendet zu werden. Zur Strukturierung der Ergebnisqualitätsmessung wurde die ICF-Systematik (WHO 2005) zur systematischen Einteilung der in der geriatrischen Rehabilitation behandelten Störungen herangezogen.

3.1.1 Ausgangspunkt: individuelles Rehabilitations-Ziel

Um ein genaueres Verständnis davon zu bekommen, wie in geriatrischen Rehabilitationseinrichtungen zu Beginn der Rehabilitation die Prozesse ablaufen, bis für einen Patienten die auf ihn zugeschnittenen Therapieziele festgelegt sind, wurden Expertengespräche in unterschiedlichen geriatrischen Rehabilitationseinrichtungen mit Ärzten, Therapeuten und Mitarbeitern der Pflege sowie der Sozialdienste geführt (s. Tabelle 3: Einrichtungen 1 bis 4). Im Klinikum Coburg erhielten wir zudem die Gelegenheit, an zwei regulären Dienstbesprechungen der Therapeuten teilzunehmen, in denen die beobachteten Therapieerfolge der ambulant bzw. stationär behandelten Patienten vorgestellt und die weiteren therapeutischen Maßnahmen besprochen wurden.

Als Ergebnis der Gespräche mit den Experten in den Rehabilitationseinrichtungen wurde folgender standardisierter hierarchischer Prozess für das Verfahren verbindlich festgelegt:

Ausgangspunkt für die individualisierte Messung der Ergebnisqualität ist das mit dem Patienten zu Rehabilitations-Beginn vereinbarte individuelle Rehabilitations-Ziel, das innerhalb der ersten drei Tage des Rehabilitations-Aufenthaltes gemeinsam mit ihm zu bestimmen und festzulegen ist. Das vereinbarte individuelle Rehabilitations-Ziel berücksichtigt Wünsche des Patienten, seine Ressourcen und seine Einschränkungen sowie den nach der Rehabilitation zu erwartenden Kontext und die Einschätzungen der Behandler, dieses Ziel im verfügbaren Zeitrahmen erreichen zu können.

Das vereinbarte individuelle Rehabilitations-Ziel wird einrichtungsseitig dokumentiert (s. Behandlerbogen, Teil A3, Frage 34), wobei das individuelle Rehabilitations-Ziel vom Wortlaut und vom Inhalt her als angestrebte Aktivität der Lebenswelt des Patienten formuliert werden soll (weitere Ausführungen dazu s. Kapitel 4.4).

3.1.2 Systematik der Rehabilitations-Zielkategorien

Bereits in der Gesamtkonzeption war eine Systematik zur Strukturierung der Leistungsaspekte festgelegt worden, die sich mit dem Verfahren differenziert messen lassen sollen. Die Systematik ist hierarchisch aufgebaut und orientiert sich an der ICF-Systematik. Die International Classification of Functioning, Disability and Health (ICF) dient fach- und länderübergreifend als einheitliche und standardisierte Klassifikation zur Beschreibung des funktionalen Gesundheitszustandes, der Behinderung, der sozialen Beeinträchtigung, der relevanten Umgebungsfaktoren eines Menschen sowie möglicher Wechselwirkungen, wobei im vorliegenden Fall die Entitäten der oberen Ebene mit den ICF-Domänen übereinstimmen und die der unteren Ebene den ICF-Kategorien entsprechen (s. Tabelle 2). Da es hier um die Verbesserung von ausgewählten Leistungsaspekten geht, werden sie als Rehabilitations-Zielkategorien der Ebene I (z.B. Mobilität) und Rehabilitations-Zielkategorien der Ebene II (z.B. Händegebrauch) bezeichnet. Rehabilitations-Zielkategorien der Ebene II werden mit RZK abgekürzt.

Die Reihenfolge der Rehabilitations-Zielkategorien der Ebene II wurde auf Anregung von externen Experten abweichend von der ICF-Systematik angeordnet. Die Rehabilitations-Zielkategorien der Ebene II, 3.2 Sprache und 4.6 Gangsicherheit, werden in der ICF-Systematik so nicht differenziert. Sie haben sich in den Gesprächen mit den Experten der besuchten Rehabilitationseinrichtungen als praxisrelevante und abgrenzbare Leistungsmerkmale herausgebildet. Insgesamt werden 33 Rehabilitations-Zielkategorien auf der Ebene II differenziert.

Um die Nützlichkeit und Praktikabilität dieser Systematik zu überprüfen, haben wir vier Einrichtungen im Nachgang der Besuche gebeten, die angesetzten Therapien für die zuletzt aufgenommenen 20 Patienten in diesem Raster einzutragen. Zusätzlich konnte für jeden Fall angegeben werden, ob zusätzliche therapeutische Leistungen geplant wurden, die in der Systematik nicht enthalten sind. Für die Überprüfung wurde die im Frühjahr 2016 gültige RZK-Liste verwendet (s. Anlage 3.1 und 3.2), die in wenigen Punkten von der finalen Liste (s. Tabelle 2) abweicht.

Insgesamt erhielten wir damit die Einordnung von 100 Patienten⁵. Im Durchschnitt wurden 7,8 Rehabilitations-Zielkategorien vergeben, wobei am häufigsten drei Rehabilitations-Zielkategorien angegeben wurden. Die am häufigsten genannten Rehabilitations-Zielkategorien fanden sich im Bereich der Mobilität, mit 64 Nennungen für die „Gangsicherheit“ (RZK 4.6), die damit am häufigsten angekreuzte Kategorie, gefolgt vom „Treppensteigen“ mit 63 Nennungen (RZK 4.8) und „Gehen im Innenbereich“ (RZK 4.7) mit 60 Nennungen. Im Bereich der Selbstorganisation war „Sich waschen“ (RZK 5.1) mit 52 Nennungen die am häufigsten vertretene Kategorie und im Bereich der Körperfunktionen dominierte die Behandlung der „Schmerzen“ (RZK 1.1) mit 48 Nennungen.

⁵ Eine Einrichtung ordnete 20 ambulante und 20 stationäre Fälle ein.

Abbildung 1 und Abbildung 2 zeigen die Verteilungen der ausgewählten Rehabilitations-Zielkategorien sortiert nach der Systematik der Rehabilitations-Zielkategorien bzw. nach der Häufigkeit.

Zusätzlich wurden folgende Rehabilitations-Maßnahmen von den teilnehmenden Einrichtungen genannt, die nicht in der Systematik der Rehabilitations-Zielkategorien aufgeführt waren: Kognition (6), psych. Stabilisation, Antrieb (6), Schwindel (5), kardiopulmonale Belastbarkeit (5), Wundheilung (4), Hilfsmittelversorgung (4), Krafterhalt (3), Hilfsmittelnutzung, -gebrauch (2), Ödeme, Kompressionsbehandlung (2) sowie Affekt, Atmung (Ausdauer), Duschen, Fahrradfahren, Gartenpflege, Prothesenhandlung mit jeweils einer Nennung.

Diese zusätzlich genannten Rehabilitations-Maßnahmen wurden in keinem Fall in die Systematik der Rehabilitations-Zielkategorien aufgenommen, weil sie zu unspezifisch waren (z.B. Krafterhalt), bereits in der Systematik enthalten waren (z.B. Duschen) oder per Definition nicht Bestandteil der Ergebnisqualität sein sollten (z.B. Wundheilung). Zudem sind sie insgesamt zu selten genannt worden.

Tabelle 2: Finale Liste der Rehabilitations-Zielkategorien der Ebenen I und II

	Rehabilitations-Zielkategorie Ebene I	Rehabilitations-Zielkategorie Ebene II	
Körperfunktion	Schmerz	1.1 Schmerzen	
	Funktionen des Verdauungstraktes	1.2 Nahrungsaufnahme, Schlucken	
		1.3 Defäkationsfunktion	
	Funktionen des Urogenitalsystems	1.4 Miktionsfunktion	
Aktivität und Partizipation (Teilhabe)	Lernen und Wissensanwendung	2.1 Lesen	
		2.2 Problemlösen	
		2.3 Konzentration	
	Stimm- und Sprechfunktion (Körperfunktion)	3.1 Sprechen	
		Kommunikation	3.2 Sprache
			3.3 Kommunikation
	Mobilität	3.4 Kommunikationsgeräte benutzen	
		4.1 Körperposition wechseln	
		4.2 In einer Körperposition verbleiben	
		4.3 Gegenstände anheben und tragen	
		4.4 Händegebrauch	
		4.5 Rollstuhlgebrauch	
		4.6 Gangsicherheit	
		4.7 Gehen im Innenbereich	
		4.8 Treppensteigen	
		4.9 Gehen im Außenbereich	
	4.10 Transportmittel nutzen		
	Selbstversorgung	5.1 Sich waschen	
		5.2 Körperpflege	
		5.3 Toilettenbenutzung	
		5.4 Sich kleiden	
		5.5 Essen	
		5.6 Trinken	
	Häusliches Leben	6.1 Mahlzeiten zubereiten	
		6.2 Hausarbeiten erledigen	
		6.3 Beschaffen von Waren und Dienstleistungen	
	Interpersonelle Interaktionen und Beziehung	7.1 Private Beziehungen	
Bedeutende Lebensbereiche	8.1 Wirtschaftliches Leben		
Gemeinschafts-, soziales und staatsbürgerliches Leben	9.1 Gemeinschaftsleben, Erholung, Freizeit		

Abbildung 1: Häufigkeitsverteilung der ausgewählten Rehabilitations-Zielkategorien nach der RZK-Systematik sortiert (Legende der Rehabilitations-Zielkategorien s. Anlage 3.2)

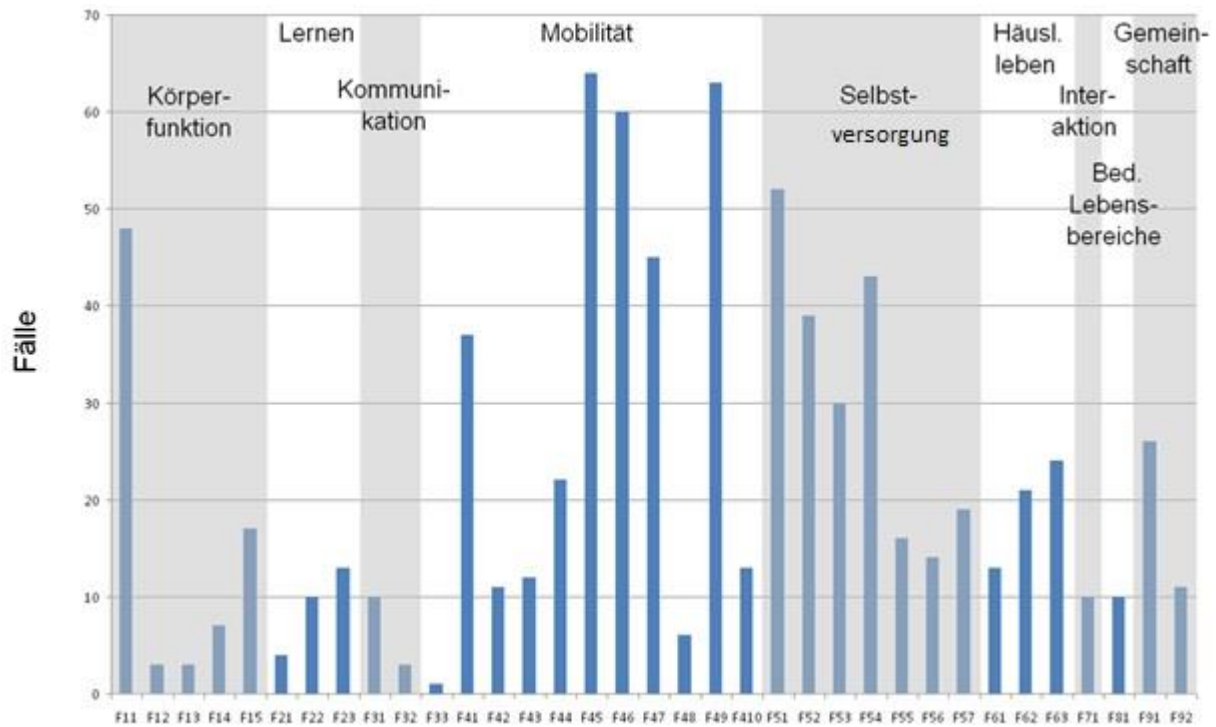
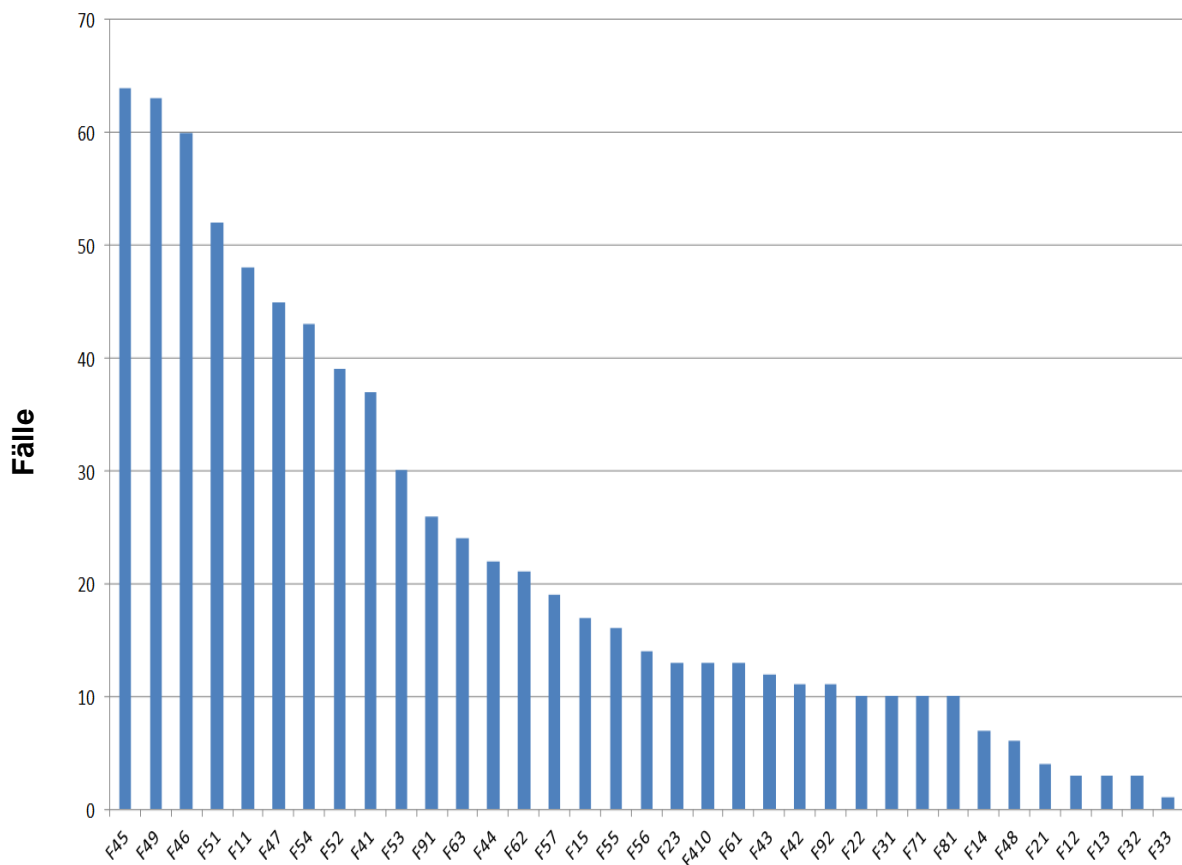


Abbildung 2: Häufigkeitsverteilung der ausgewählten Rehabilitations-Zielkategorien absteigend nach Häufigkeit sortiert (Legende der Rehabilitations-Zielkategorien s. Anlage 3.2)



3.1.3 Operationalisierung der Rehabilitations-Zielkategorien der Ebene II

Bei den Operationalisierungen der 33 Rehabilitations-Zielkategorien wurde der Ansatz verfolgt, die Abfragen ereignisorientiert anzulegen. Dazu sollen direkt beobachtbare Merkmale als Leistungsstufen abgefragt werden, die auf der Verhaltensebene bewertbar sind. Das hilft sowohl den Patienten als auch den Bezugspersonen, den aktuellen Status des Leistungsvermögens reliabel einzuschätzen.

Grundsätzlich sind in diesem Zusammenhang Fragen nach der Potenz bzw. Fragen nach der Performanz möglich, also Fragen nach der Fähigkeit etwas tun zu können ggf. auch ohne einen Realisierungsbeleg und Fragen nach der praktischen Realisierung einer Aktivität. Für dieses Verfahren haben wir uns zur Abfrage der Potenz, also der Frage nach dem Leistungsvermögen entschieden, weil der Fall eintreten kann, dass eine abgefragte Aktivität in der aktuellen Situation des Patienten nicht realisiert werden kann, weil die aktuelle Lebenssituation dies nicht zulässt, z.B. Fragen zu Aktivitäten im Haushalt, wenn man in der Klinik ist. Soweit es der betreffende Gegenstand zulässt, wird daher die Abfrage mit der Formulierung eingeleitet: „Können Sie ...“, die korrekt bleibt, wenn der Patient Erfahrungen damit hat oder nicht.

Zu beiden Messzeitpunkten, t0 und t2, wird jeweils der aktuelle Status abgefragt. Damit werden Erinnerungsfehler vermieden. Bei der Messung der Ergebnisqualität wird nicht zwischen ambulant und stationär behandelten Patienten unterschieden.

Die Therapeuten der besuchten Kliniken hatten im Vorfeld eine Tabelle mit der Bitte erhalten, sofern sie eine der aufgeführten Rehabilitations-Zielkategorien therapeutisch bearbeiten, zu notieren. Daran kann ein Patient praktisch selber gut erkennen, dass er Rehabilitations-bedingte Fortschritte gemacht hat, welche zunehmend höheren Leistungsstufen als Meilensteine betrachtet werden können und welche Hilfsmittel bzw. Hilfsmittelgruppen mit vergleichbarer unterstützender Qualität die Patienten dabei nutzen (s. Anlage 4). Die ausgefüllten Tabellen wurden mit den Therapeuten vor Ort ausführlich diskutiert. Sie bilden die Grundlage der Operationalisierung der 33 Rehabilitations-Zielkategorien.

Für die Rehabilitations-Zielkategorie 3.1 „Stimm- und Sprechfunktion“ wurde zusätzlich Kontakt mit einem Logopäden⁶ aufgenommen, mit dem die Abfragen entwickelt und in mehreren Schritten überarbeitet wurden. Für seine freundliche und kompetente Unterstützung bedanken wir uns sehr herzlich an dieser Stelle.

Die Abfragen zu den einzelnen Rehabilitations-Zielkategorien wurden soweit möglich so aufgebaut, dass die Items nach der Schwierigkeit der Aktivität aufsteigend angeordnet wurden. Die Antworten zu den Items der Potenzabfragen „Können Sie ...“ wurden überwiegend vierstufig erfasst. Die Stufen haben die Antwortlabel: „gar nicht“, „kaum“, „mittelmäßig“ und „gut“.

Beim „Gehen im Innenbereich“, dem „Treppensteigen“ und dem „Gehen im Außenbereich“ wurden Item-Blöcke für einzelne Leistungsstufen gebildet, deren Items sich auf die Nutzung bestimmter Hilfsmittel bzw. Hilfsmittelgruppen beziehen.

Beim „Gehen im Innenbereich“ z.B. wird als erster Item-Block und auf der untersten Stufe das Gehen innerhalb des Zimmers abgefragt, wobei innerhalb des Item-Blocks wie dann auch bei den folgenden dieser Rehabilitations-Zielkategorie die Fragen nach dem Gehen wie folgt differenziert werden:

- ohne Hilfsmittel und ohne Hilfe
- mit Festhalten an Gegenständen
- mit Handstock, Walkingstöcken oder Unterarmstützen
- mit Rollator oder Gehstock
- mit Hilfsperson⁷.

Die Antworten werden jeweils fünfstufig erfasst. Die Stufen haben die Antwortkategorien: „betrifft mich nicht“, „gar nicht“, „kaum“, „mittelmäßig“ und „gut“. Die zusätzliche Antwortkategorie „betrifft mich nicht“ wurde erforderlich, weil nicht davon ausgegangen werden kann, dass die Patienten immer alle Hilfsmittel zur Verfügung haben und nutzen.

In einer weiteren Variante wird mit den Antwortstufen die Unterstützung von Hilfspersonen bzw. der Verzicht darauf mit erfasst. Bei der Rehabilitations-Zielkategorie „Transportmittel nutzen“ z.B. lauten die Antwortkategorien: „nur mit Hilfsperson“, „allein kaum“, „allein mittelmäßig“ und „allein gut“.

⁶ Herr Kai Keßler, Kreiskrankenhaus Gummersbach

⁷ Nach dem Pretest geändert in „Begleitperson“ (s. Kapitel 5.4.7)

Zu der eben beschriebenen Ergebnisqualitätsmessung wurde eine Parallelversion zur Befragung der Bezugspersonen erstellt. Diese Parallelversion wurde erforderlich, weil die Bezugspersonen im Pretest trotz erläuternder Instruktion die Eingangsformulierung der Leistungspotenzabfrage „Können Sie ...“ häufig auf sich selbst und nicht auf den Patienten bezogen, also häufig Schwierigkeiten mit dem Perspektivwechsel hatten. In der Bezugspersonenversion lautet daher nun die Eingangsformulierung: „Kann Ihr(e) Angehörige(r) ...“. Weitere Formulierungen sind entsprechend angepasst worden, so dass statt einer Selbstbeobachtung deutlich eine Fremdbeobachtung abgefragt wird. Entsprechend wurde „betrifft mich nicht“ geändert in „betrifft sie/ihn nicht“. Die Items und die Antwortlabel blieben ansonsten unverändert.

Die einzelnen Operationalisierungen der 33 Rehabilitations-Zielkategorien für den Pretest sind in Anlage 5 einzusehen.

3.1.4 Standardisierter Prozess der Rehabilitations-Zielkategorien-Auswahl

Das im gegenseitigen Einvernehmen zwischen dem Patienten und den Behandlern vereinbarte individuelle Rehabilitations-Ziel wird von den Behandlern in ein Raster übertragen, das auf zwei Ebenen die Faktoren beschreibt, die die Realisierung des vereinbarten Zieles beeinträchtigen und die zur Zielerreichung in der Reha therapeutisch bearbeitet werden sollen. Dieses Raster ist das o.g. an der ICF-Systematik orientierte Ordnungssystem (s. Tabelle 2) der Rehabilitations-Zielkategorien der Ebenen I und II.

Nach diesem Verfahren wird für jeden Patienten festgestellt, welche Defizite während der Rehabilitation zur Erreichung des vereinbarten individuellen Rehabilitations-Ziels therapeutisch zu bearbeiten sind. Vorgabe für die Auswahl von Rehabilitations-Zielkategorien der Ebene II ist, dass zwischen zwei bis zehn Rehabilitations-Zielkategorien ausgewählt werden können, die zudem mindestens aus zwei unterschiedlichen Rehabilitations-Zielkategorien der Ebene I, kommen⁸.

Diese Prozedur formalisiert und standardisiert klinische Abläufe, wie sie nach unserer Beobachtung in ähnlicher Weise in allen sechs besuchten Rehabilitationseinrichtungen zu finden waren.

Jeder Rehabilitations-Zielkategorie der Ebene II steht ein Fragenblock gegenüber, mit dem das aktuelle Leistungsvermögen für den betreffenden Bereich erfasst wird.

Nach diesem Vorgehen erfolgt die individualisierte Ergebnisqualitätsmessung, indem der Leistungsstatus der Patienten nur zu den Leistungsaspekten zu t0 und zu t2 abgefragt wird, die therapeutisch bearbeitet werden, damit sie nach der Entlassung aus der Rehabilitation ihr individuelles Rehabilitations-Ziel ausführen können.

3.2 Prozessqualität und Zufriedenheit

Die Prozessqualität und die Zufriedenheit sollen für Patienten und Bezugspersonen spezifisch abgefragt werden. D.h. die Patienten werden zu den Prozessen befragt, die sie erlebt haben können bzw. erlebt haben sollten. Die Patientenzufriedenheit bezieht sich auf das, was sie während ihres Rehabilitations-Aufenthaltes erlebt haben. Dabei wird zwischen ambulanter und stationärer Behandlung unterschieden.

Bezugspersonen werden zu den patientenspezifischen Aspekten der Prozessqualität und der Patientenzufriedenheit nicht befragt, da sie selbst bei enger Begleitung des Patienten, beim Rooming-in z.B., nicht alle Prozesse wie der Patient miterleben und die Zufriedenheit des Patienten damit bestenfalls aus Äußerungen oder dem Verhalten des Patienten ableiten können. Bezugspersonen werden zu solchen Prozessen befragt, die sie selbst erlebt haben bzw. erlebt haben könnten oder sollten. Die Abfrage der Zufriedenheit bezieht sich entsprechend auf das, was sie als Bezugspersonen erlebt haben bzw. erlebt haben könnten⁹.

Zunächst wurde durch die Befragung von Vertretern unterschiedlicher Interessengruppen, Fachgesellschaften und Professionen der geriatrischen Rehabilitation (s. Anlage 2) ein möglichst umfassender Item-Pool generiert, um dann eine systematische Item-Auswahl zu treffen, die in den ersten Entwurf des Fragebogens eingeflossen ist (s. Kapitel 4.1). Nach mehrfachen, mit dem Steuerungskreis

⁸ Nach dem Pretest geändert in 3 bis 10 Rehabilitations-Zielkategorien der Ebene II aus mindestens drei Rehabilitations-Zielkategorien der Ebene I.

⁹ Gemäß dem Gesamtkonzept wurden für Bezugspersonen spezifische Instrumente zur Messung der Prozessqualität und Zufriedenheit entwickelt und im Pretest und im Piloten geprüft. Wegen geringer Fallzahlen wird auf diese Abfrage in der Routine verzichtet.

erarbeiteten, Überarbeitungen folgten schließlich kognitive Pretests und anschließend die erste empirische Überprüfung der Instrumente im Pretest.

Generierung des Item-Pools

Ziel dieser Phase war die Generierung eines umfassenden Item-Pools, der inhaltlich ein möglichst vollständiges Bild der in der geriatrischen Rehabilitation erlebbaren Prozesse abbildet. Die Maßgabe zur Vollständigkeit ist darin begründet, dass das am Ende einsatzbereite Instrument grundsätzlich nur solche Aspekte bei der Qualitätssicherung berücksichtigen kann, die im Item-Pool enthalten sind, da in allen nachfolgenden Prozessschritten der Fragebogenentwicklung Items aus inhaltlichen und messtheoretischen Gründen davon ausgeschlossen werden und so der Item-Pool sukzessiv auf die Items reduziert wird, die final ins Befragungsinstrument, also die Fragebogen-Erstversion eingehen¹⁰. Nachträglich einzelne Items in ein Instrument aufzunehmen ist aus methodischer Sicht nicht zulässig.

Zur Generierung des Item-Pools wurden in der Zeit vom 03.02. bis zum 28.02.2014 19 Experteninterviews (s. Anlage 2) unter Anwendung eines Leitfadens (für Experten Anlage 2.1 und für Patienten- und Angehörigenvertreter Anlage 2.2) durchgeführt (halbstrukturierte Leitfadeninterviews).

Die Leitfäden umfassen fünf Blöcke:

1. Allgemeine Frage, was unter guter Qualität in der Rehabilitation verstanden wird
2. Klinische Erfolgsmarker für die Vorher-Nachher-Messung
3. Inhalte für die Patientenbefragung zu Rehabilitations-Beginn
4. Inhalte für die Patientenbefragung nach Rehabilitations-Ende
5. Inhalte für die Bezugspersonenbefragung nach Rehabilitations-Ende

Alle Interviews wurden jeweils gemeinsam von einem BQS-Mitarbeiter der Abteilung Integrierte Gesundheitsversorgung (IGV) und einem Mitarbeiter der Abteilung Biometrie und Statistik geführt, nach Erlaubnis der Befragten auditiv aufgezeichnet und simultan verschriftet. Im Nachgang wurde die simultan erstellte Mitschrift für die spätere inhaltsanalytische Auswertung überarbeitet, wobei fehlende Gesprächspassagen ergänzt und fehlerhafte korrigiert wurden.

Gesprächspartner waren Experten des Steuerungskreises, externe Experten (Ärzte, Therapeuten und Vertreter der Pflege) sowie weitere Experten aus Patienten- und Angehörigenvertretungseinrichtungen, die vom Steuerungskreis benannt wurden. Die Liste der Experten sowie der vertretenen Institutionen ist in der Anlage 2 einsehbar.

Die verschrifteten Interviews wurden danach ausgewertet, Prozesse zu identifizieren, die aus Sicht der Befragten auf gute Qualität bei der Versorgung von Patienten der geriatrischen Rehabilitation hindeuten bzw. Patientenzufriedenheit bei den Patienten herstellen. Entsprechend wurden die Angaben der Bezugspersonen ausgewertet.

Für jeden isolierten Prozess wurde jeweils ein Item formuliert, insgesamt 537 Items (s. Anlage 6.1). Diese Items wurden sukzessive in ein hierarchisches Ordnungssystem gebracht (s. Anlage 6.2). Das finale Ordnungssystem diente dann dazu, eine systematische Auswahl aus allen vorliegenden Items vorzunehmen, in der alle wesentlichen Inhalte des Item-Pools vertreten waren. Aus den 537 Items des Item-Pools wurden für den Patientenbogen 91 Items und für den Bezugspersonenbogen 76 Items systematisch so ausgewählt, dass jede Kategorie mindestens mit einem wichtigen Item vertreten ist.

3.3 Confounder

In den einzelnen Fragen-Blöcken der Experteninterviews wurde zudem nach patientenbezogenen Faktoren gefragt, die nach Ansicht der befragten Experten einen deutlichen Einfluss auf den Rehabilitations-Erfolg ausüben können, aber von der Rehabilitationseinrichtung nicht beeinflusst werden, sog. Confounder. Damit wurden potentielle Confounder für eine spätere Risikoadjustierung gesammelt.

In den Verschriftungen der Interviews wurden Confounder identifiziert, anschließend nach deren Auftreten im Versorgungsprozess chronologisch geordnet und dann danach bewertet bzw. eingestuft, ob sie durch Einrichtungen beeinflussbar sind, ob der vermutete Effekt als bedeutsam einzuschätzen ist, ob diese Confounder in den Einrichtungen heterogen verteilt und mit vertretbarem Messaufwand

¹⁰ Entsprechend können Faktoren, die im Item-Pool nicht enthalten sind, von der Faktorenanalyse auch nicht entdeckt werden.

(idealerweise mit einem Item) reliabel messbar sind oder ob ein Messinstrument vorhanden ist (s. Anlage 6.3). Bei durchgehend akzeptablen Angaben wurde der Confounder in den Behandlerbogen aufgenommen (s. Kapitel 4.2). Confounder, die erst nach Entlassung aus der Rehabilitationseinrichtung auftreten können, wurden im Nachbefragungsbogen der Patienten- bzw. Bezugspersonenbefragung (N-Bogen) erfasst.

4 Erstentwürfe der Fragebögen und weitere Unterlagen

Insgesamt gibt es in diesem Verfahren sechs unterschiedliche Fragebögen, drei Patientenbögen

- Beginn-Bogen
- Nachbefragungsbogen ambulant
- Nachbefragungsbogen stationär

zwei Bezugspersonenbögen

- Beginn-Bogen
- Nachbefragungsbogen

und einen Behandlerbogen.

Die individualisierte Messung der Ergebnisqualität ist Bestandteil der Beginn- und der Nachbefragungsbögen (s. Tabelle 1.a und Tabelle 1.b). Die Rehabilitanden werden zu den ausgewählten RZK jeweils zu t0 und zu t2 identisch abgefragt. Dies gilt auch für die Bezugspersonen mit den entsprechenden Parallelversionen der Fragebögen.

Mit den Beginn-Bögen erfolgt die Nullpunktmessung und mit den Nachbefragungsbögen die Wiederholungsmessung für die Bestimmung der Ergebnisqualität. Die Patienten werden bei der Nachbefragung zusätzlich nach der Prozessqualität und nach der Patientenzufriedenheit gefragt. Hierzu gibt es zwei Nachbefragungsbögen für Patienten, die ambulant bzw. stationär behandelt wurden. Für die Bezugspersonen waren spezifische Abfragen zur Prozessqualität und zu deren Zufriedenheit vorgesehen¹¹.

4.1 Patienten- und Bezugspersonenbögen

Der Beginn-Bogen besteht aus drei Teilen,

- 1) der Nullpunktmessung zur Bestimmung der Ergebnisqualität mit der patientenspezifischen Auswahl an RZK,
- 2) der Abfrage, ob das dokumentierte und an dieser Stelle abgedruckte individuelle Rehabilitations-Ziel korrekt wiedergegeben ist, wie wichtig es dem Patienten ist und ob und ggf. welche anderen Ziele er in der Rehabilitation noch erreichen wollte sowie
- 3) schließlich Fragen zum Ausfüllen des Fragebogens.

Der Beginn-Bogen ist für ambulant und stationär behandelte Patienten identisch. In der Bezugspersonenversion werden für Teil eins und zwei Parallelversionen verwendet, die dem Umstand Rechnung tragen, dass mit ihnen Fremdbeobachtungen abgefragt werden.

Die aus dem Item-Pool ausgewählten Items bilden in den Nachbefragungsbögen folgende Teile zur Messung der Prozessqualität (a) und der Patientenzufriedenheit (b):

- a) Reporting-Fragen (Faktenfragen)
- b) Rating-Fragen (Bewertungsfragen)
- b) Fragen zur Einrichtung
- b) Fragen zum Beschwerdemanagement
- b) Globalfragen.

Zusätzlich abgefragt werden

- Ereignisse zwischen der Entlassung aus der Rehabilitation und dem zweiten Befragungstermin mit Einfluss auf das Rehabilitationsergebnis (mögliche Confounder wie z.B. Unfälle)
- Fragen zum Ausfüllen des Fragebogens.

¹¹ Wegfall in der Routine

Die individuelle Auswahl an Fragen für die Nachbefragung (t2-Messung) zur Bestimmung der Ergebnisqualität erfolgt beim Patientennachbefragungsbogen zwischen den Fragen zum Beschwerdemanagement und den Globalfragen.

Die Bezugspersonenversion beginnt mit der Abfrage der ausgewählten RZK zur Bestimmung der Ergebnisqualität. Der folgen die Frage nach dem Zielerreichungsgrad des individuellen Rehabilitations-Ziels (Globalfrage), den zwischen Rehabilitations-Ende und t2-Messung aufgetretenen Confounder-Ereignissen sowie abschließend Fragen zur Person.

Die Items wurden zur leichteren Bearbeitung für die Patienten und die Bezugspersonen zu Blöcken zusammengefasst, wobei die Überschriften für die Blöcke so gewählt wurden, dass sie keinen inhaltlichen Rahmen bezeichnen, um einen Halo-Effekt¹² zu vermeiden.

Im Patientennachbefragungsbogen werden die Reporting-Fragen dichotom abgefragt, die ja-nein-Antworten erlauben. Rating-Fragen werden 3-stufig mit den Antwortmöglichkeiten „nein“, „teils, teils“ und „ja“ abgefragt. Beide Antwortformate liefern Daten auf Intervallskalenniveau. In wenigen Fällen konnte „betraf mich nicht“ oder in einem Fall „es gab keine Änderungen“ geantwortet werden¹³.

Im Bezugspersonennachbefragungsbogen wurden zum Beantworten der Fragen 5-stufige verbal gelabelte Antwortskalen angeboten, die Daten auf Intervallskalenniveau liefern (Rohrman, 1978; aus Bortz, 1984). Bei den Reporting-Fragen lauteten sie: „trifft gar nicht zu“, „trifft weniger zu“, „teils-teils“, „trifft ziemlich zu“, „trifft völlig zu“¹⁴. Bei den Rating-Fragen lauteten die Label: „nie“, „selten“, „gelegentlich“, „oft“, „immer“¹⁵.

Die Anzahl der unterschiedlichen Antwortformate in den einzelnen Fragebögen wurde bewusst klein gehalten, um die Schwierigkeit beim Ausfüllen durch Formatwechsel nicht unnötig zu erhöhen. Ebenso wurde die Anzahl der vorgelegten Antwortstufen dem Differenzierungsvermögen der Befragten angepasst. Insgesamt wurde darauf geachtet, die Antwortformate so zu gestalten, dass die Angaben äquidistante Stufen bilden. Die damit erhobenen Daten weisen Intervallskalenniveau auf, was für die anstehenden statistischen Auswertungen zumeist von Vorteil und in einigen Fällen als Voraussetzung erforderlich ist.

Auf dieser Grundlage wurden die Erstentwürfe der Nachbefragungsbögen (N-Bögen) erstellt, die bis zur ersten Anwendung im Pretest ausführlich im Steuerungskreis diskutiert und ohne eine Änderung der Grundkonzeption mehrfach überarbeitet wurden.

Die Reihenfolge der RZK-Abfrage kann von der Systematik in Tabelle 2 insofern abweichen, als auf Anraten aus den Kliniken die Abfrage der RZK zu den Funktionen des Verdauungstraktes (1.3 Defäkationsfunktion und 1.4 Miktionsfunktion) stets am Ende der RZK-Abfrage steht.

4.2 Behandlerbogen

Mit dem Behandlerbogen wird im QS-Reha[®]-Verfahren jeder Fall dokumentiert, der an der Befragung teilnimmt bzw. nicht teilnimmt. Dabei werden pro Einrichtung mindestens 50 vollständig auswertbare Fälle erwartet. Von den Patienten, die nicht an der Befragung teilnehmen wollen oder können und für die auch keine Bezugsperson für die Teilnahme an der Befragung gewonnen werden konnte, sind maximal 30 Fälle teilweise zu dokumentieren, um auf dieser Basis eine Drop-Out-Analyse erstellen zu können. Die Analyse soll abschätzen, ob systematische Verzerrungen in den Einrichtungsergebnissen vorliegen können.

Mit dem Behandlerbogen werden Patientenmerkmale erfasst, die einen Einfluss auf die Messergebnisse haben können. Solche Patientenmerkmale werden Confounder genannt. Confounder, die von der behandelnden Einrichtung nicht beeinflusst werden können (z.B. Alter und Geschlecht), können relevant für die Risikoadjustierung der Einrichtungsergebnisse sein und in mathematische Risikoadjustierungsmodelle einfließen. Der für geriatrische Rehabilitationseinrichtungen zu entwickelnde Behandlerbogen soll in der Routine ebenfalls diese Funktionen übernehmen. In dem

¹² Die Überschrift „Abfrage Ihrer Zufriedenheit mit dem Klinikpersonal“ würde eine Globalbewertung beim Befragten erzeugen, auf deren Grundlage er dann die nachfolgenden Items beantworten würde.

¹³ Diese Antwortmöglichkeit gab es bei der Abfrage, ob über Änderungen im Therapieplan informiert wurde.

¹⁴ Beispiel-Item: Ich bin vom Klinikpersonal informiert worden, wer die zuständigen Ansprechpartner auf der Station sind.

¹⁵ Beispiel-Item: Ich habe die Erläuterungen des Klinikpersonals verstanden.

Behandlerbogen für geriatrische Rehabilitationseinrichtungen werden die in den Interviews identifizierten Confounder aufgenommen.

Der Behandlerbogen gliedert sich in vier Teile:

Teil A1 - Teilnahme

In diesem Teil erfolgt die Basisdokumentation der Befragung: wer nimmt teil, ggf. Drop-Out-Grund, Alter und Geschlecht des Patienten, Aufnahmedatum, Bewilligungsdauer und Zugangsweg.

Teil A2 - Rehabilitations-Beginn

In diesem Abschnitt werden die Basisdaten des Patienten erfasst: Vorgeschichte, Rehabilitationsbegründende und weitere Rehabilitations-relevante Diagnosen, Medikamente, Pflegestufe/Pflegegrad, Sturzanamnese, vorausgegangene Krankenhausaufenthalte, Sozialanamnese sowie die Dokumentation der Basis-Assessments bei Aufnahme zu den Bereichen: Sehen, Hören, Schlucken, Kognition (bei Verdacht auf kognitive Probleme Angabe der fünf MMSE-Subskalen-Scores¹⁶), Delir, psychisches Befinden, Ernährungszustand, Körpergewicht, Körpergröße, Dekubitus, alle 10 Scores des motorischen Barthel-Index(Mahoney et al., 1965) sowie die behandlerseitige Einschätzung der Rehabilitations-Motivation.

Teil A3 - Rehabilitations-Ziel

Im dritten Teil wird das mit dem Patienten vereinbarte individuelle Rehabilitations-Ziel als Fließtext eingegeben (Erläuterungen zu dessen Formulierung s. Kapitel 4.4), die Auswahl der RZK dokumentiert und ggf. weitere Rehabilitations-Ziele angegeben. Bei der Online-Erfassung des Behandlerbogens wird die Erstellung der für diesen Fall erforderlichen individualisierten Fragebögen durch Abschluss des A3-Bogens gestartet. Gleichzeitig wird dieser Teil gegen nachträgliche Änderungen gesperrt.

Die bezogen auf jeden Patienten individuellen Fragebögen sind eine wesentliche Neuerung im QS-Reha®-Verfahren. Diese trägt den besonders vielfältigen Krankheitskonstellationen (Multimorbidität) in der geriatrischen Rehabilitation Rechnung.

Teil A4 - Rehabilitations-Ende

Im vierten Teil werden der Rehabilitations-Verlauf (Verlängerung, unerwünschte Ereignisse), behandlerseitige Einschätzungen des Therapieerfolges sowie Angaben zur Entlassung erfasst.

In der Routine werden im Drop-Out-Fall die Teile A1 und A2 vollständig zu dokumentieren sein. Im Pretest wurde darauf verzichtet, weil es hierbei vornehmlich um die Daten der teilnehmenden Patienten ging, um die Instrumente hinsichtlich ihrer Messeigenschaften empirisch zu überprüfen.

4.3 Einwilligungserklärung

Ein Patient bzw. eine Bezugsperson kann nur dann an der Befragung teilnehmen, wenn eine von ihm bzw. von ihr unterschriebene Einwilligungserklärung vorliegt. Diese Vorgabe bestand bei der Entwicklung des Verfahrens und besteht ebenso für seine Anwendung in der Routine¹⁷.

Die Einwilligungserklärung besteht aus der Information über den Hintergrund der Befragung (Seite 1) und der Einwilligungserklärung selbst (Seite 2).

Im Informationsteil wird ausgeführt:

- Das Ziel der Befragung ist die Verbesserung der medizinischen Versorgung aus Sicht der Patienten.
- Das BQS-Institut ist mit der Durchführung der Befragung beauftragt.
- Die Teilnahme umfasst zwei Befragungen, eine in der Klinik und eine weitere zwei Wochen nach Entlassung aus der Klinik.
- Die Teilnahme an der Befragung ist freiwillig und ein Widerruf ohne Angabe von Gründen jederzeit möglich.
- Der Datenschutz wird vollumfänglich gewährleistet, um die Anonymität des Teilnehmenden zu wahren.

¹⁶ MMSE: Mini-Mental-Status-Test, ist ein für den klinischen Alltag geeignetes Screening-Verfahren zur Feststellung kognitiver Defizite nach Folstein et al. (1975)

¹⁷ Es besteht der gesetzlich vorgegebene Datenschutz, inklusive EU-DSGVO ab 25.05.2018.

In der Einwilligungserklärung wird ausgeführt:

- Die Aufklärung über Inhalt, Vorgehensweise und Ziele der Qualitätssicherungsmaßnahme ist durch das Klinikpersonal erfolgt.
- Es gab Gelegenheit und ausreichend Zeit Fragen zu stellen, die auch beantwortet wurden.
- Eine Kopie der Einwilligungserklärung wurde ausgehändigt.
- Die Speicherung, Auswertung, Weitergabe der erhobenen Daten erfolgt nach den gesetzlichen Bestimmungen.

Patienten bzw. Bezugspersonen erhalten das Formular in 2-facher Ausführung, ein Blatt zum Verbleib beim Patienten bzw. bei der Bezugsperson, das andere Blatt wird unterschrieben in der Patientenakte archiviert.

Die Einwilligungserklärung liegt in zwei Versionen vor, in einer Version für Patienten, Angehörige, Freunde und Bekannte des Patienten und in einer Version für gesetzlich bestellte Betreuer (s. Anlage 7.1 und 7.2). Die Betreuer-Version wurde erforderlich, um den Text sprachlich der Betreuungssituation anzupassen und um den Betreuerstatus explizit bestätigt zu bekommen.

4.4 Handlungsanweisung für die Rehabilitations-Zielvereinbarung sowie die Auswahl und Dokumentation der Rehabilitations-Zielkategorien

Das mit dem Patienten/der Patientin und den Bezugspersonen individuell vereinbarte Rehabilitations-Ziel dient im Verfahren als Grundlage für die Messung der Ergebnisqualität mit einem individualisierten Patientenfragebogen. Jeder Patient/jede Patientin erhält dadurch nur Fragen zu Ergebnis-Aspekten, die für ihn oder sie persönlich während der Rehabilitation relevant waren. Dazu sollen aus dem individuell vereinbarten Rehabilitations-Ziel die relevanten Rehabilitations-Zielkategorien abgeleitet werden.

Ziel der Handlungsanweisung ist es, eine einheitliche Vorgehensweise bei der Rehabilitations-Zielvereinbarung und der Ableitung der Rehabilitations-Zielkategorien in den Einrichtungen zu fördern, um ein standardisiertes Vorgehen im QS-Verfahren zu ermöglichen. Die Handlungsanweisung hat den Rang einer Empfehlung. Die Handlungsanweisung für die Einrichtungen ist Anlage 8 zu entnehmen.

Das vereinbarte individuelle Rehabilitations-Ziel wird einrichtungsseitig im Behandlerbogen dokumentiert (s. Teil A3, Frage 34). Das individuelle Rehabilitations-Ziel soll vom Wortlaut und vom Inhalt her so formuliert werden, wie es der Patient als angestrebte Aktivität innerhalb seiner Lebenswelt beschreibt.

Ein Beispiel für ein individuelles Rehabilitations-Ziel könnte lauten: Der Patient möchte seinen Hund wieder täglich ausführen können. Hier als vereinbartes individuelles Rehabilitations-Ziel zu dokumentieren: Der Patient möchte wieder 200 m am Stück laufen können und wieder über ausreichend Handkraft verfügen, wäre unpassend. Es wäre die Übersetzung des individuellen Rehabilitations-Ziels in therapeutisch gängige Größen ohne Bezug zur konkreten angestrebten Aktivität des Patienten.

Die dokumentierte Formulierung wird dem Patienten bzw. entsprechend der Bezugsperson dann bei den Befragungen an zwei Stellen vorgelegt. Bei der Eingangsbefragung wird dem Patienten bzw. der Bezugsperson im personalisierten Fragebogen die im Behandlerbogen dokumentierte Formulierung vorgelegt, verbunden mit den Fragen, ob das individuelle Rehabilitations-Ziel korrekt wiedergegeben sei, wie wichtig es ihm sei und ob und ggf. welche anderen Ziele er, der Patient, in der Rehabilitation noch erreichen wollte. Am Ende der Nachbefragung wird diese Formulierung wieder vorgelegt mit der Bitte anzugeben, in wie weit er, der Patient, dieses Rehabilitations-Ziel erreicht hat. Damit der Patient bzw. die Bezugsperson diese Fragen sinngemäß beantworten kann, soll die dokumentierte Formulierung des individuellen Rehabilitations-Ziels mit dem übereinstimmen, was mit dem Patienten tatsächlich vereinbart wurde. Wird nach Einschätzung der Befragten das Rehabilitations-Ziel nicht korrekt wiedergegeben und wird es zudem als „kaum“ oder „gar nicht“ wichtig eingestuft, dann sollte dieser Fall nicht in die Auswertung der Ergebnisqualität einbezogen werden. Außerdem sollte die Anzahl solcher Fälle für jede Einrichtung vermerkt werden.

4.5 Kognitive Pretests

Nachdem die Fragebögen im Steuerungskreis konsentiert waren, erfolgte die Überprüfung der Instrumente auf Verständlichkeit für Patienten und Bezugspersonen mit kognitiven Pretests.

In sechs geriatrischen Einrichtungen wurden kognitive Pretests mit insgesamt 38 geriatrischen Patienten durchgeführt (s. Tabelle 3). Sofern zufällig anwesend und willens wurden auch sechs

Bezugspersonen getestet. In der Regel wurden dazu von den Ärzten Patienten auf der Station vorgeschlagen. Teilweise wurden seitens des BQS-Instituts auch vorab oder vor Ort Wünsche geäußert, welche Einschränkungen abgefragt werden sollten, um zu möglichst jeder Rehabilitations-Zielkategorie Erfahrungen mit der Beantwortung der jeweiligen Erfassungsvariante sammeln zu können. Im Vorfeld hatten sich die Einrichtungen die Einwilligungserklärungen unterschreiben lassen.

Die kognitiven Pretests liefen zumeist in der Form ab, dass uns ein Arzt zunächst über das Krankheitsbild des Patienten und den Verlauf der Therapie informierte, um uns anschließend dem Patienten vorzustellen und ihm zu erklären, was wir vorhaben. Die kognitiven Pretests wurden dann in der Regel von zwei BQS-Mitarbeitern durchgeführt (die letzten 7 Tests in Meerbusch nur von einem Mitarbeiter). Die Patienten wurden gebeten, beim Ausfüllen der Fragebögen alles, was ihnen dabei in den Sinn kommt, laut auszusprechen (vgl. Think Aloud Technik, Häder 2006), um nachvollziehen zu können, ob der Patient die Aufgabe wie beabsichtigt korrekt versteht oder ob sich Fehler im Verständnis oder in der praktischen Umsetzung zeigen.

Auf Grundlage der dabei gemachten Beobachtungen wurden die Fragebögen überarbeitet. Dies betraf die Item-Formulierungen, die räumliche Anordnung der Fragen, die Antwortkategorien und das Layout z.B. der Antwortkästchen. Die überarbeiteten Fragebögen wurden dann mit Patienten der nächsten Einrichtung überprüft. Auf diese Weise wurden die Fragebögen von Einrichtung zu Einrichtung optimiert, um am Ende pretesttaugliche Versionen zu erhalten. Rückblickend lässt sich feststellen, dass die kognitiven Pretests absolut erforderlich waren, da die geriatrischen Patienten mit vielen Formulierungen Verständnisprobleme hatten. Außerdem zeigte sich, dass sie ein extrem eingeschränktes Gesichtsfeld hatten, was enge Vorgaben für das Layout bedeutete. Diese aufwändigen Vorarbeiten waren eine Voraussetzung dafür, dass der Pretest zuverlässige Ergebnisse geliefert hat.

Tabelle 3: Zur Exploration der Prozesse und zur Durchführung von kognitiven Pretests besuchte geriatrische Rehabilitationseinrichtungen

	Einrichtung	Ort	Ansprechpartner	Zeitraum	Anzahl Patienten
1.	Robert-Bosch-Krankenhaus	Stuttgart	CA Prof. Dr. Becker	28. und 29.09.2015	4
2.	Geriatrische Fachklinik Rheinhessen-Nahe	Bad Münster am See	ÄD Dr. Heckmann	27. und 28.10.2015	3
3.	Weserland-Klinik Bad Hopfenberg	Petershagen	CA Dr. Voigt	18. und 19.11.2015	5
4.	Klinikum Coburg	Coburg	CA Prof. Dr. Kraft	11. bis 13.01.2016	8
5.	Petrus-Krankenhaus	Wuppertal	LOA Dr. Nieding-Schoppen	27.01.2016	6
6.	St. Mauritius Therapieklinik	Meerbusch	CA Dr. Woelki	02. und 09.02.2016	12

5 Pretest

Zur empirischen Überprüfung der Messeigenschaften der Erstversionen der Fragebögen wurde ein Pretest durchgeführt. Auf Grundlage der dadurch gewonnenen Ergebnisse wurden die Instrumente zur Steigerung der Güte der Messungen revidiert. In der anschließenden Pilotphase wurden die Messeigenschaften der revidierten Instrumente an einer zweiten unabhängigen Stichprobe erneut empirisch überprüft, um die Stabilität der Ergebnisse oder Verbesserung feststellen zu können. Zum anderen hatte die Pilotphase zum Ziel, die Abläufe im Qualitätssicherungsverfahren, wie sie für die Routine geplant sind, zu überprüfen.

Der Pretest wurde dazu durchgeführt, die entwickelten Instrumente empirisch zu überprüfen und ihre praktische Tauglichkeit zu testen. Auf Grundlage der erhobenen Daten sollten insbesondere deren Messeigenschaften durch Betrachtung diverser Item-Kennwerte sowie der Reliabilität und Validität der erfassten Konstrukte bestimmt werden. Zudem sollten die Befunde Hinweise zur Überarbeitung der Instrumente liefern.

5.1 Ziele

Im Vordergrund des Pretests stand die empirische Überprüfung der Messeigenschaften der Erstversionen der Fragebögen. Diese wurden zunächst durch die Vollständigkeit und Vollständigkeit der ausgefüllten Fragebögen und der dort gemachten Angaben sowie durch die Item-Kennwerte „Item-Schwierigkeit“ und „Trennschärfe“ beschrieben.

Der Pretest diente weiterhin dazu, die Faktorenstruktur der Items zur Messung der Prozessqualität und der Zufriedenheit festzustellen. Die dadurch gefundenen Faktoren wurden als vorläufige Qualitätsunterdimensionen betrachtet. Für diese wurde die interne Konsistenz als Reliabilitätsmaß bestimmt.

Einzelne Rehabilitations-Zielkategorien, für die genügend Fälle vorlagen, wurden zusätzlich hinsichtlich ihrer Kriteriumsvalidität geprüft¹⁸. Die Prüfung der Kriteriumsvalidität der restlichen Rehabilitations-Zielkategorien bleibt nachfolgenden empirischen Erhebungen wie der Pilotphase und der Routine vorbehalten.

Ein besonderes Anliegen des Pretests bestand darin zu überprüfen, ob Patienten und Bezugspersonen unabhängig voneinander zu übereinstimmenden Einschätzungen der Leistungspotenz der betreffenden Patienten kommen. Ziel war empirisch abzusichern, dass die gewählte Konzeption der Ergebnisqualitätsmessung dazu führt, dass Patienten- und Bezugspersoneneinschätzungen soweit übereinstimmen, dass bei Nicht-Teilnahme der Patienten die entsprechenden Angaben der Bezugspersonen als adäquater Ersatz verwendet werden können.

Zudem wurden einzelne Items, die bestimmte Aspekte der Prozessqualität, der Zufriedenheit sowie Globalbewertungen bei den Patienten erfassen, mit Parallelversionen der Items in der Bezugspersonennachbefragung abgefragt. Dem Patienten wurde z.B. vorgelegt: „Ich wurde vom Klinikpersonal motiviert, meine Rehabilitations-Ziele aktiv zu verfolgen.“ Das parallele Item für die Bezugspersonen lautete: „Die Behandler haben meinen Angehörigen motiviert, seine Rehabilitations-Ziele aktiv zu verfolgen.“

Der Pretest sollte Aufschluss darüber geben, inwieweit diese Items, 14 insgesamt, von Patienten und Bezugspersonen vergleichbar beantwortet werden, um bei Nicht-Teilnahme der Patienten durch Angaben der Bezugspersonen ersetzt werden zu können.

5.2 Methode

Der Pretest wurde als Parallelerhebung angelegt, um die Vergleichbarkeit der Patienten- und Bezugspersoneneinschätzungen bei der Ergebnisqualitätsmessung empirisch zu prüfen. Daher sollte in einem definierten Zeitfenster möglichst zu jedem teilnehmenden Patienten eine Bezugsperson akquiriert werden. Drop-Out-Fälle wurden nicht einbezogen und entsprechend keine Drop-Out-Behandlerbögen angelegt, wie sonst im QS-Reha[®]-Verfahren üblich, da der Pretest vornehmlich zur empirischen Überprüfung der Instrumente durchgeführt wurde, und nur die Daten der ausgefüllten Fragebögen der teilnehmenden Patienten und Bezugspersonen interessierten.

Das Erhebungsdesign unterscheidet als Datenquellen: Behandler, Patienten und Bezugspersonen, die Messzeitpunkte t0 (Rehabilitations-Beginn), t1 (Rehabilitations-Ende) und t2 (zwei Wochen nach Rehabilitations-Ende) sowie die zu erhebenden Qualitätsdimensionen: Ergebnisqualität (EQ), Prozessqualität (PQ) und Patienten- und Bezugspersonenzufriedenheit (Z) sowie diverse Patientenmerkmale (s. Tabelle 4).

Tabelle 4: Erhebungsdesign des Pretests

¹⁸ RZK Schmerzen, Körperposition wechseln, Gangsicherheit, Gehen im Innenbereich, Treppen steigen, Gehen im Außenbereich, Körperpflege

	t0	t1	t2
Patient	EQ	jeweils identische RZK-Auswahl abgefragt	EQ, PQ, PZ
Bezugs- person	EQ		EQ, PQ, PZ
Behand- ler	Patienten- merkmale	Reha-Verlauf	

5.3 Ablauf

5.3.1 Einrichtungsakquise

Für die Einrichtungsakquise wurden alle ambulanten und stationären Geriatrie-Einrichtungen, die im QS-Reha[®]-Register hinterlegt waren, durch einen BQS-Mitarbeiter per E-Mail und telefonisch kontaktiert. Insgesamt wurden auf diesem Wege ca. 200 geriatrische Rehabilitationseinrichtungen über das Verfahren informiert. Parallel dazu hat der Bundesverband Geriatrie e.V. (BVG) eine eigene Akquise durchgeführt. Nach finaler Rücksprache mit den interessierten Einrichtungen wurden in der Summe 17 Teilnehmer aus stationären und ambulanten Einrichtungen bundesweit für den Pretest akquiriert.

5.3.2 Schulung der Teilnehmer

Die Schulungen der zentralen Koordinatoren der Rehabilitationseinrichtungen für den Pretest fanden im Zeitraum vom 01.06.2016 bis zum 06.07.2016 statt. Am 01.06.2016 und am 15.06.2016 wurden Präsenzs Schulungen im Haus des GKV-Spitzenverbandes in Berlin durchgeführt. Da nicht alle Einrichtungen einen dieser Termine wahrnehmen konnten, wurden zudem zwei Einrichtungen am 06.07.2016 telefonisch geschult (s. Tabelle 5).

Tabelle 5: Durchgeführte Schulungen

Datum	Ort	Anzahl Teilnehmer
01.06.2016	GKV-SV, Berlin	5
15.06.2016	GKV-SV, Berlin	13
06.07.2016	telefonisch	2
Summe		20

Im Rahmen der Schulungen wurde im Allgemeinen auf das Konzept des Entwicklungsprojektes Geriatrie eingegangen und die Bedeutung des Pretests und der Pilotphase dargestellt. Im Zuge dessen wurden die Hintergründe des Entwicklungsprojekts aufgezeigt und die Besonderheiten in der Geriatrie beleuchtet. Der Fokus der Schulungen lag insbesondere bei der Erläuterung des individuellen Rehabilitations-Ziels und den Rehabilitations-Zielkategorien. Um den Einrichtungen die praktische Durchführung zu erleichtern, wurde darüber hinaus über die technische Realisierung und Organisation in der Einrichtung informiert.

Den Schulungsteilnehmern wurden jeweils folgende Schulungsunterlagen zur Verfügung gestellt:

1. Schulungsprogramm
2. Vortragsfolien
 - 2.1. Konzept des Entwicklungsprojekts Geriatrie und Konzept des Pretests und der Pilotphase

- 2.2. Das individuelle Rehabilitations-Ziel und die Rehabilitations-Zielkategorien
- 2.3. Technische Realisierung
- 2.4. Organisation in der Einrichtung
3. Überblick über das Entwicklungsprojekt Geriatrie
4. Handlungsanweisung Rehabilitations-Zielvereinbarung
5. Systematik der Rehabilitations-Zielkategorien
6. Auszug der Rehabilitations-Zielkategorien - Operationalisierung
7. Muster-Fragebögen

Diese Dokumente wurden den Schulungsteilnehmern auch im Extranet als Download zur Verfügung gestellt.

Die Pretest-Schulungen wurden mittels Fragebögen evaluiert (s. Anlage 9). An der Evaluation der Schulungen haben sich insgesamt 11 Schulungsteilnehmer beteiligt. Insgesamt sind die Schulungen von den Teilnehmern recht positiv bewertet worden. Auf einer Skala von 1 (sehr gut) bis 6 (völlig unzureichend) lag der Mittelwert des Gesamteindrucks bei 1,45. Das Programm der Schulung entsprach den Vorstellungen von 73% der Schulungsteilnehmer. 27% der Teilnehmer hatten hingegen zuvor keine konkrete Vorstellung, jedoch im Nachhinein einen guten Eindruck durch das Schulungsprogramm gewonnen.

5.3.3 Durchführung des Pretests

Die Patientenakquise erfolgte in den teilnehmenden Rehabilitations-Einrichtungen vom 18.07. bis zum 16.09.2016.

Die Behandler füllten für jeden akquirierten Fall einen Behandlerbogen aus, mit dem potentielle Confounder für die Risikoadjustierung erhoben wurden. Dazu wurden Personenmerkmale sowie Angaben zum Verlauf der Rehabilitation zu t0 und zu t1 erfasst. Der Behandlerbogen liegt als Online-Version vor. Durch die Eingabe des individuellen Rehabilitations-Ziels und der ausgewählten Rehabilitations-Zielkategorien in Teil A3 des Behandlerbogens sowie dem definitiven Abschluss dieser Eingaben werden die für diesen Fall individuell zusammengestellten Fragebögen für die t0- und die t2-Befragung des Patienten und der Bezugsperson als PDF bereitgestellt.

Patienten und Bezugspersonen wurden zu t2 spezifisch zur Prozessqualität und zur Patienten- bzw. Bezugspersonenzufriedenheit befragt. Für die Messung der Ergebnisqualität wurde die patientenspezifische und damit individuelle Auswahl an RZK den Patienten und den Bezugspersonen zu t0 und zu t2 vorgelegt. Die patientenspezifische RZK-Auswahl wird also zur Messung der Ergebnisqualität an vier Stellen im Erhebungsdesign identisch vorgelegt. Mit der Parallelerhebung von Patienten und Bezugspersonen unterscheidet sich das Erhebungsdesign des Pretests vom Routineverfahren. Bezugspersonen sollen im Routineverfahren nicht mehr parallel, sondern alternativ bei Bedarf und Zusage befragt werden.

Die in den Einrichtungen gesammelten ausgefüllten Fragebögen wurden an zwei Lieferterminen (16.09. und 28.10.2016) an das BQS-Institut in Hamburg zum Erfassen der Angaben gesendet. Anders als in der Routine wurden die t2-Bögen an die Einrichtung zurückgesendet, die dort verschlossen für den Versand an das BQS-Institut gesammelt wurden.

Die Angaben der maschinenlesbaren Fragebögen wurden elektronisch erfasst und in die Datenbank des BQS-Instituts eingelesen. Zur Auswertung der Daten wurde diese mit denen der Behandlerbögen zusammengeführt und bereitgestellt, so dass für jeden Fall alle Angaben der sechs eingesetzten Instrumente¹⁹ zur Auswertung vorlagen.

5.4 Ergebnisse

5.4.1 Stichprobe

Zur Auswertung kamen 629 vollständig dokumentierte Behandlerbögen aus 13 Einrichtungen mit 618 Patienten und 239 Bezugspersonen. Bei den Patienten lag die Rücklaufquote (RQ) zum Messzeitpunkt t2 insgesamt bei 41,3%, bei den Bezugspersonen insgesamt bei 30,1%. Für 60 Fälle lagen Patienten- und Bezugspersonenangaben für t0 und für t2 vor (RQ=26,3%).

¹⁹ Behandlerbogen, Patienten- und Bezugspersonenbogen zu Rehabilitations-Beginn, Patientenbogen ambulant und stationär sowie Bezugspersonenbogen Nachbefragung

In der Stichprobe gab es nur 10 ambulante Fälle mit Angaben zu t0 und zu t2, so dass wegen der geringen Aussagekraft der auf dieser Grundlage ermittelten Ergebnisse für diese Patientengruppe keine statistischen Auswertungen durchgeführt werden konnten.

Die folgenden statistischen Auswertungen beziehen sich auf die Angaben aus dem Patientenachbefragungsbogen stationär (PNS) und dem Bezugspersonennachbefragungsbogen (BN), und zwar zunächst bezogen auf die Items zur Messung der Prozessqualität und der Patientenzufriedenheit. Die Messung der Ergebnisqualität erfolgt auf Grundlage eines methodisch aufwändigen Algorithmus, so dass die statistische Überprüfung der Messeigenschaften getrennt verläuft (s. Kapitel 5.4.6 ff).

Vollzähligkeit und Vollständigkeit der Befragungsdaten

Bei der Befragung der stationär behandelten Patienten gab es 600 Beginn-Bögen und 245 Nachbefragungsbögen, was einer Rücklaufquote von 40,8% entspricht.

Die statistische Analyse der Befragungsdaten zeigt, dass der Anteil fehlender Werte überwiegend unter 5% liegt. Zudem zeigt sich kein Ermüdungseffekt, also keine Zunahme fehlender Werte zum Fragebogenende hin (s. Anlage 10).

Im Bezugspersonennachbefragungsbogen wurde bei einer Reihe von Items zusätzlich zu den inhaltlichen Antwortstufen die Antwortkategorie „betrifft mich nicht“ und „weiß ich nicht“ aufgenommen. Diese Antwortkategorie wurde zwischen 20% bis 49% (Item 33: „Mit mir vereinbarte Termine fanden eine halbe Stunde und mehr verspätet statt.“) angekreuzt. Diese Antworten besagen, dass der Befragte häufig nicht antworten konnte, weil er den thematisierten Vorgang nicht erlebt hat. Der Anteil verbleibender tatsächlich fehlender Antworten war bei diesen Items ebenfalls zumeist unter 5%.

Es wurde kein Item auf Grund eines hohen Anteils fehlender Werte gestrichen oder umformuliert.

5.4.2 Item-Analyse

Die *Item-Schwierigkeit*²⁰ lag bei zahlreichen Items unter 0,2 bzw. über 0,8, was eine eingeschränkte Differenzierungsfähigkeit bedeutet. Auf Grund extremer Item-Schwierigkeiten wurde jedoch kein Item gestrichen oder umformuliert.

Die *Trennschärfe* der Items war überwiegend hoch ($TS \geq 0,6$). Geringe Trennschärfen wurden damit erklärt, dass die Item-Blöcke Items unterschiedliche Faktoren enthalten und damit inhomogen sind.

Mit der *explorativen Faktorenanalyse* wird die Item-Gruppierung überprüft, die dann auch für die Score²¹-Bildung relevant ist. Die Faktorenanalysen wurden differenziert nach Patienten und Bezugspersonen und getrennt für Reporting- und Rating-Fragen (Prozessqualität und Zufriedenheit) sowie weiter getrennt für spezifische und globale Fragen gerechnet.

Die durchgeführten Faktorenanalysen liefern insgesamt gut interpretierbare Faktoren, hohe Varianzaufklärungen (VA) sowie hohe Kaiser-Meyer-Olkin-Kriterium (KMO) Werte bei häufig hohen Fallzahlen (N)²². Die Varianzaufklärung der einbezogenen Items gibt an, inwieweit durch die Faktoren die gesamte Varianz der beobachteten Werte erklärt wird. Ein Faktor muss daher genügend Varianz erklären, um seine weitere Verwendung zu rechtfertigen. Das KMO gibt an, ob ein Datensatz für eine Faktorenanalyse geeignet ist und kann Werte zwischen null und eins annehmen. Wünschenswert ist ein KMO-Wert von $> 0,8$.

Zusätzlich zur Bezeichnung der extrahierten Faktoren wird die interne Konsistenz der Faktor-Items mit Cronbachs Alpha (α)²³ angegeben, das als Maß für die Reliabilität der aus den Faktor-Items

²⁰ Item-Schwierigkeit: Mittelwert aller Antworten eines Items, bei dem die unterste Antwortkategorie den Wert 0 und die oberste den Wert 1 hat. Items mit extrem niedriger bzw. extrem hoher Item-Schwierigkeit ($IS < 0,2$ bzw. $IS > 0,8$) sind für Messinstrumente ungeeignet.

²¹ Score: Damit wird der Messwert bezeichnet, der aus den Messergebnissen einer definierten Gruppe von Items gebildet wird.

²² Bis auf in den Faktorenanalysen der Reporting- und der Rating-Fragen der Bezugspersonen waren weitaus mehr als die doppelte Anzahl von Fällen als Items in den Faktorlösungen zu finden.

²³ Cronbachs Alpha entspricht der mittleren Testhalbierungs-Reliabilität eines Tests (hier den Items eines Faktors) für alle möglichen Halbierungen (vgl. Bortz & Döring, 2006).

gebildeten Scores zu betrachten ist. Alpha-Werte zwischen 0 und 1 sind interpretierbar, wobei als Faustregel $\alpha=0,65$ oder größer sein sollte. In einzelnen Fällen lässt sich die Reliabilität solcher Item-Gruppen durch das Streichen inhomogener bzw. das Hinzunehmen homogener Items vergrößern. In wenigen Fällen ist Cronbachs Alpha kleiner als $\alpha=0,65^{24}$, was besagt, dass die damit errechneten Scores das zu messende Konstrukt mit einer höheren Messgenauigkeit beschreiben.

Die Ergebnisse der einzelnen Faktorenanalysen werden nachfolgend dargestellt (Details s. Anlage 11):

Patientennachbefragungsbogen stationär (PNS)

1. Faktorenanalyse aller Reporting-Fragen (Prozessqualität)

Für die Reporting-Fragen zur Prozessqualität wurde eine 5-Faktoren-Lösung ermittelt, die Faktoren waren inhaltlich gut interpretierbar. Die Varianzaufklärung (VA) betrug 68,8%, das Kaiser-Meyer-Olkin-Kriterium (KMO) betrug 0,837 bei einer Fallzahl N=80. Die ermittelten Faktoren sind:

Faktor 1: Medizinische Informationen mit Patient besprechen	($\alpha=0,83$)
Faktor 2: Informationen für die Zeit nach der Rehabilitation geben	($\alpha=0,76$)
Faktor 3: Rehabilitations-Grundlagen erfragen und vermitteln	($\alpha=0,67$)
Faktor 4: Allgemeine Informationen zu Rehabilitations-Beginn geben	($\alpha=0,68^*$)
Faktor 5: Therapieplanänderungen	($\alpha=0,42$)

2. Faktorenanalyse der spezifischen Rating-Fragen (Patientenzufriedenheit), ausgenommen Hotel-Fragen

Für die spezifischen Rating-Fragen zur Patientenzufriedenheit (ausgenommen Hotel-Fragen) wurde eine 6-Faktoren-Lösung ermittelt, die Faktoren waren inhaltlich gut interpretierbar. In diesem Fall betrug die Varianzaufklärung VA=75,2%, das Kaiser-Meyer-Olkin-Kriterium (KMO) betrug 0,817 bei einer Fallzahl N=76. Die ermittelten Faktoren sind:

Faktor 1: Respekt des Klinikpersonals	($\alpha=0,82^*$)
Faktor 2: Zuverlässigkeit / Zugewandtheit des Klinikpersonals	($\alpha=0,80$)
Faktor 3: Aktivitäten der Therapeuten	($\alpha=0,76^*$)
Faktor 4: Last / Anstrengung der Therapien	($\alpha=0,64$)
Faktor 5: Diskretion des Klinikpersonals	($\alpha=0,67^{\ddagger}$)
Faktor 6: Verständlichkeit des Klinikpersonals	($\alpha=^{**}$)

3. Faktorenanalyse der Hotel-Rating-Fragen (Patientenzufriedenheit)

Für die Hotel-Rating-Fragen zur Patientenzufriedenheit wurde eine 1-Faktor-Lösung ermittelt. In diesem Fall betrug die Varianzaufklärung VA=50,7%, das Kaiser-Meyer-Olkin-Kriterium (KMO) betrug 0,745 bei einer Fallzahl N=60. Der ermittelte Faktor ist:

Faktor 1: Hotelleistungen	($\alpha=0,71$)
---------------------------	-------------------

4. Faktorenanalyse der globalen Rating-Fragen (Patientenzufriedenheit, Ergebnisqualität)

Für die globalen Rating-Fragen zur Patientenzufriedenheit und Ergebnisqualität wurde eine 3-Faktoren-Lösung ermittelt, die Faktoren waren inhaltlich gut interpretierbar. In diesem Fall betrug die Varianzaufklärung VA=68,6%, das Kaiser-Meyer-Olkin-Kriterium (KMO) betrug 0,833 bei einer Fallzahl N=147. Die ermittelten Faktoren sind:

Faktor 1: Globalbewertung der Rehabilitations-Einrichtung (PZ)	($\alpha=0,85^*$)
Faktor 2: Rehabilitations-Zielerreichung (EQ)	($\alpha=0,81^*$)
Faktor 3: Bewertung sekundärer Dienstleistungen (Z)	($\alpha=0,49^{\ddagger}$)

Bezugspersonennachbefragungsbogen (BN)

5. Faktorenanalyse der Reporting-Fragen (Prozessqualität)

Für die Reporting-Fragen zur Prozessqualität wurde eine 5-Faktoren-Lösung ermittelt, die Faktoren waren inhaltlich gut interpretierbar. In diesem Fall betrug die Varianzaufklärung VA=81,7%, das

*) maximales Cronbachs Alpha bei Streichung mindestens eines inhomogenen Items

‡) maximales Cronbachs Alpha bei Hinzunahme mindestens eines homogenen Items

***) keine Berechnung möglich, weil nur ein Item vorliegt

²⁴ Cronbachs Alpha ist ein Maß auf Fallebene, Die QS-Ergebnisse sind jedoch auf Einrichtungsebene relevant. Die Reliabilität von QS-Ergebnissen steigt mit zunehmender Fallzahl an.

Kaiser-Meyer-Olkin-Kriterium (KMO)²⁵ betrug 0,803 bei einer Fallzahl N=42. Die ermittelten Faktoren sind:

Faktor 1: Medizinische Informationen zum Rehabilitations-Ende erhalten	($\alpha=0,95$)
Faktor 2: Informationen zum med. Status erklärt bekommen	($\alpha=0,89$)
Faktor 3: Informationen zur Unterstützung des Patienten erhalten	($\alpha=0,91$)
Faktor 4: Informationen über Leistungen nach der Rehabilitation erhalten	($\alpha=0,91$)
Faktor 5: Informationen zur Prävention und Weiterversorgung erhalten	($\alpha=0,92$)

6. Faktorenanalyse der spezifischen Rating-Fragen (Bezugspersonenzufriedenheit)

Für die spezifischen Rating-Fragen zur Bezugspersonenzufriedenheit wurde eine 4-Faktoren-Lösung ermittelt. Die Faktoren waren wegen der kleinen Fallzahl eingeschränkt interpretierbar. In diesem Fall betrug die Varianzaufklärung VA=86,0%, das Kaiser-Meyer-Olkin-Kriterium (KMO) betrug 0,538 bei einer Fallzahl N=19. Die ermittelten Faktoren sind:

Faktor 1: Vertrauenswürdigkeit / Zugewandtheit	($\alpha=0,94$)
Faktor 2: Gesprächsmodalitäten (pünktlich, vertraulich, ernsthaft)	($\alpha=0,86$)
Faktor 3: Respekt / Freundlichkeit	($\alpha=0,83$)
Faktor 4: keine inhaltliche Interpretation möglich	

7. Faktorenanalyse der globalen Rating-Fragen (Bezugspersonenzufriedenheit)

Für die globalen Rating-Fragen zur Bezugspersonenzufriedenheit wurde eine 3-Faktoren-Lösung ermittelt, die Faktoren waren inhaltlich gut interpretierbar. In diesem Fall betrug die Varianzaufklärung VA=72,7%, das Kaiser-Meyer-Olkin-Kriterium (KMO) betrug 0,574 bei einer Fallzahl N=45. Die ermittelten Faktoren sind:

Faktor 1: Hotelleistungen	($\alpha=0,76$)
Faktor 2: Bewertung sekundärer Dienstleistungen	($\alpha=0,54$)
Faktor 3: Transfer nach Hause	($\alpha=**$)

8. Faktorenanalyse der globalen Rating-Fragen (Ergebnisqualität, Bezugspersonenzufriedenheit)

Für die globalen Rating-Fragen zur Ergebnisqualität und Bezugspersonenzufriedenheit wurde eine 2-Faktoren-Lösung ermittelt, die Faktoren waren inhaltlich gut interpretierbar. In diesem Fall betrug die Varianzaufklärung VA=73,8%, das Kaiser-Meyer-Olkin-Kriterium (KMO) betrug 0,726 bei einer Fallzahl N=35. Die ermittelten Faktoren sind:

Faktor 1: Globalbewertung des Rehabilitations-Ergebnisses (EQ-Bez.)	($\alpha=0,82$)
Faktor 2: Globalbewertung der Rehabilitations-Einrichtung (Z-Bez.)	($\alpha=0,90$)

Die gefundenen Faktoren bilden den Ausgangspunkt für die Bezeichnung und Operationalisierung der Prozess- und Zufriedenheits-Scores. Für jeden Faktor wurde einzeln überprüft, welche auf ihm hoch ladenden Items eine maximal reliable Messung des betreffenden Konstrukts erlauben. Dazu wurde für diese Item-Gruppen Cronbachs Alpha (α) bestimmt und es wurden ggf. zur Optimierung der Messgenauigkeit inkonsistente Items sukzessiv herausgenommen. Die so ermittelten optimierten Item-Gruppen definieren dann die verschiedenen Scores, die im QS-Reha[®]-Verfahren als vorläufige Qualitätsunterdimensionen bezeichnet werden.

Zur vertiefenden Betrachtung der Ergebnisse ist in Anlage 11 die Zusammenstellung der Faktorenanalysen mit allen Items und Faktorladungen zu finden.

5.4.3 Item-Revision

Aus inhaltlichen Überlegungen wurden drei Items umformuliert.

Im Patientenbogen Nachbefragung stationär wurden die Items:

- 10. Mir wurde der Umgang mit den verordneten Hilfsmitteln erläutert und gezeigt.
- 15. Ich habe Therapieanleitungen für Zuhause erhalten (z.B. gymnastische Übungen, Ernährungstipps, Trainingsmaßnahmen).
- 51. Ich war mit der Betreuung durch den Sozialdienst zufrieden.

umformuliert in:

- 10. Der Umgang mit den verordneten Hilfsmitteln wurde mir erläutert und mit mir geübt.

²⁵ Das Kaiser-Meyer-Olkin-Kriterium bildet in der Faktorenanalyse die Grundlage zur Prüfung, ob ein Datensatz sich durch Faktoren darstellen lässt und somit für eine Faktorenanalyse geeignet ist.

*) maximales Cronbachs Alpha bei Streichung mindestens eines inhomogenen Items

‡) maximales Cronbachs Alpha bei Hinzunahme mindestens eines homogenen Items

***) keine Berechnung möglich, weil nur ein Item vorliegt

- 15. Ich habe Anleitungen für Zuhause erhalten (z.B. gymnastische Übungen, Ernährungstipps, Trainingsmaßnahmen).
- 51. Ich war mit der Sozialberatung zufrieden.

Diese Items wurden im Patientenbogen Nachbefragung ambulant und im Bezugspersonenbogen Nachbefragung entsprechend geändert.

Item 53 (Die Hilfsmittel sind richtig angepasst, so dass ich sie gut benutzen kann.) wurde als einziges Item gestrichen, weil die Anpassung der Hilfsmittel keine Leistung der Einrichtungen ist und somit keinen Qualitätsaspekt der Rehabilitations-Einrichtung darstellt.

Im Bezugspersonenbogen Nachbefragung sind vor dem 3. Fragen-Block zur Abfrage des Informationsstandes der Bezugsperson („Welche Informationen haben Sie während der Reha erhalten?“) drei Eingangsfragen mit anschließender Sprunganweisung platziert. Bezugspersonen, die diese drei Fragen mit „nein“ beantworten, sollen den 3. Fragen-Block überspringen, weil sie damit angeben, über keine Grundlage zu verfügen über medizinische Details des Patienten informiert zu werden. Für 3 von 70 Bezugspersonenbögen traf dieser Fall zu. Daher wurde im Steuerungskreis beschlossen, diese drei Fragen (Frage 15 bis 17) im Fragebogen zu belassen, auf die Sprunganweisung jedoch zu verzichten und sie an anderer Stelle abzufragen.

Auf Grundlage der Faktorenanalysen wird die Reihenfolge der Items in den Fragebögen geändert. Die gefundenen Faktoren werden inhaltlich chronologisch angeordnet. Innerhalb der Faktoren werden die Items nach den Faktorenladungen absteigend sortiert.

Ergebnisqualität

Zur Messung der Ergebnisqualität wurden in Anlehnung an die ICF-Systematik 33 Rehabilitations-Zielkategorien aus neun ICF-Domänen operationalisiert.

5.4.4 Auswahl der Rehabilitations-Zielkategorien

Im Pretest konnten mindestens zwei und maximal zehn RZK aus mindestens zwei Rehabilitations-Zielkategorien der Ebene I ausgewählt werden. Eine entsprechende Programmierung war im Extranet für den online abrufbaren Behandlerbogen als Plausibilitätsregel hinterlegt. Die Mehrheit der Einrichtungen hat 5 bis 8 RZK ausgewählt. Im Mittel wurden 6,14 RZK ausgewählt (s. Abbildung 3). Die Anzahl ausgewählter RZK ist annähernd symmetrisch um den Mittelwert verteilt.

Abbildung 3: Häufigkeitsverteilung der Anzahl ausgewählter Rehabilitations-Zielkategorien der Ebene II

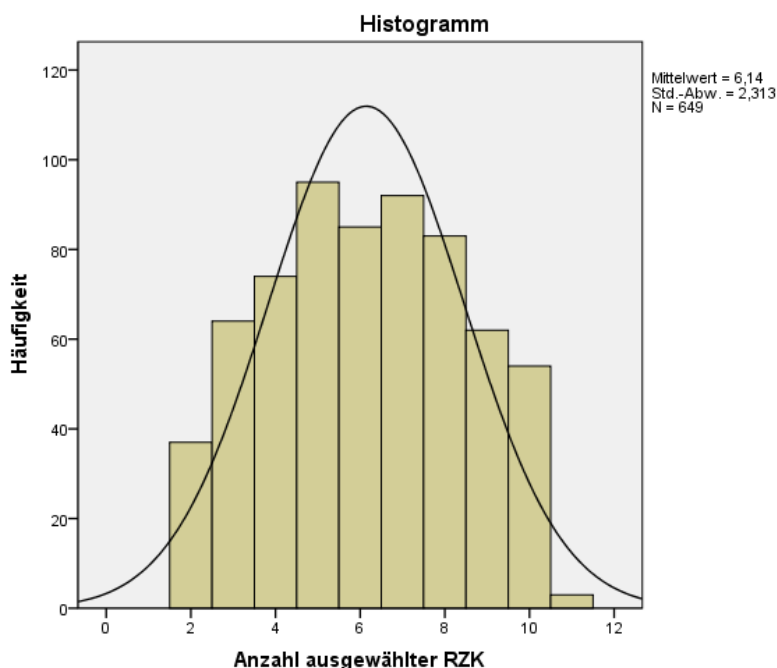
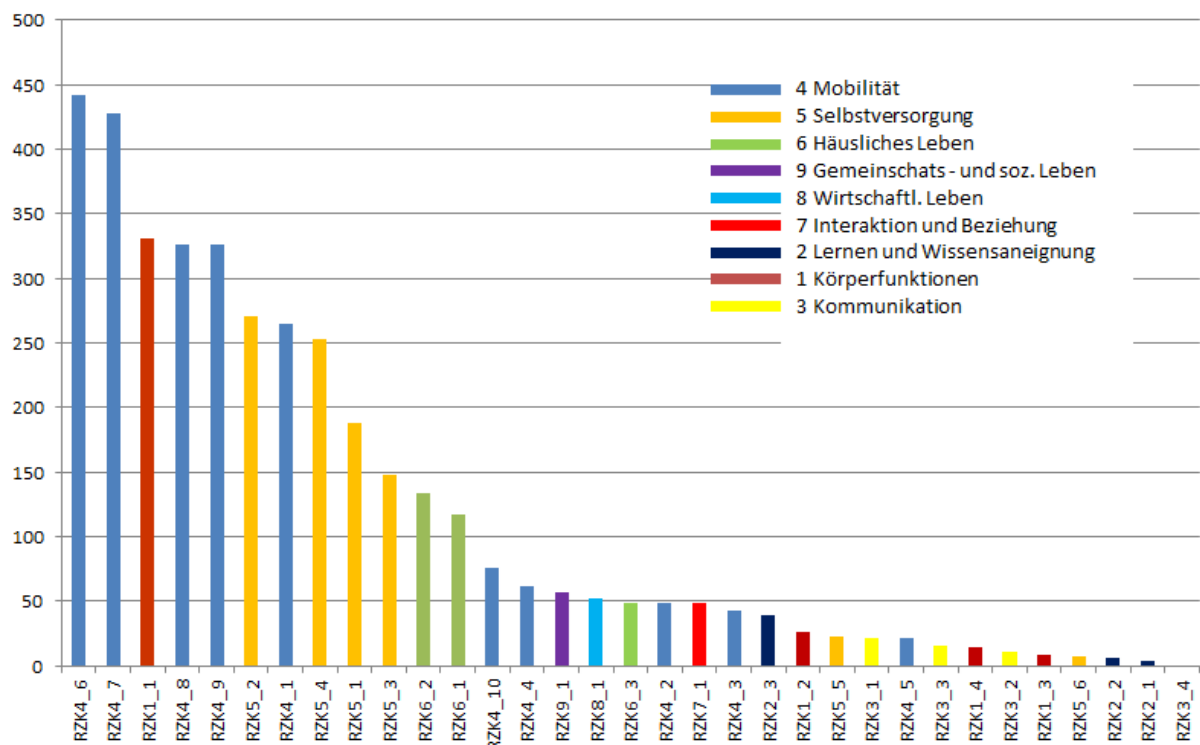


Abbildung 4: Häufigkeitsverteilung der ausgewählten RZK absteigend nach Häufigkeit sortiert



Die 33 RZK wurden unterschiedlich häufig ausgewählt (s. Abbildung 4). An der Spitze der Nennungen liegen die RZK der Rehabilitations-Zielkategorie der Ebene I Mobilität, wobei hier „Gangssicherheit“ (4.6) und „Gehen im Innenbereich“ (4.7) am häufigsten genannt wurden.

Nach dem Mobilitätsblock schließen sich die Selbstversorgungs-RZK „Sich kleiden“ (5.4), „Sich waschen“ (5.1) und „Toilettenbenutzung“ (5.3) an. Denen folgen die RZK „Hausarbeiten erledigen“ (6.2) und „Mahlzeiten zubereiten“ (6.1) der Rehabilitations-Zielkategorie der Ebene I „Häusliches Leben“.

Auf dem 3. Platz folgen „Schmerzen“, diese wurden nur geringfügig häufiger ausgewählt als „Treppensteigen“ (4.8) und „Gehen im Außenbereich“ (4.9). Die weiteren RZK werden eher selten ausgewählt. Die Rehabilitations-Zielkategorie der Ebene I „Lernen und Wissensaneignung“ sowie „Kommunikation“ sind gering vertreten. „Kommunikationsgeräte nutzen“ (3.4) wurde im Pretest als einzige RZK nicht einmal ausgewählt.

Der Steuerungskreis hat daher empfohlen, bei zukünftigen Erhebungen soll die Mindestzahl auszuwählender RZK auf mindestens drei Rehabilitations-Zielkategorien der Ebene I angehoben werden.

5.4.5 Nachträgliche Änderungen der ausgewählten Rehabilitations-Zielkategorien

Im Vorfeld des Pretests wurde im Steuerungskreis die methodische Schwierigkeit von nachträglichen Änderungen der zu t0 ausgewählten RZK intensiv erörtert. Zur Abklärung der Problematik wurden im Pretest die Häufigkeit und die Gründe für nachträgliche Änderungen im Behandlerbogen erfasst.

Bei insgesamt 649 einbezogenen Fällen und 3.986 ausgewählten RZK wurden insgesamt 15 Änderungen dokumentiert, die im Wesentlichen mit einem Herabstufen des Leistungsniveaus der Rehabilitanden begründet wurden. Damit wurde in 2,3% aller Fälle bzw. in 0,4% aller ausgewählten RZK eine Änderung der zu t0 vorgenommenen Auswahl als erforderlich angegeben.

Vor diesem Hintergrund empfiehlt der Steuerungskreis, wegen der geringen Fallzahl und der damit einhergehenden geringen Bedeutung für das Verfahren auf weitere Beobachtungen und Überlegungen für eine gesonderte methodische Lösung zu verzichten.

5.4.6 Berechnung der Ergebnisqualitäts-Scores

Zur Bestimmung der Ergebnisqualität wird ein Vorher-Nachher-Vergleich der für die einzelnen Patienten spezifisch ausgewählten RZK durchgeführt. Zur Deltamessung wird der Leistungsstatus für die ausgewählten RZK zu t0 und zu t2 gemessen. Aus den Unterschieden wird der Score der Ergebnisqualität bestimmt, was als indirekte Veränderungsmessung bezeichnet wird. Der hier zu realisierende Auswertungsalgorithmus muss berücksichtigen, dass die Statusabfragen in der Regel auf Basis ordinaler Antwortskalen²⁶ erfolgen, dass die unterschiedlichen RZK mit unterschiedlichen spezifischen Verbesserungspotentialen behaftet sind und die Deltawerte risikoadjustiert werden sollen. Damit wird ein Algorithmus erforderlich, der die Differenzen ordinaler Daten²⁷ auf Intervallskalenniveau standardisiert darstellt.

Ausgangspunkt ist die Wiederholungsmessung einzelner spezifisch für den Patienten ausgewählter bestimmter RZK. In Abbildung 5 wird beispielhaft gezeigt, wie sich die Fähigkeit eines Patienten Nahrung zu sich zu nehmen zwischen t0 und t2 verändert hat. Beim ersten Item der RZK 1.2 „Nahrungsaufnahme, Schlucken“ zeigt sich z.B., dass der Patient seine Fähigkeit (a) „Wasser, Saft, Kaffee, Tee oder andere Flüssigkeiten“ zu sich zu nehmen bzw. hinunterzuschlucken zu t0 als „mittelmäßig“ einstuft und zu t2 als „gut“. Bei den Fragen a bis e werden unterschiedlich große Verbesserungen (Δ_i) angegeben, die von unterschiedlichen Ausgangspunkten ausgehen ($x_i(t_0)$). Hier stellt sich die Frage, wie die unterschiedlichen Ergebnisse zu einem RZK-Score für den Therapieerfolg bei der Nahrungsaufnahme bzw. beim Schlucken eines Patienten verdichtet werden können.

Abbildung 5: RZK 1.2 Körperfunktion: Nahrungsaufnahme, Schlucken; Antwortbeispiel

Im Folgenden sind unterschiedliche Nahrungsmittel aufgeführt. Wir wollen nun von Ihnen wissen: Können Sie Folgendes zu sich nehmen (hinunterschlucken):					
		gar nicht	schlecht	mittelmäßig	gut
a	Wasser, Saft, Kaffee, Tee oder andere Flüssigkeiten?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
b	Kartoffelbrei, Pudding oder Joghurt?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
c	weich gekochtes Gemüse (z.B. Kartoffeln, Blumenkohl)?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d	feste Speisen (z.B. Brot, Kekse oder Fleisch)?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
e	Mahlzeiten mit unterschiedlichen Bestandteilen (z.B. Eintopf, Suppe mit Einlage)?	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

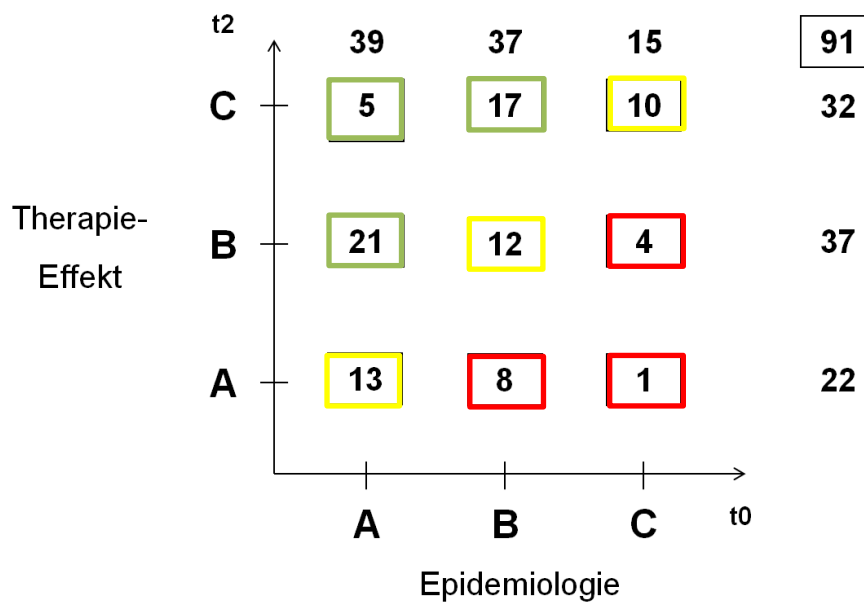
Bezogen auf eine Frage mit drei ordinalen Antwortstufen (A, B, C) soll der Algorithmus beispielhaft verdeutlicht werden.

²⁶ Die Stufen ordinaler Antwortskalen sind nicht äquidistant. Selbst gleiche Stufensprünge bei gleichen Antwortskalen unterschiedlicher Fragen können nicht als äquidistant angenommen werden.

²⁷ Wegen fehlender Äquidistanz zwischen den Stufen sind bei ordinalen Daten mathematische Operationen wie Addition und Subtraktion und darauf aufbauende Operationen nicht zulässig.

Abbildung 6 zeigt die Häufigkeitsverteilung der Deltamessung.

Abbildung 6: Häufigkeitsverteilung für ein 3-stufiges Item und Vorher-Nachher-Messung



In diesem Beispiel haben 91 Patienten eine bestimmte Frage i zu t_0 (x-Achse) und zu t_2 (y-Achse) beantwortet. In dem resultierenden Neun-Zellendesign sind die Häufigkeiten der realisierten Deltawerte eingetragen. So gibt es z.B. fünf Patienten, die zu t_0 Antwort A angegeben haben und zu t_2 die Antwort C. Auf der Hauptdiagonalen sind die Patienten eingetragen, bei denen sich zwischen t_0 und t_2 keine Veränderung eingestellt hat (13, 12, 10). Die Zellen über der Hauptdiagonalen zeigen Verbesserungen an, die unterhalb liegenden Verschlechterungen. Von den 15 Patienten, die zu t_0 die höchste Stufe ankreuzen konnten, haben zu t_2 nur noch 10 ihr Niveau gehalten. Bei vier Patienten hat sich eine Verschlechterung um eine Stufe ergeben, bei einer ist sogar eine Verschlechterung um zwei Stufen eingetreten.

Bei diesem Auswertungsansatz wird die Verteilung zu t_0 als gegeben betrachtet, denn auf welchem Ausgangsniveau die Patienten ihre Therapien beginnen, ist vor allen Dingen epidemiologisch bedingt. Der zu betrachtende Therapieeffekt zeigt sich in den drei vertikalen t_2 -Verteilungen. Ausgehend von der vorliegenden Häufigkeitsverteilung wird für jede einzelne vertikale Verteilung auf Basis der Standardnormalverteilung ein flächenbezogener z-Wert²⁸ für jede Zelle ermittelt. Dazu werden vor der Flächentransformation für jede Zelle die bis zum Mittelwert der betrachteten Zelle kumulierten und zu 100% relativierten Zellenhäufigkeiten ermittelt.

$$Z = F \{ (\text{cum } f - f / 2) * (100 / N) \}$$

wobei

F die Quantilfunktion der Standardnormalverteilung ist.²⁹

Die z-Werte werden getrennt für jede einzelne vertikale Häufigkeitsverteilung bestimmt.

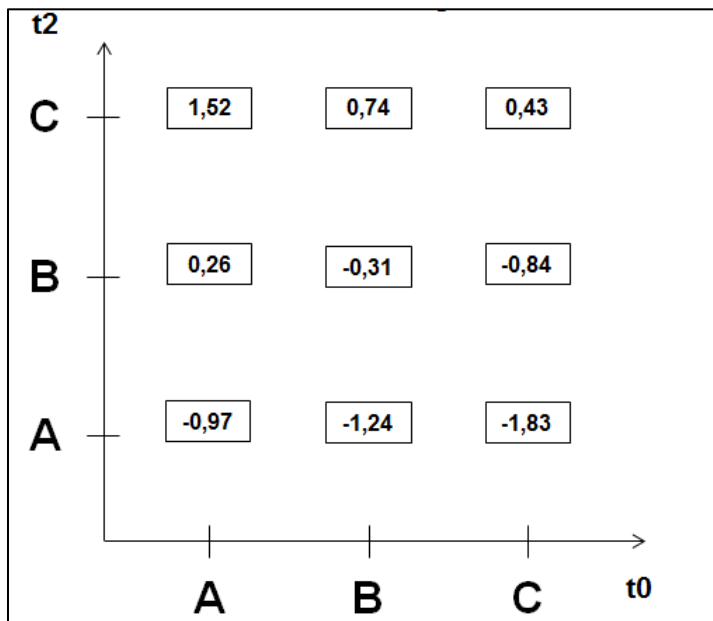
Mit diesem Verfahren wird für jede Zelle ein spezifischer z-Wert ermittelt, der zum Ausdruck bringt, wie häufig die einzelnen Stufen durch therapeutische Interventionen bei einem bestimmten Ausgangsniveau erreicht wurden. Je seltener z.B. eine Zelle besetzt ist, die ein extrem positives Ergebnis darstellt, desto schwieriger war es offenbar dieses Niveau therapeutisch zu erreichen. Der

²⁸ Unter dem flächenbezogenen z-Wert versteht man einen Wert z auf der reellen Zahlenachse, bis zu welchem der Flächeninhalt, den die Dichtefunktion der Standardnormalverteilung mit der reellen Achse einschließt, einen vorgegebenen Prozentwert erreicht. Zur Verdeutlichung ein Rechenbeispiel aus Abbildung 6: Zelle (A, C) [oben links]=5, halbe Zellbesetzung wäre dann 2,5 und die Kumulation bis dort $13+21+2,5=36,5$. Hieraus lässt sich ableiten, dass dieser Wert ca. 93,6% der Zellbesetzungen über Antwort A aus t_0 entspricht. Und einen Flächenanteil von 93,6% erhält man unter der Standardnormalverteilung bei einem z-Wert von $z=1,52$.

²⁹ Siehe Lienert und Raatz, 1998, S. 294

resultierende z-Wert wird entsprechend extremer positiv. Die resultierenden z-Werte aus dem Beispiel sind Abbildung 7 zu entnehmen.

Abbildung 7: z-Werte-Verteilungen für jede der drei t₀-Stufen und den t₂-Ergebnissen



Der Therapieeffekt der fünf Patienten, die zu t₀ Antwort A angegeben haben und zu t₂ die Antwort C, wird auf dieser Grundlage mit einem z-Wert von z=1,52 beschrieben. Die von einer Person besetzte Zelle C->A erhält den z-Wert z=-1,83. Die z-Werte der Hauptdiagonalen im Beispiel besagen, dass keine Änderung erfahren zu haben umso schlechter bewertet wird, je geringer das Ausgangsniveau war (-0,97, -0,31, 0,43).

Die Grundlage der z-Werte bilden die ordinale Information der Antwortstufen sowie die Häufigkeiten, mit denen die theoretisch realisierbaren Vorher-Nachher-Konstellationen tatsächlich gefunden werden. Die so ermittelten z-Werte einer Frage beschreiben damit die in der erhobenen Grundgesamtheit empirisch festgestellten für die einzelnen Ausgangsniveaus spezifischen Therapieeffekte.

Gemäß dem beschriebenen Verfahren werden für alle Fragen mit ordinalen Antwortskalen auf Grundlage der Häufigkeitsverteilungen aller Ausgangsniveaus z-Werte bestimmt. Die z-Werte haben Intervallskalenniveau und können miteinander verrechnet werden. Der RZK-Score für eine Person wird dann als Mittelwert der z-Werte der gefundenen Deltawerte der einzelnen Fragen einer RZK definiert. Ein Patient, dessen drei Antworten zu t₀ und zu t₂ auf der Hauptdiagonalen im 3-stufigen Beispiel liegen, hätte einen z-Wert von -0,28. Der so definierte RZK-Score hat Intervallskalenniveau und ist standardisiert und kann daher mit entsprechenden Scores anderer RZK verrechnet werden. Außerdem kann er als Kriterium für Risikoadjustierungsmodelle (O-Wert) verwendet werden.

Diese Rechenregel wird immer dann angewendet, wenn der Leistungsstatus mit ordinalen Antwortskalen abgefragt wird. Wenn intervallskalierte Antwortskalen zur Statusmessung verwendet werden, wie z.B. bei „Schmerzen“ (RZK 1.1), dann werden die Deltawerte direkt als Differenz der Item-Scores von t₀ und t₂ gebildet ($\Delta_i = x_i(t_1) - x_i(t_0)$) und auf Grundlage der Verteilungskennwerte Mittelwert und Streuung der Differenzenverteilung z-standardisiert.

Diese Form der Ergebnisstandardisierung bewirkt, dass weder die Auswahl der RZK noch die Auswahl von Patienten mit bestimmten Fähigkeitsniveaus einen Einfluss auf die RZK-Scores haben. Eine Verbesserung der Rehabilitationseinrichtungsergebnisse ist durch eine entsprechende strategische Auswahl nicht zu erzielen. Der Einfluss des einrichtungsspezifischen Patientenkollektivs wird durch die Risikoadjustierung nivelliert.

5.4.7 Reliabilität der Ergebnisqualitäts-Scores

RZK-Scores werden als Mittelwert von z-Werten gebildet. In einigen Fällen bilden die z-Werte der einzelnen Fragen dazu die Basis (z.B. RZK 1.2 „Nahrungsaufnahme / Schlucken“: a) flüssige, b) breiige, c) weiche, d) feste, e) unterschiedlich feste Lebensmittel). In anderen Fällen gibt es eine weitere Aggregationsebene, wie beim „Gehen im Innenbereich“, dem „Treppensteigen“ und dem „Gehen im Außenbereich“. Hier bilden die Einzelfragen zunächst Fragenblöcke (z.B. „Gehen im Innenbereich“: 1. Block: „Gehen innerhalb des Zimmers“ mit den Einzelfragen bei Nutzung unterschiedlicher Hilfsmittel a) bis e), deren Ergebnisse dann zum RZK-Score verdichtet werden.

Bei der Prüfung der internen Konsistenz mit Cronbachs Alpha der so aggregierten RZK-Scores der RZK „Gehen im Innenbereich“, „Treppensteigen“ und „Gehen im Außenbereich“ auf beiden Ebenen hat sich gezeigt, dass innerhalb aller Blöcke die jeweils letzte Frage bezogen auf das Leistungsvermögen bei Unterstützung durch eine Hilfsperson den Score dieser Ebene unreliabler macht. Die Frage nach den Hilfspersonen ist im Vergleich zu den anderen Fragen des jeweiligen Blockes also eher inkonsistent. Beim Treppensteigen mindert zudem die erste Frage die Konsistenz, bei der das Leistungsvermögen „ohne Hilfsmittel und ohne Hilfe“ abgefragt wird.

Bei der Aggregation der Blöcke zu den RZK-Scores hat sich zudem gezeigt, dass beim „Gehen im Innenbereich“ der 3. Block (das Gebäude verlassen können) den Gesamtwert unreliabler macht. Beim „Treppensteigen“ mindert der 4. Block (zwei oder mehr Etagen auf- und abwärts steigen können) den Gesamtwert.

Die Inkonsistenz durch die Abfrage der Hilfsperson wird darin vermutet, dass Hilfspersonen nicht in der Konstanz zur Verfügung stehen wie die anderen abgefragten Hilfsmittel (u.a. weil bezogen auf einen Patienten unterschiedliche Hilfspersonen unterschiedlich wirksame Unterstützungen darstellen). Außerdem dürften Hilfspersonen nicht immer zum therapeutischen Setting gehören und damit keine therapeutische Leistung der Einrichtung darstellen und damit wiederum zur Messung der Ergebnisqualität nicht berücksichtigt werden.

Die Inkonsistenz bei der Abfrage der 1. Frage („ohne Hilfsmittel und ohne Hilfe“) wird auf ein schwer überschaubares Layout zurückgeführt.

Die Inkonsistenz des 3. Blockes beim „Gehen in Innenbereich“ (Gebäude verlassen) wird in tatsächlich vorliegenden Unterschieden der Gebäude und der dazu erforderlichen Aktivitäten, sie zu verlassen, gesehen.

Die Inkonsistenz des 4. Blockes beim „Treppensteigen“ (zwei oder mehr Etagen) wird darauf zurückgeführt, dass bei den stationär behandelten Patienten diese Leistungsstufe nicht mehr oder nur gering differenziert (die Patienten können zu t0 und zu t2 überwiegend keine zwei Etagen überwinden), was methodisch bedingt die Reliabilität mindert.

Fazit

Bei den Untersuchungen zur Reliabilität unterschiedlicher Ergebnisqualitätsmessungen mit den RZK-Scores wurden hohe Cronbachs Alpha-Werte gefunden, die durch einzelne Fragen oder Blöcke geringfügig gemindert wurden. Die Konsistenzminderungen durch diese einzelnen Fragen und Blöcke waren in allen Fällen erklärbar.

Daher wird davon ausgegangen, dass die Operationalisierungen der RZK reliable Messergebnisse zur Bestimmung der Ergebnisqualität liefern. Einzelne gezielte Änderungen der Abfragen sollen die Reliabilität weiter verbessern.

Der Steuerungskreis des Projektes gibt daher zum weiteren Verfahren folgende Empfehlungen:

- Der Einbezug der Hilfsperson bei der Abfrage des Leistungsvermögens wird von den externen Experten als erforderlich bewertet, weil diese Unterstützung in der Realität eine bedeutende Rolle spielt. Daher bleibt die Abfrage bestehen, jedoch mit geänderter Wortlaut, um dem Sprachgebrauch der Patienten mehr zu entsprechen. Hilfsperson soll daher durch Begleitperson ausgetauscht werden.
- Das Layout der jeweils ersten Frage soll übersichtlicher gestaltet werden.
- Der 3. Block der RZK „Gehen im Innenbereich“ soll gestrichen werden.

5.4.8 Validierung der Ergebnisqualitäts-Scores

Bei der Konstruktion der Fragebögen wurden für die spätere Prüfung der kriteriumsbezogenen Validität globale Ergebnis-Items als Außenkriterien (Fisseni 2004) aufgenommen. Sie erfassen

- a) das Erreichen der persönlichen Rehabilitations-Ziele (PNS-F56³⁰),
- b) das Erreichen des individuellen Rehabilitations-Ziels (PNS-F59) sowie
- c) die aktuelle Gesamtlebenszufriedenheit (PNS-F60).

Diese Fragen wurden zu t2 abgefragt und werden bei der Auswertung als Außenkriterien verwendet. Betrachtet wird der korrelative Zusammenhang zwischen den RZK-Scores diverser RZK und den drei Außenkriterien.

Die Auswertung von 13 RZK-Scores („Körperposition wechseln“, „Gegenstände anheben und tragen“, „Händegebrauch“, „Gangsicherheit“, „Gehen im Innenbereich“, „Treppensteigen“, „Gehen im Außenbereich“, „Transportmittel nutzen“, „Sich waschen“, „Körperpflege“, „Toilettenbenutzung“, „Mahlzeiten zubereiten“ und „Hausarbeiten erledigen“) zeigt, dass entweder der Gesamt-Score, ein um einen Teilaspekt reduzierter RZK-Score oder ein Aspekt des RZK-Scores mit mindestens einem Außenkriterium hoch korreliert (s. Anlage 12). Mit der Analyse dieser 13 am häufigsten ausgewählten RZK werden 87% aller ausgewählten RZK einbezogen.

Diese Befunde sprechen dafür, dass die hoch komplexe Aggregation der RZK-Scores sowie die gewählten Operationalisierungen valide Ergebnisse liefern. Die hier untersuchten RZK-Scores sind damit inhaltlich valide.

Die Überprüfung der restlichen RZK-Scores steht noch aus und bedarf einer Stichprobe, die möglichst nahe an das im Routineverfahren zu erwartende Datenvolumen reicht.

5.4.9 Risikoadjustierung der Ergebnisqualitäts-Scores

Auf Grundlage der berechneten RZK-Scores „Gangsicherheit“ und „Treppensteigen“ und den im Behandlerbogen erfassten Confoundern soll exemplarisch gezeigt werden, dass damit taugliche Modelle zur Risikoadjustierung der Ergebnisqualitäts-Scores gefunden werden können.

In das Startset der Modellbildung werden solche Confounder aufgenommen, deren Stufen einen möglichst signifikanten Unterschied zwischen den betrachteten RZK-Scores erzeugen, wie z.B. die Nutzung von Inkontinenzvorlagen bei der „Gangsicherheit“. Hier besteht zwischen Patienten, die vor der Rehabilitation Inkontinenzvorlagen genutzt haben bzw. nicht genutzt haben, bei der Verbesserung der „Gangsicherheit“ ein hoch signifikanter Unterschied ($p < 0,001$, s. Anlage 13). Von allen möglichen signifikanten Confoundern fließen schließlich nur die in das Regressionsmodell ein, die zusammen genommen ein hoch prädiktives und sparsames Vorhersagemodell³¹ bilden.

In das Modell zur Vorhersage der Verbesserung der Gangsicherheit fließen die Prädiktoren:

- Bewilligungsdauer (Beta=-0,16, $p=0,04$),
- Pflegestufe³² (Beta=-0,24, $p < 0,01$),
- Barthel-Subskala Sich waschen (Beta=0,20, $p=0,03$),
- Barthel-Subskala Aufsetzen/Umsetzen (Beta=0,17, $p=0,06$),
- Hinweise auf eine Depression (Beta=-0,16, $p=0,04$) und
- neue medizinische Probleme nach Rehabilitations-Ende (Beta=-0,15, $p=0,07$) ein.

Das Multiple R=0,513 und ist dabei hoch signifikant. Der Determinationskoeffizient $R^2=0,263$. Die Modellgüte ist damit noch akzeptabel.

Wenn anstatt der einzelnen Barthel-Index-Subskalen nur der Barthel-Index als Gesamtscore (BI) zur Modellbildung für die Vorhersage der Verbesserung der Gangsicherheit aufgenommen wird, dann fließen die Prädiktoren: Bewilligungsdauer (Beta=-0,14, $p=0,08$), Pflegestufe (Beta=-0,22, $p < 0,01$), Barthel-Index (Beta=0,23, $p < 0,01$), Hinweise auf eine Depression (Beta=-0,15, $p=0,06$) und neue medizinische Probleme nach Rehabilitations-Ende (Beta=-0,19, $p=0,02$) ein. Das Multiple R=0,475 und ist dabei hoch signifikant. Der Determinationskoeffizient $R^2=0,226$. Die Modellgüte ist damit noch

³⁰ Patientennachbefragungsbogen, Frage 56

³¹ Regressionsmodelle sollen hoch prädiktiv sein (entspricht einem hohen R^2), das jedoch mit möglichst wenigen Prädiktoren (Sparsamkeit), die dann einen inhaltlich umgrenzten Beitrag zur Vorhersage beitragen.

³² Ab dem Jahr 2017 wurden die 3 Pflegestufen durch 5 Pflegegrade abgelöst.

akzeptabel, liegt aber deutlich unter dem o.g. Modell mit den einzelnen BI-Subskalen-Scores. Dieser Befund spricht dafür, alle 10 Subskalen-Scores des Barthel-Index im Behandlerbogen weiterhin differenziert abzufragen und nicht nur den Gesamtscore des Barthel-Index.

In das Modell zur Vorhersage der Verbesserung in der RZK „Treppensteigen“ fließen die Prädiktoren:

- Hilfsmittelnutzung zu t0, Handstock (Beta=-0,55, p<0,001),
- Hilfe im Alltag vor der Rehabilitation (Beta=-0,40, p<0,01),
- Gesamtscore des kognitiven FIM³³ (Beta=0,38, p<0,01) und
- neue medizinische Probleme nach Rehabilitations-Ende (Beta=-0,45, p=0,001) ein.

Das Multiple R=0,833 und ist dabei hoch signifikant. Der Determinationskoeffizient R²=0,694. Die Modellgüte ist als gut zu bezeichnen.

Die Modelle könnten zur Risikoadjustierung der Pretestergebnisse verwendet werden. Sie könnten jedoch nicht ungeprüft auf den Routinebetrieb übertragen werden. In der Routine sind für alle Qualitätsunterdimensions-Scores, die risikoadjustiert werden sollen, neue Modelle zu bestimmen.

5.4.10 Prüfung alternativer Abfragen bei Ausfall von Patienten

Unter den Patienten der geriatrischen Rehabilitation ist die Zahl derer mit kognitiven Einschränkungen nicht unbedeutend. Umfangreiche Erhebungen gehen von ca. 40% kognitiv beeinträchtigter Patienten aus (Gassmann 2007). Daher wurden unterschiedliche Ansätze entwickelt, den dadurch hervorgerufenen Datenverlust möglichst gut zu kompensieren. Im Pretest wurden diese Ansätze auf ihre Tauglichkeit hin empirisch überprüft. Dazu war der Pretest als Parallelerhebung konzipiert worden, um zu jeder Patientenangabe korrespondierende Angaben der Bezugspersonen vorliegen zu haben, die unabhängig voneinander erhoben wurden. Im Vergleich der Angaben beider Quellen wurde dann deren Übereinstimmung untersucht.

Bei der Messung der *Ergebnisqualität* mit der patientenspezifischen Auswahl an RZK gab es eine vollständige Parallelerhebung bei den Bezugspersonen. Patienten und Bezugspersonen erhielten zu t0 und zu t2 genau dieselbe RZK-Auswahl mit der Bitte vorgelegt, sie unabhängig voneinander zu beantworten.

Bezogen auf die *Prozessqualität* und die *Patientenzufriedenheit* wurden im Bezugspersonen-nachbefragungsbogen einzelne Items aufgenommen, die inhaltlich mit denen im Patienten-nachbefragungsbogen übereinstimmten, deren Formulierungen aber deutlich machten, dass die Bezugsperson hier eine Fremdbeurteilung für den Patienten vornimmt.

Schließlich wurden *globale Einschätzungen* der Behandler *zum Erfolg der Rehabilitations-Maßnahme* im Vergleich der entsprechenden Patientenselbsteinschätzungen untersucht. Nachfolgend werden die Ergebnisse dieser drei Ansätze vorgelegt.

5.4.10.1 Alternativen bei der Ergebnisqualitätsmessung

Das beschlossene Gesamtkonzept für das Entwicklungsprojekt Geriatrie sieht vor, dass bei Nicht-Teilnahme von Patienten, sofern verfügbar, Bezugspersonen als Datenquelle zur Messung der Ergebnisqualität herangezogen werden können. Die Bezugspersonen sollen dieselben Fragen zum Leistungsstatus der Patienten zu t0 und zu t2 zur Beantwortung vorgelegt bekommen.

Auf Grundlage der Pretest-Daten wurde daher überprüft, ob Bezugspersonen und Patienten die Fragen zu den einzelnen RZK unabhängig voneinander übereinstimmend beantworten, wozu eine Parallelerhebung durchgeführt wurde. Den Bezugspersonen wurden, anders als im Routinebetrieb, auch wenn Patienten die Fragebögen ausfüllen wollen und können, die Fragen zur Messung der Ergebnisqualität zu t0 und zu t2 vorgelegt.

Für 60 Patienten und Bezugspersonen liegen ausgefüllte t0- und t2-Fragebögen vor. Auf dieser Grundlage konnten die RZK-Scores von sieben RZK („Schmerzen“, „Körperposition wechseln“, „Gangsicherheit“, „Gehen im Innenbereich“, „Treppensteigen“, „Gehen im Außenbereich“ und „Körperpflege“) für Patienten und Bezugspersonen berechnet werden.

³³ Functional Independence Measure (FIM), kognitiver Teil (Granger et al., 1993)

Für die RZK „Schmerzen“ werden Schmerzen 1) in Ruhe, 2) bei Belastung und 3) bei Bewegung abgefragt. Die Angaben von Patienten und Bezugspersonen zu den Schmerzen in Ruhe korrelieren kaum miteinander ($r=0,37$, $p=0,24$). Bei den Angaben zu den Schmerzen bei Belastung und bei Bewegung finden sich dagegen hohe Korrelationen ($r=0,89$, $p<0,001$ und $r=0,79$, $p=0,004$).

Ebenso zeigen sich bei fünf von sechs RZK hohe Korrelationen zwischen den Patienten- und den Bezugspersonenangaben. Beim „Gehen im Außenbereich“ erreicht die Korrelation den Wert $r=0,23$. Hier liegt vermutlich ein Artefakt vor, das darin besteht, dass bei den stationären Patienten eine geringere Variation in den Ergebnissen vorliegt, was die Korrelationen aus methodischen Gründen mindert.

Neben dem Vorliegen hoher Korrelationen ist zudem notwendig, dass zwischen den Mittelwerten der Patienten- und Bezugspersonenangaben keine Unterschiede bestehen. Die dazu durchgeführten t-Tests für abhängige Stichproben sollten daher einen Alpha-Fehler von mindestens 20% feststellen, was bis auf einen der Fälle auch so gefunden wurde. Nur bei RZK 5.2 „Körperpflege“ wurde eine größere Differenz zwischen dem Patienten- und dem Bezugspersonendelta gefunden, die jedoch weit davon entfernt war signifikant zu sein ($p=0,11$ ³⁴).

Fazit dieser Auswertungen ist, dass davon ausgegangen werden kann, dass gut beobachtbare Ereignisse von Patienten und Bezugspersonen überwiegend übereinstimmend eingeschätzt werden und sich die so operationalisierten RZK dazu eignen, bei Nicht-Teilnahme der Patienten, von den Bezugspersonen valide beantwortet zu werden. Die Maßgabe, die Ergebnisqualität bei Ausfall der Patienten auf Basis von Bezugspersonenangaben zu messen, kann damit als tragfähig angesehen werden.

5.4.10.2 Alternativen bei der Messung von Prozessqualität und Patientenzufriedenheit

Der Pretest wurde auch dazu verwendet, um zu überprüfen, inwieweit einzelne spezifische Items zur Prozessqualität sowie Items zur globalen Patientenzufriedenheit und globalen Ergebnisqualität von Patienten und Bezugspersonen übereinstimmend beantwortet werden, um ggf. bei Nicht-Teilnahme der Patienten ersatzweise auf die entsprechenden Antworten der Bezugspersonen zurückgreifen zu können.

Im Bezugspersonennachbefragungsbogen wurden dazu 14 Items aufgenommen, die nahezu identisch zum Patientennachbefragungsbogen stationär (PNS) formuliert waren. Für jedes Item-Paar wird eine Produkt-Moment-Korrelation sowie ein t-Test für abhängige Stichproben gerechnet. Mit der Korrelation wird die lineare Assoziation zwischen den entsprechenden Patienten- und Bezugspersonenantworten bestimmt. Mit dem t-Test³⁵ wird abgesichert, dass beide Items im Mittel auf gleichem Niveau beantwortet werden.

Die Korrelationen variieren zwischen $r=0,114$ und $r=0,755$ bei 34 bis 56 Messwert-Paaren und sind dabei teilweise hochsignifikant (s. Tabelle 6). Der Alpha-Fehler der t-Tests liegt durchgehend bei über 20%. Damit könnten die hoch korrelierenden Items als brauchbarer Ersatz für ausgefallene Patientenangaben verwendet werden, da sich die Angaben vom Bewertungsniveau her nicht bedeutsam unterscheiden.

Tabelle 6: Korrelationen zwischen Patienten- und Bezugspersonenangaben, Alpha-Fehler (p) und Fallzahl (N)

³⁴ Auf einer breiteren Datenbasis sollte geprüft werden, ob sich hier ein Stichprobenartefakt oder ein inhaltlich substantieller Unterschied zeigt, der eine Item-Revision oder neu formulierte Items erforderlich macht.

³⁵ Um Mittelwerte mit dem t-Test auf Gleichheit zu prüfen, wird ein Alpha-Fehler von mindestens 20 % gefordert (Bortz, 2005, S. 122).

BN		PNS			
Item-Nr.	Inhalt	Item-Nr.	Korrelation	p	N
44	möglichst viel selbst machen	6	0,114	0,434	49
45	auf Bitten eingegangen	34	0,294	0,041	49
46	motiviert Reha-Ziele zu verfolgen	20	0,502	< 0,001	51
47	Sorgen & Ängste äußern	39	0,398	0,013	38
48	Intimsphäre geachtet	31	0,235	0,100	50
49	Privatsphäre geachtet	32	0,652	< 0,001	48
50	freundlicher Umgang	37	0,643	< 0,001	51
54	gut gekümmert	50	0,609	< 0,001	56
60	Essen	44	0,581	< 0,001	52
63	persönl. Reha-Ziele erreicht	56	0,411	0,002	55
64	HiMi gut angepasst	53	0,221	0,209	34
65	Gelerntes anwenden	55	0,202	0,135	56
66	profitiert	54	0,233	0,087	55
83	indiv. Reha-Ziel erreicht	59	0,755	< 0,001	47

Auf Beschluss des Steuerungskreises soll im Bezugspersonennachbefragungsbogen bis auf die Einschätzung der Erreichung des individuellen Rehabilitations-Ziels auf alle Angehörigen-Items verzichtet werden, da sie mit $r=0,755$ ($p<0,001$, $N=47$) und einem Mittelwert der gepaarten Differenzen von 0,02 ($p=0,844$) bestens als Alternative für die Patientenbefragung geeignet ist.

5.4.10.3 Behandler einschätzungen als Alternative für globale Ergebnisbewertungen der Patienten

Am Ende der Rehabilitation werden von den Behandlern Einschätzungen des Rehabilitations-Erfolges erfasst (s. Behandlerbogen, Teil A4 - Rehabilitations-Ende). Es wird zum einen das Ausmaß erfragt, in dem der Patient nach Einschätzung der Behandler von der Rehabilitation profitiert hat. Zum anderen soll die Wahrscheinlichkeit angegeben werden, mit der der Patient sein individuelles Rehabilitations-Ziel nach der Entlassung realisieren können wird.

Diese beiden Abfragen wurden im Behandlerbogen aufgenommen, um zu überprüfen, in wie weit die Behandlerurteile mit denen der Patienten übereinstimmen. Angedacht war die Möglichkeit, bei Ausfall der Patientenangaben auf die Behandlerurteile zurückgreifen zu können.

Bezogen auf die erste Frage besteht zwischen den behandlerseitigen Einschätzungen zu denen der Patienten nur ein geringer Zusammenhang, $r=0,26$ ($p<0,001$). Bei der zweiten Frage liegt er bei $r=0,36$ ($p<0,001$). Zudem unterscheiden sich die Mittelwerte zwischen Behandler- und Patienteneinschätzung hoch signifikant. Die Behandler unterschätzen den Rehabilitations-Profit um 0,3 Skaleneinheiten³⁶ hoch signifikant ($p=0,001$). Gleichzeitig überschätzen sie die Wahrscheinlichkeit für das Erreichen des individuellen Rehabilitations-Ziels extrem um 0,8 Skaleneinheiten³⁷ hoch signifikant ($p<0,001$).

Diese Befunde zeigen, dass die behandlerseitigen Einschätzungen sich nicht dazu eignen, als Ersatz für fehlende Patientenangaben herangezogen zu werden.

Der Steuerungskreis empfiehlt: Beide Fragen sollen weiterhin im Behandlerbogen abgefragt und die Entwicklung der Antworten beobachtet und ausgewertet werden.

5.4.11 Feedback der Befragten zur Befragung

Patienten und Bezugspersonen fanden die Fragebögen überwiegend inhaltlich angemessen und nur in wenigen Fällen zu lang. Die mittlere Bearbeitungszeit für den Nachbefragungsbogen, mit dem die Prozessqualität, die Zufriedenheit und die individualisiert gemessene Ergebnisqualität abgefragt wird, lag für stationäre Patienten bei 32 Minuten (80%-Perzentil³⁸=45Min, 95%-Perzentil=60Min), für

³⁶ auf einer Skala von 1 bis 5

³⁷ auf einer Skala von 1 bis 5

³⁸ 80%-Perzentil bedeutet, dass 80% der beobachteten Werte unter diesem Wert liegen.

ambulante Patienten bei 36 Minuten (80%-Perzentil=60Min) und für Bezugspersonen bei 16,5 Minuten (80%-Perzentil=23Min, 95%-Perzentil=40Min). Diese Ergebnisse besagen, dass die Fragebogenlänge für die Befragten überwiegend akzeptabel war und kein Gebot zur Kürzung zu erkennen ist.

5.4.12 Änderungen für die Pilotphase

Das *Erhebungsdesign* in der Pilotphase wurde gegenüber dem Pretest geändert. Im Pretest galt die Maßgabe, zu jedem Patienten eine Bezugsperson für die Befragung zu akquirieren, um eine Parallelerhebung zu realisieren. Diese Vorgabe gibt es in der Pilotphase nicht, wie es sie auch im Routinebetrieb nicht gibt. Es wird aber nicht ausgeschlossen, dass beide am Verfahren teilnehmen, da dann die Bezugspersonen zur Prozessqualität und zur Zufriedenheit spezifisch befragt werden können.

Eine weitere Änderung besteht darin, dass in der Pilotphase die Drop-Out-Fälle im Behandlerbogen (Teil A1 und Teil A2) dokumentiert werden. Anders als in der Routine wird in der Pilotphase, wie schon im Pretest, darauf verzichtet, eine Mindestanzahl an zu akquirierende Fälle vorzugeben. Daher gibt es in der Pilotphase ebenfalls keine Nachfassaktionen.

Neben redaktionellen Änderungen wird die Abfrage der Basis-Assessments bei Aufnahme im *Behandlerbogen* grundlegend geändert (Teil A2). Die Abfrage der Beeinträchtigungen (Sehen, Hören, Schlucken usw.) erfolgt in einer einheitlichen Form. Sie beginnt mit der Eingangsfrage: „Gibt es Hinweise auf ein Problem ...?“, gefolgt von den Antwortmöglichkeiten: „ja“ und „nein“. Falls Hinweise dafür vorliegen, folgen vier weitere Fragen:

1. Mit welchem Verfahren bzw. Instrument wurde der Verdacht abgeklärt?
2. Welchen Befund, welches Ergebnis lieferte das verwendete Verfahren?
3. Wurde der Verdacht damit bestätigt oder nicht? und
4. Erläuterung, falls keine Abklärung durchgeführt wurde.

Im Zuge der Vereinheitlichung dieser Abfrage wurde in drei Fällen auf die Vorgabe bestimmter Instrumente für bestimmte Beeinträchtigungen verzichtet. Beim Hören, dem Bestehen eines Delirs bzw. einer Depression wird nicht mehr vorgegeben, den Verdacht nach Lachs (Lachs et al., 1990) oder mit dem CAM (Inouye SK et al., 1990) oder dem GDS-15 (Yesavage et al., 1982) abzuklären. Auf die Abfrage des motorischen FIM (Granger et al., 1993) und des DEMMI (De Morton et al., 2008) wird in der Pilotphase ganz verzichtet.

In Teil A3 des Behandlerbogens, in dem die Auswahl der Rehabilitations-Zielkategorien erfolgt, wird der Auswahlmodus so geändert, dass aus den 33 Rehabilitations-Zielkategorien nun mindestens drei und maximal zehn Rehabilitations-Zielkategorien ausgewählt werden können, die aus mindesten drei ICF-Domänen stammen.

In Teil A4 des Behandlerbogens wird zukünftig darauf verzichtet, Änderungen von Rehabilitations-Zielkategorien angeben zu können, die sich im Laufe der Rehabilitation als erforderlich herausgestellt haben. Weiterhin werden keine behandlerseitigen Einschätzungen des Patienten-Bezugspersonen-Verhältnisses und der Beteiligung der Bezugspersonen an Rehabilitations-Maßnahmen erfasst.

Zur Messung der *Prozessqualität* und der *Zufriedenheit* werden drei Items umformuliert und eingestrichen, weil es sich auf Leistungen bezieht, die nicht von den Rehabilitations-Einrichtungen erbracht werden (s. Item-Revision, Kapitel 5.4.3). Diese Änderungen betreffen jeweils entsprechend alle drei Nachbefragungsbögen.

Im Bezugspersonennachbefragungsbogen werden zusätzlich alle Items gestrichen, die eine Fremdbeurteilung des Patienten darstellen, mit Ausnahme des Erreichungsgrades des individuellen Rehabilitations-Zieles des Patienten (s. Kapitel 5.4.10.2).

Bei der Messung der *Ergebnisqualität* gibt es bei den RZK 4.7 „Gehen im Innenbereich“, 4.8 „Treppensteigen“ und 4.9 „Gehen im Außenbereich“ folgende Änderungen: In der höchsten Unterstützungsklasse dieser drei RZK-Abfragen wird „Hilfsperson“ durch „Begleitperson“ ersetzt (s. Kapitel 5.4.7). Beim „Gehen im Innenbereich“ wird der 3. Block („das Gebäude verlassen“) gestrichen und beim „Treppensteigen“ wird das Layout an das der anderen RZK angeglichen.

Zur alternativen Messung der Ergebnisqualität durch die Befragung der Bezugspersonen wird eine parallele Abfrage-Version für jeden der 33 RZK erstellt, um den befragten Bezugspersonen einen Perspektivwechsel zu ersparen und damit eine Fehlerquelle auszuschließen. Dazu werden nur die Eingangsformulierungen von „Können Sie ...“ in „Kann Ihr(e) Angehörige(r) ...“ sowie die Antwortlabel entsprechend geändert („betrifft mich nicht“ zu „betrifft sie/ihn nicht“). Die abgefragten Inhalte bleiben ansonsten unverändert.

6 Pilotphase

Die Pilotphase stellte in erster Linie eine Simulation der Routine dar, bei der die Abläufe in den Einrichtungen und in den Institutionen in denen eine Befragung durchgeführt wurde, die unterstützenden IT-Strukturen und die eingesetzten Instrumente im Routinemodus an einer zweiten unabhängigen Patientenstichprobe überprüft wurden. Auf Grundlage der dabei erhobenen Daten wurden die revidierten Instrumente erneut hinsichtlich ihrer Messeigenschaften empirisch untersucht.

6.1 Ziele

Im Vordergrund der Pilotphase stand die Überprüfung der Prozessabwicklung, wie sie für den Routinebetrieb geplant ist. Dazu gehörte auch eine Einschätzung der mit der Durchführung verbundenen Aufwände, die auf Seiten der Rehabilitations-Einrichtung und auf der Seite der Auswertungsstelle nach § 299 Abs. 3 SGB V, aktuell das BQS-Institut, anfallen, um die Aufwände für den Routinebetrieb kommunizieren, planen und kalkulieren zu können.

Die empirische Überprüfung der revidierten Befragungsinstrumente hatte zum Ziel, die im Pretest festgestellten Ergebnisse an einer neuen unabhängigen Stichprobe zu überprüfen. Damit soll gezeigt werden, dass die Pretestergebnisse keinen Zufallsbefund darstellen. Dazu wurden wieder die Item-Kennwerte betrachtet, außerdem die Faktorenstrukturen sowie die Reliabilität und Validität der auf dieser Grundlage für den Routinebetrieb abschließend definierten Qualitätsunterdimensionen.

Im QS-Reha[®]-Verfahren gehört zum Routinebetrieb eine Ergebnisdarstellung für die teilnehmenden Rehabilitations-Einrichtungen mit Drop-Out-Analyse, Qualitätssynopse sowie Befragungsergebnissen auf der Ebene der Qualitätsunterdimensionen und auf Item-Ebene für die betrachteten Qualitätsdimensionen: Ergebnisqualität, Prozessqualität und Zufriedenheit (für Details s. Kapitel 7).

Ein Musterergebnisbericht wird auf Grundlage der hier definierten Qualitätsunterdimensionen vorgelegt und diskutiert, der die Struktur der Ergebnisberichte der bereits in der Routine des QS-Reha[®]-Verfahrens vertretenen Indikationsbereiche aufweist.

6.2 Methode

Das Design der zweiten Studie weicht in einzelnen Punkten von dem im Pretest realisierten ab, um die Umsetzung des Designs, wie es in der Routine vorgesehen ist, zu überprüfen.

Für die Befragungen sollen in den Einrichtungen Patienten und Bezugspersonen akquiriert werden. Für die teilnehmenden Bezugspersonen gibt es dann zwei Varianten, je nachdem, ob der dazugehörige Patient an der Befragung teilnimmt oder nicht. Im ersten Fall wird die Ergebnisqualität über den Patienten gemessen und die Bezugsperson wird nur zu t2 befragt und zwar zur Prozessqualität und Zufriedenheit spezifisch aus seiner Sicht. Nimmt der Patient an der Befragung nicht teil, dann wird die Ergebnisqualität über die teilnehmende Bezugsperson ermittelt. Sie erhält zu t0 den Fragebogen zur Nullpunktmessung. Zu t2 erhält sie einen Fragebogen mit der Wiederholungsmessung sowie der Abfrage der Prozessqualität und Zufriedenheit aus ihrer Sicht (vgl. Tabelle 4).

Die Einrichtungen wurden nicht angewiesen, mit besonderem Aufwand zu jedem akquirierten Patienten auch eine Bezugsperson für die Befragungen zu gewinnen³⁹. Denn dies wurde von den am Pretest teilnehmenden Einrichtungen als sehr aufwändig beschrieben und daher in der Pilotphase nicht verlangt. Dieses Vorgehen wird ebenfalls für die Routinephase des QS-Reha[®]-Verfahrens empfohlen.

Wie in der Routine werden hier die Drop-Out-Fälle im Behandlerbogen dokumentiert, wobei im Indikationsbereich geriatrische Rehabilitation der Drop-Out-Fall dann eintritt, wenn weder der Patient noch eine Bezugsperson für das Verfahren gewonnen werden konnte.

Die ausgefüllten Nachbefragungsbögen wurden von den Patienten bzw. den Bezugspersonen, wie im Routineverfahren, direkt mit den beiliegenden frankierten und adressierten Rückumschlägen an die zentrale Datenerfassungsstelle beim BQS-Institut in Hamburg gesendet. Die Einrichtungen konnten sich über das Extranet über den Fragebogeneingang der Nachbefragung informieren und ggf. Nachfassaktionen starten.

³⁹ Die Maßgabe Bezugspersonen forciert zu akquirieren gab es nur im Pretest, um eine Parallelerhebung zu realisieren, die in der Pilotphase ebenso wie in der Routine nicht vorgesehen war.

6.3 Ablauf

6.3.1 Einrichtungsakquise

Die Akquise der Rehabilitations-Einrichtungen wurde vom Bundesverband Geriatrie e.V. (BVG) unterstützt. Insgesamt wurden 12 Einrichtungen für die Pilotphase rekrutiert, 9 Einrichtungen, die stationäre geriatrische Rehabilitation durchgeführt sowie 3 Einrichtungen, die mit ihrer ambulanten geriatrischen Rehabilitation an der Pilotphase teilnahmen.

6.3.2 Schulung der Teilnehmer

Es gab eine Schulung der zentralen Koordinatoren am 06.04.2017 in Berlin in den Räumen des GKV-Spitzenverbandes, an der sechs Personen teilgenommen haben. Die Schulung wurde von zwei Mitarbeitern des BQS-Instituts durchgeführt.

Es wurden das Grundkonzept des Qualitätssicherungsverfahrens ausführlich vorgestellt, die Ergebnisse des Pretests und die Ziele der Pilotphase dargestellt sowie die Abläufe und technischen Realisierungen eingehend erläutert.

Zudem erhielten die Teilnehmer einen Ordner mit Unterlagen, zum einen, einen Überblick über das Verfahren, Handlungsanweisungen zur Rehabilitations-Zielvereinbarung, der Struktur der Rehabilitations-Zielkategorien, sowie zum anderen, einen Auszug der Operationalisierung einer Rehabilitations-Zielkategorie und den Behandlerbogen.

Abschließend wurde die Schulung mit einem Fragebogen evaluiert. An der Evaluation der Schulung haben sich 4 Schulungsteilnehmer beteiligt. Insgesamt ist die Schulung von den Teilnehmern sehr positiv bewertet worden. Auf einer Skala von 1 (sehr gut) bis 6 (völlig unzureichend) lag der Mittelwert des Gesamteindrucks bei 1,0, der Mittelwert über alle neun Bewertungskriterien bei 1,4.

6.3.3 Durchführung der Pilotphase

In den Einrichtungen wurde die Akquise der Patienten und Bezugspersonen am 11.05.2017 gestartet. Für ambulant behandelte Patienten endete der Akquisezeitraum am 15.07.2017, für stationär behandelte Patienten am 22.09.2017. Der für die ambulante Reha verkürzte Zeitraum sollte dem Umstand Rechnung tragen, dass sich die Dauer der Rehabilitation für stationäre Patienten von z.B. 3 Wochen bei ambulanten Patienten auf 5 Wochen verlängern kann, weil die Behandlungstage nicht unmittelbar aufeinander folgen müssen. Durch die unterschiedlich langen Akquisezeiträume gab es am Ende für ambulant und stationär behandelte Patienten wieder einheitliche Abgabetermine. Am 22.09.2017 wurden die Einrichtungen gebeten, die Behandlerbögen abzuschließen und die gesammelten Beginn-Bögen dem BQS-Institut zum Einlesen zuzusenden. Annahmeschluss für die Nachbefragungsbögen war der 06.10.2017.

Wie schon beim Pretest wurden die Angaben der Fragebögen elektronisch erfasst, in die Datenbank des BQS-Instituts eingelesen und mit den Behandlerbögen zusammengeführt. Nach der Bereitstellung der kompletten Datensätze, bestehend aus Behandlerbogen, den Patienten- (alternativ Bezugspersonenbogen) zu Rehabilitations-Beginn, den Patientenbogen ambulant und stationär (alternativ den entsprechenden Bezugspersonenbogen) der Nachbefragung, folgten die statistischen Auswertungen.

6.4 Ergebnisse

6.4.1 Stichprobe

Insgesamt wurden 568 Fälle mit dem Behandlerbogen dokumentiert. Davon gab es 464 Zusagen zur Teilnahme (81,7%) und 104 Drop-Out-Fälle (18,3%). Von den 464 Zusagen kamen 451 von Patienten (79,4%) und 58 von Bezugspersonen (10,2%), wobei 45 Zusagen gemeinsam von Patienten und Bezugspersonen (7,9%) kamen.

Insgesamt lagen 210 ausgefüllte Nachbefragungsbögen vor, was einer Rücklaufquote (RQ) von 45,3% (210 von 464) entspricht. Davon waren 201 Bögen von Patienten (RQ=44,6%) und 9 von Bezugspersonen (15,5%). Von den 201 auswertbaren Patientennachbefragungsbögen waren 171 von stationär behandelten Patienten (RQ=43,4%) und 30 von ambulant behandelten Patienten (RQ=52,6%) beantwortet worden.

Insgesamt gab es 117 nicht-teilnehmende Patienten. Als Begründung für die Nicht-Teilnahme gaben 72 Patienten an, die Teilnahme abzulehnen (61,5%), für 25 war die Teilnahme wegen kognitiver oder

psychischer Probleme nicht möglich (21,4%) und für 11 war sie aus sprachlichen Gründen nicht möglich (9,4%). 8 Patienten hatten „andere Gründe“ (6,8%) und in einem Fall gab es keine Angabe (0,9%, s. Anlage 14).

Insgesamt gab es 510 nicht-teilnehmende Bezugspersonen. Gründe für die Nicht-Teilnahme von Bezugspersonen lagen in 23 Fällen daran, dass sie die Teilnahme ablehnten (4,5%). In 14 Fällen lehnte der Patient die Teilnahme seiner Bezugsperson ab (2,7%). Für 2 Bezugspersonen war die Teilnahme wegen kognitiver oder psychischer Probleme nicht möglich (0,4%) und für eine andere war sie aus sprachlichen Gründen nicht möglich (0,2%). 197 Bezugspersonen waren nicht verfügbar (38,6%) und bei 260 lagen „andere Gründe“ vor (51,0%)⁴⁰. Bei 13 Bezugspersonen gab es keine Angaben (2,5%, s. Anlage 14).

Wichtige Kennwerte für die Datenqualität sind die Vollzähligkeit und die Vollständigkeit der Daten. Die Vollzähligkeit wird mit den Rücklaufquoten beschrieben (s.o.), die Vollständigkeit mit den Anteilen valider Antworten innerhalb der Bögen. Der mittlere Anteil valider Antworten in den Fragebögen (ausgenommen der Abfrage zu den RZK) liegt bei den Beginnbögen bei 81,0% für Patienten und 89,4% für Bezugspersonen. Bei den Nachbefragungsbögen beträgt der Anteil 87,0% bzw. 89,8% für ambulant und stationär behandelte Patienten, bei Bezugspersonen beträgt er 93,1%.

Die Rücklaufquoten fallen etwas geringer aus im Vergleich zu anderen Indikationen des QS-Reha[®]-Verfahrens. Das kann damit zusammenhängen, dass die Einrichtungen auch in der Pilotphase keine Mindestmengenvorgabe hatten und keine Nachfassaktionen durchgeführt wurden. In diesem Licht betrachtet, sind die erzielten Rücklaufquoten bei den Patienten nicht schlecht. Anders verhält es sich mit der Rücklaufquote der Bezugspersonen, die bemerkenswert niedrig ist. Dazu, inwieweit Nachfassaktionen den Rücklauf bei ihnen erhöhen, liegen bisher keine Erfahrungen vor.

Die Vollständigkeit der Daten, also der Umfang, in dem die Fragen im Fragebogen beantwortet wurden, kann als überwiegend hoch betrachtet werden. Das spricht dafür, dass die Patienten und Bezugspersonen die Fragebögen als sinnvolle Instrumente der Qualitätssicherung annehmen.

Die nachfolgenden Auswertungen beziehen sich auf die Angaben der 164 stationär behandelten Patienten, für die der vollständige Datensatz bestehend aus Behandlerbogen, Beginn- und Nachbefragungsbogen vorliegt.

6.4.2 Item-Analyse

Die Item-Analyse bezieht sich auf die Items zur Messung der Prozessqualität und der Patientenzufriedenheit. Die Messung der Ergebnisqualität ist methodisch aufwändiger und wird daher gesondert betrachtet.

Die *Item-Schwierigkeit*⁴¹ lag bei zahlreichen Items unter 0,2 bzw. über 0,8, was eine eingeschränkte Differenzierungsfähigkeit bedeutet (s. Anlage 15.1 bis 15.3). Auf Grund extremer Item-Schwierigkeiten wurde jedoch kein Item gestrichen oder umformuliert.

Die *Trennschärfe* der Items war überwiegend hoch ($TS \geq 0,6$). Geringe Trennschärfen wurden damit erklärt, dass die Item-Blöcke Items unterschiedlicher Faktoren enthalten und damit inhomogen sind.

Mit *explorativen Faktorenanalysen* werden die Item-Gruppierungen, die nach dem Pretest faktorenanalytisch begründet geändert wurden, erneut überprüft. Wenn es sich bei den im Pretest gefundenen Faktoren um substantielle Entitäten handelt, dann sollten sie sich auch in einer weiteren unabhängigen Stichprobe replizieren lassen. Die Faktorenanalysen wurden nur für die Angaben in den Nachbefragungsbögen stationär behandelter Patienten (PNS) durchgeführt, weil die Fallzahlen für ambulant behandelte Patienten und Bezugspersonen zu gering waren. Die Analysen wurden getrennt für Reporting- und Rating-Fragen (Prozessqualität und Patientenzufriedenheit) sowie weiter getrennt für spezifische und globale Fragen gerechnet.

⁴⁰ Bisher liegen keine Erkenntnisse dazu vor, welche Gründe sich im Einzelnen hinter diesem Label „andere Gründe“ verbergen. Für eine differenzierte Betrachtung müsste in der Routine eine differenzierte Abfrage erfolgen, die entsprechende Vorarbeiten erforderlich macht.

⁴¹ Item-Schwierigkeit: Mittelwert aller Antworten eines Items, bei dem die unterste Antwortkategorie den Wert 0 und die oberste den Wert 1 hat. Items mit extrem niedriger bzw. extrem hoher Item-Schwierigkeit ($IS < 0,2$ bzw. $IS > 0,8$) sind für Messinstrumente ungeeignet.

Die durchgeführten Faktorenanalysen lieferten, wie im Pretest, überwiegend gut interpretierbare Faktoren, hohe Varianzaufklärungen (VA) sowie überwiegend akzeptable Kaiser-Meyer-Olkin-Kriterium (KMO) Werte bei hohen Fallzahlen (N)⁴² (s. Anlage 16):

1. Faktorenanalyse aller Reporting-Fragen (Prozessqualität)

Für die Reporting-Fragen zur Prozessqualität wurde eine 5-Faktoren-Lösung ermittelt. Die Faktoren waren inhaltlich gut interpretierbar. Die Varianzaufklärung (VA) betrug 67,5%, das Kaiser-Meyer-Olkin-Kriterium (KMO) betrug 0,720 bei einer Fallzahl N=82. Die ermittelten Faktoren sind:

Faktor 1: Medizinische Informationen mit Patient besprechen	($\alpha=0,77$)
Faktor 2: Informationen für die Zeit nach der Reha geben	($\alpha=0,74$)
Faktor 3: Allgemeine Informationen zu Rehabilitations-Beginn geben	($\alpha=0,69^*$)
Faktor 4: Reha-Grundlagen erfragen und vermitteln	($\alpha=0,46$)
Faktor 5: Therapieplanänderungen	($\alpha=0,36^*$)

Die Faktoren der Reporting-Fragen im Pretests wurden mit der 2. Stichprobe der Pilotphase exakt repliziert. Die Ergebnisse beider Faktorenanalysen unterscheiden sich nur in der Reihenfolge des 3. und des 4. Faktors, sowie in der Reihenfolge der Items innerhalb der Faktoren, was für die inhaltliche Interpretation ohne Belang ist.

Die interne Konsistenz der Faktoren, bestimmt mit Cronbachs Alpha (α) und damit deren Reliabilität, lag bei beiden Erhebungen überwiegend zwischen $0,67 \leq \alpha \leq 0,83$, was als akzeptabel bis gut zu bewerten ist. Einzig die Therapieplanänderungen hoben sich in beiden Erhebungen davon mit geringen Werten ab. Sie unterschieden sich von den anderen Reporting-Faktoren dadurch, dass sie mit nur zwei Items erfasst wurden, bei denen teilweise viele auswertbare Angaben fehlten (52% der Antworten in der Pilotphase waren: „es gab keine Änderungen“). Beim Faktor: „Reha-Grundlagen erfragen und vermitteln“ schwanken die α -Werte zwischen 0,67 und 0,46. Diese Befunde waren messtechnisch nicht optimal⁴³. Die inhaltliche Zusammengehörigkeit der Items sprach jedoch für eine gemeinsame Darstellung in der Berichterstattung der Qualitätsergebnisse.

Die im Pretests gefundenen Faktoren wurden zunächst als vorläufige Qualitätsunterdimensionen deklariert. Mit der Replikation der Faktoren in der Pilotphase werden sie als gültige Qualitätsunterdimensionen zur Bestimmung der Prozessqualität für das Geriatriemodul im QS-Reha[®]-Verfahren festgelegt.

2. Faktorenanalyse der spezifischen Rating-Fragen (Patientenzufriedenheit), ausgenommen Hotel-Fragen

Für die spezifischen Rating-Fragen zur Patientenzufriedenheit (ausgenommen Hotel-Fragen) wurde eine 5-Faktoren-Lösung ermittelt. Die Faktoren waren inhaltlich gut interpretierbar. Hier betrug die Varianzaufklärung VA=71,3%, das Kaiser-Meyer-Olkin-Kriterium (KMO) betrug 0,632 bei einer Fallzahl N=75. Die ermittelten Faktoren sind:

Faktor 1: Respekt des Klinikpersonals	($\alpha=0,92$)
Faktor 2: Zuverlässigkeit / Zugewandtheit des Klinikpersonals	($\alpha=0,82$)
Faktor 3: Diskretion des Klinikpersonals	($\alpha=0,66$)
Faktor 4: Belastung durch die Therapien	($\alpha=0,44^*$)
Faktor 5: Leistungen der Therapeuten	($\alpha=0,65$)

⁴² In allen Faktorenanalysen waren weitaus mehr als die doppelte Anzahl von Fällen als Items in den Faktorenlösungen zu finden.

⁴³ Diese Befunde beschreiben die Fallebene, im Rahmen der Qualitätssicherung ist die Einrichtungsebene von Relevanz, bei der die Reliabilität mit zunehmender Fallzahl ansteigt. Dessen ungeachtet ist die Steigerung der internen Konsistenz der Qualitätsunterdimensionen zu begrüßen, indem zusätzliche homogene Items aufgenommen werden.

*) maximales Cronbachs Alpha bei Streichung mindestens eines inhomogenen Items

‡) maximales Cronbachs Alpha bei Hinzunahme mindestens eines homogenen Items

***) keine Berechnung möglich, weil nur ein Item vorliegt

Auch bei den spezifischen Rating-Fragen (außer Hotel-Fragen) konnten in der Pilotphase die Faktoren, wie sie im Pretest gefunden wurden, weitgehend repliziert werden. Jeweils die ersten fünf Faktoren lassen sich im Pretest und in der Pilotphase finden. Faktor 6 im Pretest wurde durch das Item: „Ich habe die Erläuterungen des Klinikpersonals verstanden.“ gebildet. Dieses Item fällt in der Faktorenanalyse der 2. Stichprobe dem Faktor 5 „Leistungen der Therapeuten“ mit hoher Faktorladung zu (FL=0,76). U.a. auch deshalb wurde der Faktor von „Aktivitäten der Therapeuten“ in „Leistungen der Therapeuten“ umbenannt.

Die Änderung der Bezeichnung von Faktor 4 von „Last/Anstrengung der Therapien“ in „Belastung durch die Therapien“ stellt eine sprachliche Präzisierung und Vereinfachung der Faktorenbezeichnung dar, sie bedeutet keine inhaltliche Änderung.

Innerhalb der Faktoren beider Erhebungen gab es geringfügige Unterschiede, die zumeist darin bestanden, dass insgesamt gering ladende Items zwischen den Faktoren fluktuierten. Im Pretest fand sich z.B. das Item: „Es wurden zu wenige Therapien am Tag durchgeführt“ unter Faktor 3 „Aktivitäten der Therapeuten“ an letzter Stelle mit einer geringen höchsten Faktorladung (FL=-0,54). In der Pilotphase war es das letzte Item im Faktor 1 „Respekt des Klinikpersonals“ mit einer sehr geringen höchsten Faktorladung (FL=-0,46). Wegen deren Relevanz für die Qualität der Versorgung wurden solche Items nicht gestrichen, sondern den inhaltlich entsprechenden Qualitätsunterdimensionen zugeordnet (s. Anlage 16).

Die Reliabilität der Faktoren (bestimmt mit Cronbachs Alpha) stellte sich in beiden Erhebungen als überwiegend akzeptabel bis exzellent dar. Nur in der Pilotphase fiel sie bei Faktor 4 „Belastung durch die Therapien“ ab. Im Pretest erreichte sie noch einen Wert von $\alpha=0,64$. Der Hintergrund für die indifferenten Befunde war noch unklar⁴⁴. Die inhaltliche Kohärenz der Items aus Faktor 4 sprach jedoch für eine gemeinsame Darstellung in den Qualitätsergebnissen.

Die in der Pilotphase replizierten Faktoren wurden als gültige Qualitätsunterdimensionen der Patientenzufriedenheit für das Geriatriemodul im QS-Reha[®]-Verfahren festgelegt.

3. Faktorenanalyse der Hotel-Rating-Fragen (Patientenzufriedenheit)

Für die Hotel-Rating-Fragen zur Patientenzufriedenheit wurde eine 3-Faktor-Lösung ermittelt. Die Faktoren waren inhaltlich gut interpretierbar. In diesem Fall betrug die Varianzaufklärung VA=70,1%, das Kaiser-Meyer-Olkin-Kriterium (KMO) betrug 0,498 bei einer Fallzahl N=123. Die ermittelten Faktoren sind:

Faktor 1: Gebäude	($\alpha=0,25$)
Faktor 2: Patientenzimmer	($\alpha=0,24$)
Faktor 3: Essen	($\alpha=**$)

Die empirische Sachlage zeigte im Pretest und in der Pilotphase ein divergentes Bild. Die identische Item-Gruppe lieferte im Pretest eine 1-Faktorenlösung mit einem deutlich hohen KMO-Wert von 0,745 bei einer Fallzahl von N=60 und einer geringen Varianzaufklärung (50,7%). In der Pilotphase lieferte die Faktorenanalyse eine inhaltlich gut interpretierbare 3-Faktoren-Lösung mit einer 70%-igen Varianzaufklärung, aber geringem KMO-Wert (0,498) und höherer Fallzahl (N=123).

Zudem lag Cronbachs Alpha im Pretest für den einen Faktor bei $\alpha=0,71$, in der Pilotphase erreicht Alpha bei keinem Faktor $\alpha > 0,3$. Hier lässt sich nicht entscheiden, welche der divergenten Ergebnisse die tatsächlichen Zusammenhänge darstellen. Daher bleibt zu beobachten, wie sich die Faktorenstruktur sowie die Interne Konsistenz der Faktoren in den Daten der Routine darstellen, um dann auf dieser Grundlage ggf. Änderungen zu empfehlen.

Unabhängig von der Dimensionalität der Hotel-Rating-Items wurde der Summenscore über alle fünf Hotel-Items als gültige Qualitätsunterdimension für das Geriatriemodul im QS-Reha[®]-Verfahren festgelegt. Denn auch wenn diese Items einen 3-dimensionalen Raum aufspannen, stellt der relativ ausgeglichen gewichteten Index einen vergleichbaren Wert wie der Summenscore dar, in den Aspekte des Gebäudes und des Patientenzimmers mit gleichem Gewicht und die Zufriedenheit mit dem Essen

⁴⁴ Es bleibt zu beobachten, wie sich die Interne Konsistenz des Faktors in den Daten der Routine darstellt, um dann auf dieser Grundlage ggf. Änderungen zu empfehlen.

*) maximales Cronbachs Alpha bei Streichung mindestens eines inhomogenen Items

‡) maximales Cronbachs Alpha bei Hinzunahme mindestens eines homogenen Items

***) keine Berechnung möglich, weil nur ein Item vorliegt

mit halb so hohem Gewicht eingehen. Die Ergebnisse eines eindimensionalen Summenscores und eines dreidimensionalen Index wären also vergleichbar.

4. Faktorenanalyse der globalen Rating-Fragen (Patientenzufriedenheit, Ergebnisqualität)

Für die globalen Rating-Fragen zur Patientenzufriedenheit und Ergebnisqualität wurde eine 3-Faktoren-Lösung ermittelt. Die Faktoren waren inhaltlich gut interpretierbar. In diesem Fall betrug die Varianzaufklärung $VA=67,7\%$, das Kaiser-Meyer-Olkin-Kriterium (KMO) betrug 0,799 bei einer Fallzahl $N=140$. Die ermittelten Faktoren sind:

Faktor 1: Globalbewertung der Rehabilitations-Einrichtung (Z)	($\alpha=0,89^*$)
Faktor 2: Rehabilitations-Zielerreichung (EQ)	($\alpha=0,79^*$)
Faktor 3: Zufriedenheit mit dem Sozialdienst (Z)	($\alpha=0,18^*$)

Die Faktorenanalysen der Globalbewertungen kamen in beiden Erhebungen zu annähernd gleichen Ergebnissen. Die ersten beiden Faktoren fielen jeweils nahezu identisch aus und beschrieben als globales Zufriedenheitsurteil die „Globalbewertung der Rehabilitations-Einrichtung“ sowie die „Rehabilitations-Zielerreichung“ als Globalbewertung der Ergebnisqualität. Faktor 3 im Pretest, die „Bewertung sekundärer Dienstleistungen“ als Teilaspekt der Patientenzufriedenheit konkretisierte sich in der Pilotphase zur „Zufriedenheit mit dem Sozialdienst“.

Geringfügige Abweichungen in den Ergebnissen sind zu erwarten gewesen, weil das Item: „Die Hilfsmittel sind richtig angepasst, so dass ich sie gut benutzen kann“ (1. Item aus Faktor 3 im Pretest) nach dem Pretest gestrichen wurde, weil das Anpassen von Hilfsmitteln keine Leistung der Rehabilitations-Einrichtungen ist, sondern der Lieferanten, die die Patienten mit Hilfsmitteln versorgen. Die Reliabilität der ersten beiden stabilen Faktoren erreichte gute bis annähernd exzellente Werte ($0,79 \leq \alpha \leq 0,89$). Der um das Item: „Der Übergang von der Rehabilitation nach Hause bzw. in die jetzige Einrichtung war reibungslos“ ergänzte Faktor 3 erzielt keine befriedigende Reliabilität. Die inhaltliche Kohärenz der Items sprach jedoch für eine gemeinsame Darstellung in den Qualitätsergebnissen.

Die in der Pilotphase replizierten ersten beiden Faktoren sowie der um ein Item ergänzte Faktor 3 wurden als gültige Qualitätsunterdimensionen der Patientenzufriedenheit und der Ergebnisqualität für das Geriatriemodul im QS-Reha[®]-Verfahren festgelegt.

Zusammenfassend kann hier festgestellt werden, dass die im Pretest gefundenen Faktoren mit hoher Übereinstimmung in der Pilotphase repliziert werden konnten. Damit zeigt sich, dass mit den verwendeten Items substanziale und stabile Entitäten der Prozessqualität und der Patientenzufriedenheit der stationären geriatrischen Rehabilitation erfasst werden können. Zudem belegen die erhobenen Daten, dass die auf dieser Grundlage definierten Qualitätsunterdimensionen zumeist hinreichend akzeptabel bis gut reliabel gemessen werden können.

Die Teile der Patientennachbefragungsbögen, mit denen die Prozessqualität und die Patientenzufriedenheit abgefragt werden, werden entsprechende den Befunden aus der Pilotphase erneut umgestellt. Die Items werden danach gruppiert, dass sie eine maximal reliable Messung der gültigen Qualitätsunterdimensionen erlauben. Dazu werden vier Items anderen Qualitätsunterdimensionen als im Pretest-Fragebogen zugeordnet. Innerhalb der Qualitätsunterdimensionen werden die Items nach der Faktorenladung neu sortiert. Der Patientennachbefragungsbogen für ambulant behandelte Patienten wird in Anlehnung an die Faktorenanalyseergebnisse der stationären Patienten entsprechend geändert. An den Items selbst werden keine Änderungen vorgenommen, weder gibt es Umformulierungen noch Streichungen.

Ergebnisqualität

6.4.3 Auswahl der Rehabilitations-Zielkategorien

In der Pilotphase konnten aus den 33 Rehabilitations-Zielkategorien pro Patient mindestens drei bis maximal zehn RZK aus mindestens drei ICF-Domänen ausgewählt werden. Insgesamt wurden in der Pilotphase 3.089 Rehabilitations-Zielkategorien für 463 Patienten ausgewählt. Damit wurden im Mittel

pro Patient 6,68 Rehabilitations-Zielkategorien (Median=6) mit einer Streuung⁴⁵ von SD=2,04 (Interquartile Range: 5 bis 8) ausgewählt (s. Abbildung 8).

In den 12 teilnehmenden Rehabilitations-Einrichtungen wurden durchschnittlich im Minimum 5,24 und im Maximum 11 RZK ausgewählt. Das Einrichtungsmittel lag bei 7,64 RZK (Median=7,79) mit einer Streuung von SD=1,89 (Interquartile Range: 5,87 bis 8,82).

Abbildung 8: Häufigkeitsverteilung der Anzahl ausgewählter Rehabilitations-Zielkategorien der Ebene II

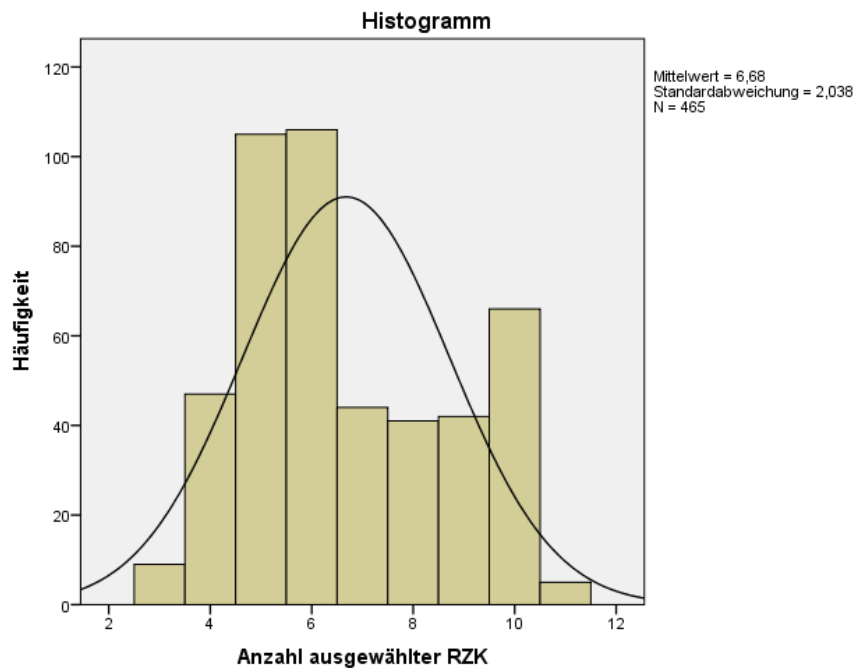
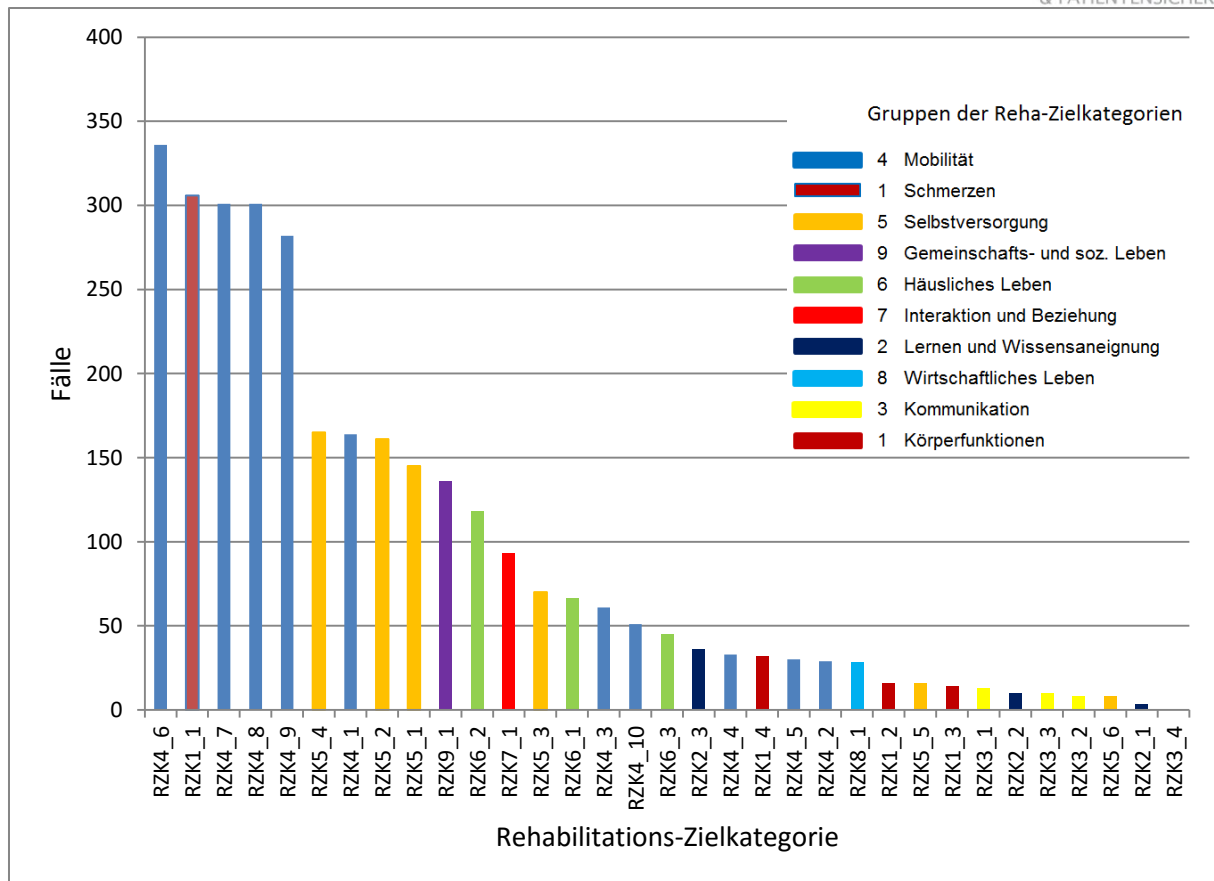


Abbildung 9: Häufigkeitsverteilung der ausgewählten RZK absteigend nach Häufigkeit sortiert

⁴⁵ Die Standardabweichung oder kurz Streuung (englisch: standard deviation; Abkürzung: SD) beschreibt die Variation (Dispersion) von Messwerten einer Verteilung. Zur Berechnung dieses Verteilungsparameters sind Messwerte auf Intervallskalenniveau erforderlich. Für Messwerte auf Ordinalskalenniveau wird dazu häufig der Interquartile Range angegeben.



Die 33 RZK wurden unterschiedlich häufig ausgewählt (s. Abbildung 9). Die Häufigkeitsverteilung ähnelte sehr der aus dem Pretest (vgl. Abbildung 4) hinsichtlich der Form der Verteilung und der Position der einzelnen RZK. Im Vergleich der Pretest- und der Pilotphasenergebnisse fielen die Positionen der einzelnen RZK vielfach identisch oder um eine oder zwei Positionen vertauscht aus.

In beiden Erhebungen dominierten RZK der „Mobilität“ und „Schmerzen“ die Auswahl. In der Gruppe der am häufigsten ausgewählten RZK stand RZK 4.6 „Gangssicherheit“ beide Male an der Spitze (bei 72,6% aller Patienten der Pilotphase) und weiter die RZK 4.7 bis 4.9, „Gehen im Innenbereich“, „Treppensteigen“ und „Gehen im Außenbereich“ (65,0%, 65,0% und 60,9%).

An 2. Position der Häufigkeit kam RZK 1.1 „Schmerzen“ (66,1%).

Im mittleren Häufigkeitsbereich traten die RZK der Selbstversorgung 5.1 bis 5.4, „Sich waschen“, „Körperpflege“, „Toilettennutzung“ und „Sich kleiden“ auf (31,3%, 34,8%, 15,1% und 35,6%).

Dann folgten RZK des häuslichen Lebens 6.1 bis 6.3, „Mahlzeiten zubereiten“, „Hausarbeiten erledigen“ und „Beschaffen von Waren und Dienstleistungen“, die schon deutlich seltener ausgewählt wurden (14,3%, 25,5% und 9,7%).

Zu den am seltensten ausgewählten RZK gehörten die zum Lernen und der Wissensaneignung, RZK 2.1 bis 2.3, „Lesen“, „Problemlösen“ und „Konzentration“ (0,6%, 2,2% und 7,8%). Wie schon im Pretest wurde RZK 3.4 „Kommunikationsgeräte nutzen“ kein einziges Mal für einen Patienten ausgewählt.

Von den teilnehmenden Einrichtungen kamen keine Rückmeldungen zur neuen Auswahlmodalität, dass mindestens drei RZK aus drei ICF-Domänen auszuwählen sind. Dieser Umstand wird so interpretiert, dass diese Vorgabe bei der Falldokumentation kein Problem hervorgerufen hat.

6.4.4 Reliabilität der Ergebnisqualitäts-Scores

Bei der Ergebnisqualität wurden die Ergebnisse der Qualitätsunterdimensionen kurz als Ergebnisqualitäts-Scores bezeichnet.

Für die am häufigsten ausgewählten RZK 4.6 bis 4.9 „Gangssicherheit“, „Gehen im Innenbereich“, „Treppensteigen“ und „Gehen im Außenbereich“ und RZK 1.1 „Schmerzen“ (s. Abbildung 9) wurden die Ergebnisqualitäts-Scores auf Basis der Deltawerte der Vorher-Nachher-Messungen berechnet (s. Kapitel 5.4.6 Berechnung der Ergebnisqualitäts-Scores). Diese fünf RZK deckten 49,4% aller RZK-Auswahlen ab. Die Auswertungen wurden hier exemplarisch an diesen fünf Qualitätsunterdimensionen durchgeführt, weil die anderen RZK zu schwach mit Fällen besetzt waren, als dass auf deren Grundlage statistisch abgesicherte Ergebnisse gewonnen werden konnten. Die geringen Zellenbesetzungen waren in der Auswahl der RZK⁴⁶ und in den geringen Rücklaufzahlen begründet, die für die Vorher-Nachher-Messungen das Limit setzten. In der Routine wird dieser Umstand wahrscheinlich nicht zum Tragen kommen, weil alle rund 190 Rehabilitations-Einrichtungen⁴⁷ Patienten über einen deutlich längeren Zeitraum mit einer definierten Mindestfallzahl akquirieren werden, so dass selbst für selten ausgewählte RZK und bei geringem Rücklauf genügend Fälle zur Berechnung der Ergebnisqualitäts-Scores vorliegen werden.

Der Ergebnisqualitäts-Scores der RZK „Gangssicherheit“ berechnete sich aus den fünf Items RZK 4.6 a bis e. Hier wurde ein Cronbachs Alpha von $\alpha=0,851$ bei 110 Messwertpaaren erzielt. Durch Löschen von einzelnen Items konnte keine Verbesserung der internen Konsistenz erreicht werden.

Der Ergebnisqualitäts-Scores der RZK „Gehen im Innenbereich“ errechnete sich aus allen 10 Items RZK 4.7 1a bis 2e. Auf dieser Grundlage wurde ein Cronbachs Alpha von $\alpha=0,767$ bei 99 Messwertpaaren ermittelt. Durch Löschen von einzelnen Items konnte keine Verbesserung der internen Konsistenz erreicht werden.

Der Ergebnisqualitäts-Scores der RZK „Treppensteigen“ wurde aus allen 16 Items RZK 4.8 1a bis 4d berechnet. Auf dieser Grundlage wurde ein Cronbachs Alpha von $\alpha=0,870$ bei 78 Messwertpaaren ermittelt. Auch hier konnte durch Löschen von einzelnen Items keine Verbesserung der internen Konsistenz erreicht werden.

Der Ergebnisqualitäts-Scores der RZK „Gehen im Außenbereich“ wurde aus den 16 Items RZK 4.9 1a bis 4d berechnet. Auf dieser Grundlage wurde ein Cronbachs Alpha von $\alpha=0,800$ bei 46 Messwertpaaren ermittelt. Durch Löschen von einzelnen Items konnte auch bei diesem Score keine Verbesserung der internen Konsistenz erreicht werden.

Zur Messung des Ergebnisqualitäts-Scores der RZK „Schmerzen“ erfolgten vier Abfragen:

- a) Schmerzmitteleinnahme: ja/nein
- b) Stärke der Schmerzen in Ruhe?
- c) Stärke der Schmerzen bei Belastung?
- d) Stärke der Schmerzen bei Bewegung?

Als Antwortformat für die Fragen b bis d wird eine 11-stufige Antwortskala angeboten, deren Stufen von 0 (linker Pol der Skala) bis 10 (rechter Pol der Skala) numerisch bezeichnet sind und die Pole selbst zusätzlich verbal gekennzeichnet sind mit „keine Schmerzen“ (0) und „stärkste vorstellbare Schmerzen“ (10).

Bei der Schmerzmitteleinnahme gibt es durch die Wiederholungsmessung der dichotomen ja-nein-Abfrage vier Outcome-Konstellationen. Die Häufigkeiten der Zellenbesetzungen zeigt Tabelle 7. In der Pilotphase haben von den 307 Patienten, für die die Verbesserung ihrer Schmerzen bzw. deren Behandlung zu t0 ein Therapieziel darstellte, zu t2 noch 91 von ihnen dazu Auskunft gegeben⁴⁸. 62 davon nahmen zu t0 und zu t2 Schmerzmittel ein, 15 Patienten zu beiden Zeitpunkten keine. Acht

⁴⁶ In inhomogenen Indikationsbereichen wie der geriatrischen Rehabilitation mit unterschiedlichen Krankheitsbildern und unterschiedlichen Therapieansätzen sind die Fallzahlen in den differenzierten Kategorien von der Verteilung der Fälle in der Population abhängig.

⁴⁷ Im 2. Erhebungszyklus des QS-Reha[®]-Verfahrens von 2015 bis 2017 haben 150 stationäre und 44 ambulante geriatrische Rehabilitations-Einrichtungen an der Erhebung der Strukturqualität teilgenommen.

⁴⁸ Das entspricht einer Rücklaufquote von 29,6%. Von 107 Patienten lagen t2-Bögen vor, die zu t0 das Schmerzmodul ausgefüllt hatten. Von denen haben 91 Patienten (85,0%) Angaben zur Schmerzmitteleinnahme gemacht.

Patienten nahmen vorher aber nachher keine Schmerzmittel ein und bei 6 Patienten lag der umgekehrte Fall vor, die vorher keine aber nachher Schmerzmittel einnahmen.

Den Ergebnisqualitäts-Score der RZK „Schmerzen“ erhält man, indem für die einzelnen Items b bis d die Differenzen zwischen den t0- und t2-Messungen bestimmt und anschließend die Differenzen z-transformiert werden und über die z-Werte der drei Items der Mittelwert errechnet wird. Das resultierende Cronbachs Alpha der drei Items, die in den Ergebnisqualitäts-Score einfließen, unterscheidet sich, je nach dem, welche Schmerzmitteleinnahme zu beiden Messzeitpunkten vorlag (s. Tabelle 8).

Tabelle 7: Häufigkeitsverteilung der Schmerzmitteleinnahme zu Rehabilitations-Beginn (t0) und zwei Wochen nach Rehabilitations-Ende (t2)

		t2				Σ
		kein FB vorhanden	FB ohne Schmerz	nein	ja	
t0	FB ohne Schmerz	21	0	0	0	21
	nein	21	7	15	6	49
	ja	108	9	8	62	187
	Σ	150	16	23	68	257

Wenn die Schmerzmitteleinnahme bei der Berechnung der Scores nicht berücksichtigt wird (Zeile 1), resultiert ein Cronbachs Alpha von $\alpha=0,877$ bei 88 Messwertpaaren. Wenn der Score nur dann berechnet wird, wenn die Schmerzmitteleinnahme zu t2 genauso erfolgte wie zu t0 (Zeile 2), dann steigt Cronbachs Alpha auf $\alpha=0,884$ bei 66 Messwertpaaren. Wenn vorher und nachher Schmerzmittel eingenommen werden müssen, um den Score berechnen zu können (Zeile 3), dann liegt Cronbachs Alpha bei $\alpha=0,883$ bei 54 Messwertpaaren. Für Patienten, die vorher und nachher keine Schmerzmittel eingenommen haben (Zeile 4), wird die Veränderung der Schmerzen am reliabelsten gemessen. Für sie wird ein Cronbachs Alpha von $\alpha=0,899$ bei dann nur noch 12 Messwertpaaren errechnet.

Tabelle 8: Cronbachs Alpha für den Ergebnisqualitäts-Score der RZK „Schmerzen“ in Abhängigkeit von der Schmerzmitteleinnahme zu Rehabilitations-Beginn (t0) und zwei Wochen nach Rehabilitations-Ende (t2)

Schmerzmittel-einnahme		Cronbachs Alpha	Fallzahl n
zu t0	zu t2		
-	-	0,877	88
wie t2	wie t0	0,884	66
ja	ja	0,883	54
nein	nein	0,899	12

Da bei Nicht-Berücksichtigung der Schmerzmitteleinnahme noch ein sehr guter Wert für die interne Konsistenz der Veränderungsmessung mit den drei Items RZK 1.1 b bis d erzielt wird und der Verlust an Ergebnisqualitäts-Scores damit minimal ist, wird empfohlen den Score unabhängig von der Schmerzmitteleinnahme (Item a) zu bilden. Bei der Bestimmung geeigneter Risikoadjustierungsmodelle des Scores sollten dann beide Angaben zur Schmerzmitteleinnahme zu t0 und zu t2 in das Startset der Prädiktoren (s. Anlage 13) aufgenommen werden.

Zusammenfassend lässt sich feststellen, dass für die fünf am häufigsten ausgewählten Rehabilitations-Zielkategorien, die 49,4% aller RZK-Auswahlen ausmachen, durchgängig gute Werte für die interne Konsistenz der Ergebnisqualitäts-Scores gefunden wurden, bei ausreichend hohen Fallzahlen für die Berechnung der diversen Faktorenanalysen. Dabei ist zu unterstreichen, dass die hier betrachteten Scores auf Differenzwerten beruhen, deren Grundlage Befragungsdaten von geriatrischen Patienten sind, die zu Beginn der Rehabilitation und zwei Wochen nach Ende der Rehabilitation erhoben wurden. D.h. mit dem zugrunde gelegten Konzept für die Messung der Ergebnisqualität und den entsprechenden Operationalisierungen werden wichtige Qualitätsunterdimensionen der Ergebnisqualität mit hoher Reliabilität gemessen. Das belegen die Befunde aus dem Pretest und der Pilotphase. Zudem zeigt sich, dass die durch die Pretest-Ergebnisse motivierten Änderungen an der Abfrage der RZK die interne Konsistenz der damit erhaltenen Ergebnisse befördert hat.

6.4.5 Validität der Ergebnisqualitäts-Scores

Bei der Konstruktion der Patienten- und Bezugspersonennachbefragungsbögen wurden für die spätere Prüfung der kriteriumsbezogenen Validität⁴⁹ globale Ergebnis-Items als Außenkriterien aufgenommen. Sie erfassen

- a) das Erreichen der persönlichen Rehabilitations-Ziele,
- b) das Erreichen des vereinbarten individuellen Rehabilitations-Ziels⁵⁰ sowie
- c) die aktuelle Gesamtlebenszufriedenheit.

Diese drei Fragen wurden zu t2 abgefragt und werden bei der Analyse als Außenkriterien verwendet. Betrachtet wird der korrelative Zusammenhang zwischen den Ergebnisqualitäts-Scores für alle Aggregationsstufen bis zum Gesamt-Score der fünf am häufigsten ausgewählten RZK und den drei Außenkriterien. Je höher der korrelative Zusammenhang zwischen den RZK-Scores und den Außenkriterien ist, desto höher fällt die Kriteriumsvalidität aus.

Bei „Schmerzen“ zeigten sich in den Daten nur unbedeutende Korrelationen. Der höchste Wert fand sich bei Schmerzen in Ruhe (RZK 1.1 b) und dem Erreichen des individuellen Rehabilitations-Ziels mit $r=0,23$ ($p=0,03$, $n=92$).

Bei „Gangsisicherheit“ lag die höchste Korrelationen zwischen dem Gesamt-Score und dem Erreichen des individuellen Rehabilitations-Ziels mit $r=0,30$ ($p=0,001$, $n=112$).

Bei der RZK „Gehen im Innenbereich“ lagen die Korrelationen auf Itemebene häufig zwischen 0,3 und 0,4, maximal bei $r=0,46$. Zwischen dem Gesamt-Score und dem Erreichen des individuellen Rehabilitations-Ziels lag sie bei $r=0,38$ ($p<0,001$, $n=99$).

Auch beim „Treppensteigen“ lagen die Korrelationen auf Itemebene häufig zwischen 0,3 und 0,4, maximal bei $r=0,46$. Zwischen dem Gesamt-Score und dem Erreichen des individuellen Rehabilitations-Ziels lag sie bei $r=0,34$ ($p=0,001$, $n=91$).

Beim „Gehen im Außenbereich“ fiel auf, dass die Korrelationen mit zunehmender Schwierigkeit der Item-Blöcke abnehmen. Der Gesamt-Score für den 1. Block („15 Minuten am Stück gehen“) korrelierte mit der Einschätzung des Erreichens des individuellen Rehabilitations-Ziels mit $r=0,27$. Die entsprechende Korrelation mit dem Gesamt-Score für den 2. Block („30 Minuten am Stück gehen“) lag bei $r=0,23$. Beim 3. Block („eine Stunde am Stück gehen“) lag sie bei $r=0,11$ und beim 4. Block („über schwierige Wege gehen“) bei $r=0,07$. Entsprechend finden sich Korrelationen zwischen 0,3 und 0,4 nur im 1. Block. Der RZK-Gesamt-Score korrelierte mit dem Erreichen des individuellen Rehabilitations-Ziels mit $r=0,15$ (n.s., $n=90$).

⁴⁹ Kriteriumsvalidität liegt vor, wenn zwischen den Werten eines Scores und denen eines validen Außenkriteriums, z.B. einem Expertenurteil, substantielle Zusammenhänge festgestellt werden, s. Fisseni (2004)

⁵⁰ Das ist das individuelle Rehabilitations-Ziel, das zu Beginn der Behandlung mit dem Patienten vereinbart wurde und die Grundlage für die Auswahl der Rehabilitationsziel-Kategorien bildet.

Dieser systematische Befund ließ sich so interpretieren, dass die abgefragten Stufen des Gehens im Außenbereich für stationär behandelte Rehabilitations-Patienten außerhalb ihres Ziel- und Erfahrungsbereiches liegen und ihre Angaben daher nicht zur kriteriumsbezogenen Validierung der Operationalisierungen geeignet sind. Als ein bescheidener Beleg für die Konstruktvalidität, nämlich für den zunehmenden Schweregrad der abgefragten Aktivitäten, konnte der Befund vielleicht betrachtet werden. Die RZK wurde für die Routine unverändert übernommen, da davon ausgegangen wird, dass sie für ambulant behandelte Patienten korrekt adjustiert ist.

Die Ergebnisse der Pilotphase waren zwar in sich konsistent interpretierbar, fallen aber insgesamt gesehen sehr schwach aus. Die Ergebnisse aus dem Pretest fielen dagegen deutlicher aus. Dort fanden sich Korrelationen im Bereich deutlich über 0,4 und reichten maximal bis $r=0,89$ ($p<0,001$, $n=16$) zwischen „Gegenstände anheben und tragen“ (RZK 4.3) und der Patienteneinschätzung des Erreichens des persönlichen Rehabilitations-Ziels.

Insgesamt ließen sich Zusammenhänge zwischen den berechneten Ergebnisqualitäts-Scores und den globalen patientenseitigen Einschätzungen der Rehabilitations-Zielerreichungen in den Daten beider Erhebungen erkennen. In ihnen zeichnete sich eine Bestätigung der kriteriumsbezogenen Validität der Ergebnisqualitäts-Scores ab.

Die Außenkriterien sollten kritisch geprüft werden, inwieweit sie geeignet sind, kriteriumsbezogene Validität festzustellen. Hier ist zu berücksichtigen, dass für einen Patienten drei bis zehn RZK ausgewählt werden, so dass beim Überprüfen einzelner Ergebnisqualitäts-Scores immer Varianzanteile der anderen ausgewählten RZK im Gesamturteil des Patienten enthalten sind, die das individuelle Rehabilitations-Ziel mit abdecken und bei der Korrelation nicht differenziert werden. Auf Grund dieser Betrachtung ist klar, dass die hier berichteten Korrelationen die wahren Zusammenhänge unterschätzen. Ein Verfahren, das dieses methodische Problem in den Griff bekommt, steht derzeit nicht zur Verfügung.⁵¹

6.4.6 Risikoadjustierung der Ergebnisqualitäts-Scores

Im QS-Reha[®]-Verfahren werden die Ergebnis-Scores der Ergebnisqualität und der Patientenzufriedenheit risikoadjustiert als Differenz zwischen den beobachteten Werten (O für observed) und den erwarteten Werten (E für expected) $O - E$ numerisch und grafisch dargestellt.

Dabei wird für jeden Patienten sein Erwartungswert E auf Basis einer linearen Geradengleichung errechnet, die ein zuvor ermitteltes brauchbares Vorhersagemodell darstellt, bei dem für den betrachteten Ergebnis-Score solche Personenmerkmale optimal gewichtet einfließen, die einen signifikanten Einfluss auf diesen Score ausüben und von der Rehabilitations-Einrichtung nicht beeinflusst werden können. Solche Einflussfaktoren heißen Confounder und klassische Vertreter dafür sind das Lebensalter und das Geschlecht des Patienten.

Auf Grundlage der bei dem Patienten vorliegenden Ausprägungen der einbezogenen Risikofaktoren wäre gemäß dem ermittelten Modell der errechnete E-Wert für ihn zu erwarten gewesen. Diesem E-Wert wird der tatsächliche beobachtete O-Wert gegenübergestellt, indem die Differenz $O - E$ betrachtet wird. Wenn die Summe der $O - E$ -Differenzen einer Einrichtung positiv ausfällt, dann hat sie ein Ergebnis erzielt, das insgesamt besser ist als es für die einbezogenen Patienten und den bei ihnen vorliegenden Risikofaktoren zu erwarten gewesen wäre. Ist die Summe der Differenzen negativ, dann sind die beobachteten Ergebnisse insgesamt schlechter ausgefallen als sie zu erwarten gewesen wären.

Dass auf Grundlage der im Behandlerbogen erfassten Confounder tragfähige Risikoadjustierungsmodelle gefunden werden können, soll hier exemplarisch an den Ergebnisqualitäts-Scores der RZK „Schmerzen“ und „Gehen im Innenbereich“ vorgestellt werden.

Als Risikoadjustierungsmodell für den Ergebnisqualitäts-Score der RZK „Schmerzen“ wurde eine multiple lineare Regression, mit den beiden Prädiktoren: „Stürze innerhalb der letzten 6 Monate“ mit den Stufen ja (1) und nein (0) und dem Score der Barthel-Index-Subskala: Toilettenbenutzung mit den Stufen 0, 5, 10 und dem Gesamt-Score der RZK „Schmerzen“ als Kriterium, gefunden.

In der Linearkombination der multiplen linearen Regression gehen die Stürze-Angaben mit einem Beta-Gewicht von 0,35 ($p=0,02$) und der BI-Score der Toilettenbenutzung mit einem Beta-Gewicht von -0,31 ($p=0,03$) ein. Das Multiple R beträgt 0,52 und der Determinationskoeffizient $R^2=0,27$. Das Modell ist hoch signifikant.

⁵¹ Semipartialkorrelationen höherer Ordnung auf Basis unvollständiger Datenlagen wären erforderlich.

Die Ergebnisse besagen, dass die Veränderungen der Schmerzen nach dem Rehabilitations-Aufenthalt positiver ausfallen, wenn der Patient innerhalb der letzten 6 Monate vor Aufnahme in die Reha gestürzt war und je schlechter die Toilettennutzung zu Rehabilitations-Beginn bei ihm eingestuft wurde. Die Modellgüte kann mit dem berichteten R und R^2 als noch akzeptabel bewertet werden.

Wenn der Patient vor der Behandlung in der Rehabilitationseinrichtung einen Sturz, also ein Akutereignis erlebt hat, dann ist bei der Schmerzbehandlung mit einem deutlicheren Behandlungserfolg zu rechnen, schon wegen der Heilung der sturzbedingten Traumata. Zudem zeigt sich ein (negativer) Zusammenhang zwischen der Selbstständigkeit bei der Toilettenbenutzung zu Beginn der Rehabilitation und dem Erfolg der Schmerzbehandlung. Es wird vermutet, dass der Chronifizierungsgrad der Schmerzen in dieser Assoziation zum Ausdruck kommt, dass ein hoher Chronifizierungsgrad eher bei größerer Selbstständigkeit und ein geringerer Chronifizierungsgrad eher bei geringerer Selbstständigkeit der Toilettennutzung zu finden ist und mit abnehmendem Chronifizierungsgrad die Erfolge bei der Schmerzbehandlung zunehmen.

Eine Risikoadjustierung unter Verwendung dieser beiden Confounder erscheint nachvollziehbar, als damit der Vorteil von Einrichtungen aus den RZK-Ergebnissen herausgerechnet wird, der darin besteht, vermehrt Schmerzen bei Patienten mit einem Sturzereignis vor der Reha sowie Patienten mit geringer Selbstständigkeit bei der Toilettennutzung behandelt zu haben, was den Erfolg einer Schmerzbehandlung unabhängig von den Interventionen der Einrichtung bereits positiv fördert.

In das Modell zur Vorhersage der Ergebnisqualitäts-Scores der RZK „Gehen im Innenbereich“ fließen die Prädiktoren: Hilfsmittel (bisherige Nutzung ggf. vor dem aktuellen Akutereignis) mit den Stufen ja (1) und nein (0), der aus Körpergewicht und Körpergröße errechnete Body-Mass-Index und der Barthel-Index-Score für An- und Auskleiden mit den Stufen 0, 5, 10 ein.

Der Prädiktor Hilfsmittelnutzung geht mit einem Beta-Gewicht von -0,27 ($p < 0,01$), der BMI mit -0,27 ($p < 0,01$) und der BI-Score An- und Auskleiden mit 0,29 ($p < 0,01$) ein. Das multiple R beträgt 0,49 und der Determinationskoeffizient $R^2 = 0,24$. Das Modell ist hoch signifikant.

Die Ergebnisse besagen, dass die Veränderungen beim Gehen im Innenbereich positiver ausfallen, wenn der Patient zu Rehabilitations-Beginn bzw. vor dem Akutereignis keine Hilfsmittel genutzt hat, je niedriger sein BMI zu Rehabilitations-Beginn war und je besser er sich zu Rehabilitations-Beginn an- und auskleiden konnte. Die Modellgüte kann mit dem berichteten R und R^2 als noch akzeptabel bewertet werden.

Damit wurde beispielhaft gezeigt, dass auf Grundlage der im Behandlerbogen erfassten Confounder tragfähige Risikoadjustierungsmodelle für Ergebnisqualitäts-Scores der RZK gebildet werden können. Beide Modelle könnten zur Risikoadjustierung der Ergebnisse der Pilotphase verwendet werden. Die Modelle, die in der Routine angewendet werden, müssen auf Grundlage aller in der Routine erhobenen Patientendaten aller einbezogenen Einrichtungen erstellt werden. Damit wird die Gültigkeit der gefundenen Modelle für die Routine sichergestellt.

Nachdem alle Risikoadjustierungsmodelle für alle Qualitätsunterdimensionen der Ergebnisqualität und der Patientenzufriedenheit bestimmt wurden, kann der Behandlerbogen auf die Confounder reduziert werden, die in den Modellen verwendet werden.

6.4.7 Prüfung alternativer Abfragen bei Ausfall von Patienten

Von den drei entwickelten und im Pretest überprüften Ansätzen (s. Kapitel 5.4.10) wurde die alternative individualisierte Ergebnisqualitätsmessung durch Befragung von Bezugspersonen in der Pilotphase durchgeführt, ebenso deren globale Ergebnisqualitätsabfrage zur Zielerreichung des individuellen Rehabilitations-Ziels. Die Behandler einschätzungen zum Erfolg der Rehabilitations-Maßnahme sollten in der Pilotphase erneut überprüft werden.

6.4.7.1 Behandler einschätzungen als Alternative für globale Ergebnisbewertungen der Patienten

Am Ende der Rehabilitation wurden von den Behandlern im Behandlerbogen Einschätzungen des Rehabilitations-Erfolges erfragt (s. Teil A4 - Rehabilitations-Ende). Es wurde zum einen das Ausmaß erfragt, in dem der Patient nach Einschätzung der Behandler von der Rehabilitation profitiert hat. Zum anderen sollte die Wahrscheinlichkeit von den Behandlern angegeben werden, mit der der Patient sein individuelles Rehabilitations-Ziel nach der Entlassung realisieren können wird.

Bezogen auf die erste Frage: „Hat der Patient Ihrer Einschätzung nach von der Rehabilitation profitiert?“ bestand zwischen den behandlerseitigen Einschätzungen und den entsprechenden Selbstauskünften der Patienten nur ein geringer korrelativer Zusammenhang, $r = 0,20$ ($p = 0,01$). Bei der

zweiten Frage: „Wird der Patient Ihrer Einschätzung nach sein individuelles Rehabilitations-Ziel nach Entlassung realisieren können?“ lag der korrelative Zusammenhang bei $r=0,31$ ($p < 0,001$). Zudem unterscheiden sich die Mittelwerte zwischen Behandler- und Patienteneinschätzung hoch signifikant. Die Behandler unterschätzten den Rehabilitations-Profit um 0,4 Skaleneinheiten⁵² hoch signifikant ($p < 0,001$). Gleichzeitig überschätzten sie die Wahrscheinlichkeit für das Erreichen des individuellen Rehabilitations-Ziels extrem um 0,75 Skaleneinheiten⁵³ hoch signifikant ($p < 0,001$).

Diese Befunde stimmten mit denen des Pretests überein und bestätigten, dass die behandlerseitigen Einschätzungen sich nicht dazu eignen, als Ersatz für fehlende Patientenangaben herangezogen zu werden.

Der Steuerungskreis empfahl, beide Fragen weiterhin im Behandlerbogen zu belassen und dass die Entwicklung der Antworten weiter beobachtet und ausgewertet werden.

6.4.8 Feedback der Befragten zur Befragung

Die Fragebögen wurden überwiegend als inhaltlich angemessen bewertet (Patienten: 62,1 bis 71,5%, Bezugspersonen: 54,5%). Für den Großteil der Befragten waren die Fragebögen nicht zu lang. 90,7% der Patienten und 90,9% der Bezugspersonen fanden den Beginnbogen nicht zu lang und 77,0% der stationär und 79,3% der ambulant behandelten Patienten fanden den Nachbefragungsbogen nicht zu lang. Die Bearbeitungszeiten waren in den meisten Fällen akzeptabel. Für den Beginnbogen benötigten Patienten bzw. Bezugspersonen im Mittel 18 bzw. 17 Minuten und zum Ausfüllen des Nachbefragungsbogens 32 bzw. 25 Minuten.

Wie schon im Pretest besagten auch diese Ergebnisse aus der Pilotphase, dass die Fragebogenlänge überwiegend als akzeptabel angesehen wird. In den Ergebnissen wurde kein Anlass gesehen die Fragebögen kürzen zu müssen.

6.4.9 Empfehlungen und Vorgaben für die Routine

Als Konsequenz der in der Pilotphase erhaltenen Ergebnisse wurden vom Steuerungskreis in der Sitzung am 23.01.2018 folgende Punkte beschlossen:

Das *Erhebungsdesign* wurde dahingehend geändert, dass Bezugspersonen nur dann in die Befragung einbezogen werden, wenn der Patient an der Befragung nicht teilnehmen kann oder will. Eine parallele Erhebung bei Patienten und Bezugspersonen fand also in keinem Fall statt. Zudem wurde auf die spezifische Messung der Prozessqualität und der Zufriedenheit aus Sicht der Bezugspersonen wegen der zu erwartenden geringen Fallzahlen im Routinebetrieb verzichtet (s. Tabelle 9 und Tabelle 10). Die Erwartung für die Routine leitete sich aus den in der Pilotphase realisierten Fallzahlen der Bezugspersonen ab. Wegen geringer Fallzahlen konnte keine statistisch belastbaren Einrichtungsergebnisse ermittelt werden. Bezugspersonen nahmen nur bei Nicht-Teilnahme von Patienten als Ersatz für die EQ-Messung am Verfahren teil.

Tabelle 9: Erhebungsdesign in der Routine, wenn der Patient an der Befragung teilnimmt

Befragte	Messzeitpunkt		
	Rehabilitations-Beginn	Rehabilitations-Ende	2 Wochen nach Rehabilitations-Ende
Patient, ambulant	EQ		EQ, PQ _a , Z _a
Patient, stationär	EQ		EQ, PQ _s , Z _s
Behandler	Confounder	Confounder	

Tabelle 10: Erhebungsdesign in der Routine, wenn anstatt des Patienten die Bezugsperson an der Befragung teilnimmt

Befragte	Messzeitpunkt		
	Rehabilitations-Beginn	Rehabilitations-Ende	2 Wochen nach Rehabilitations-Ende
Bezugsperson	EQ*		EQ*
Behandler	Confounder	Confounder	

*) Parallelversionen für Bezugspersonen zur EQ-Messung (s. Kapitel 3.1.3).

⁵² auf einer Skala von 1 bis 5

⁵³ auf einer Skala von 1 bis 5

Für den *Behandlerbogen* wurden Präzisierungen einzelner Frageformulierungen vorgeschlagen und umgesetzt.

Die Abfrage der Basis-Assessments bei Aufnahme wurde unverändert in die Routine übernommen. Auf Basis der Daten aus dem Routinebetrieb sollte die Verbreitung der aufgeführten Instrumente und Methoden sowie die dazu genannten Schwellenwerte untersucht werden. Konkret sollte geprüft werden, ob bzw. unter welchen Bedingungen sie für die Risikoadjustierung verwendet werden können.

Von den 33 Rehabilitations-Zielkategorien waren auch in der Pilotphase nur wenige so häufig besetzt, dass für sie der RZK-Score berechnet wurde, um dessen Messgüte statistisch zu bestimmen. Es wird davon ausgegangen, dass in der Routine genügend Fälle in jeder RZK vorliegen werden, so dass damit statistisch belastbare Ergebnisse der Ergebnisqualität bestimmt werden können. Von den 33 RZK gehen alle in die Routine, es findet keine Selektion statt.

Die Auswahl von mindestens drei und maximal zehn RZK aus mindestens drei ICF-Domänen waren in der Pilotphase unproblematisch. Dieser Auswahlmodus wurde für die Routine beibehalten.

In wenigen Fällen wurden in der Pilotphase 11 RZK ausgewählt, die maximale Anzahl auswählbarer RZK ist jedoch auf 10 festgelegt. Es wird geprüft, wie die Abweichung auftreten konnte, um sie in der Routine abzustellen.

Die behandlerseitige Einschätzung des globalen Rehabilitations-Erfolges sowie der Wahrscheinlichkeit das individuelle Rehabilitations-Ziel nach der Rehabilitation zu erreichen, werden beibehalten.

Die Teile der *Patientennachbefragungsbögen*, mit denen die *Prozessqualität* und die *Patientenzufriedenheit* abgefragt wurden, wurden entsprechend der Ergebnisse der Faktorenanalysen aus der Pilotphase erneut umgestellt bzw. umsortiert. Die Items wurden danach gruppiert, dass sie eine maximal reliable Messung der Qualitätsunterdimensionen erlauben. Innerhalb der Item-Gruppen wurden sie nach den Faktorladungen absteigend sortiert. Dabei wurden vier Items anderen Qualitätsunterdimensionen als im Pretest-Fragebogen zugeordnet⁵⁴. An den Items selbst wurden keine Änderungen vorgenommen, weder gibt es Umformulierungen noch Streichungen. Der Patientennachbefragungsbogen für ambulant behandelte Patienten wurde in Anlehnung an die Ergebnisse der stationären Patienten entsprechend geändert.

In den *Bezugspersonenbögen Beginn* und *Nachbefragung* sollte bei der Abfrage oder bei der Fragebogeninstruktion die Möglichkeit vorgesehen oder darauf hingewiesen werden, dass einzelne Fragen von der Bezugsperson nicht beantwortet werden konnten. In solchen Fällen sollte die Bezugsperson kein Kreuz machen. In der Instruktion der Bezugspersonenbögen wurde der Text entsprechend geändert.

Die zur Messung der *Ergebnisqualität* der RZK 1.1 „Schmerzen“ bisher verwendete 11-stufige Antwortskala ist nach Angabe von Mitgliedern des Steuerungskreises für Fremdbeurteilungen nicht geeignet. Die graduellen Schmerzabfragen RZK 1.1 b bis d wurden daher in der Routine mit einer 6-stufigen Antwortskala abgefragt, deren Stufen wie folgt gelabelt wurde: „Ich habe keine Schmerzen“, „sehr leicht“, „leicht“, „mittelmäßig“, „stark“, „sehr stark“.

Bei den RZK „Gehen im Innenbereich“, „Treppensteigen“ und „Gehen im Außenbereich“ sollte das Layout der Eingangsfrage optisch leichter erfassbar gestaltet werden. Entsprechend wurde der Text u.a. nicht mehr grau unterlegt.

7 Ergebnisdarstellung für die Einrichtungen

Das in diesem Projekt entwickelte QS-Verfahren soll sich in das QS-Reha[®]-Verfahren einfügen. Daher gibt es im Routinebetrieb für die teilnehmenden Reha-Einrichtungen auch einen Ergebnisbericht, in dem die in ihrer Einrichtung erzielten Ergebnisse (Ergebnisse der Indexeinrichtung) denen aller Einrichtungen (Referenzeinrichtungen) gegenüber gestellt werden. Im Einzelnen enthalten die Ergebnisberichte die Abschnitte: Drop-Out-Analyse, Qualitätssynopse sowie die Befragungsergebnisse auf der Ebene der Qualitätsunterdimensionen und auf Item-Ebene (Detailübersicht über die Beantwortung der einzelnen Items) für die Qualitätsdimensionen Ergebnisqualität, Prozessqualität und Patientenzufriedenheit. Die Qualitätssynopse enthält zudem Angaben zur Strukturqualität. Die vollständigen Ergebnisse der Strukturhebung werden im QS-Reha[®]-Verfahren gesondert im Bericht Teil 1 präsentiert.

⁵⁴ Im Patientennachbefragungsbogen - stationär - sind das die Items 22, 31, 38 und 39 (Details s. Anlage 16)

Im Rahmen des Projektes wurde ein Musterergebnisbericht auf Grundlage der hier definierten Qualitätsunterdimensionen vorgelegt und diskutiert, der die Struktur der Ergebnisberichte der bereits in der Routine des QS-Reha[®]-Verfahrens vertretenen Indikationsbereiche aufweist (s. Anlage 17). Die Einrichtungen, die an der Pilotphase teilgenommen haben, erhalten nach Abschluss des Projektes einen individuellen Ergebnisbericht.

Bei der Berechnung von Qualitätsunterdimensionen-Scores werden maximal 33,3% fehlende Werte bei den dazu einzubeziehenden Fragen toleriert.

Für die Indexeinrichtung werden nur dann Mittelwerte von Qualitätsunterdimensionen-Scores berechnet und angegeben, wenn von mindestens 5 Patienten valide Scores vorliegen.

Im QS-Reha[®]-Verfahren wird bei der Prozessqualität in den anderen Indikationsbereichen der Erfüllungsgrad über eine Item-weise inhaltlich begründete Dichotomisierung der positiven und der negativen Stufen ermittelt. Im Geriatrie-Modul wird der Erfüllungsgrad empirisch ermittelt. Dazu gibt es auch keine Alternative, da nach Herausnahme der Bezugspersonen aus der PQ-Messung, die Patientenabfrage übrig bleibt, die bereits dichotom erfolgt.

Der Ergebnisbericht gliedert sich in die folgenden acht Abschnitte:

- I Vorbemerkung
- II Untersuchung der Stichprobe
 - 1 Qualitätssynopse
 - 2 Ergebnisqualität
 - 3 Patientenzufriedenheit
 - 4 Prozessqualität
 - 5 Patientenzufriedenheit Detailübersicht
 - 6 Prozessqualität Detailübersicht

Im Rahmen der *Untersuchung der Stichprobe* erfolgt eine Drop-Out-Analyse mit Angaben zur Anzahl der Nicht-Teilnehmer und deren Gründe für die Nicht-Teilnahme, wobei hier Gründe von nicht-teilnehmenden Patienten und nicht-teilnehmenden Bezugspersonen aufgeführt werden. Danach werden verschiedene Patientenparameter (Alter, Geschlecht, Bewilligungsdauer, Antragsstellung) der Drop-Out-Fälle und der teilnehmenden Patienten auf Signifikanz der Unterschiede geprüft. Bei der folgenden Deskription der Stichprobe werden die o.g. Patientenparameter und weitere Merkmale, die im Behandlerbogen erfasst werden, der Indexeinrichtung mit denen der Referenzeinrichtungen auf Signifikanz geprüft. U.a. werden die drei häufigsten führenden Rehabilitation-begründenden und weiteren Rehabilitation-relevanten Diagnosen aufgeführt. Zusätzlich werden die drei am häufigsten genannten medizinischen Ereignisse mit Einfluss auf den Rehabilitationsverlauf berichtet.

In der *Qualitätssynopse* werden die Struktur- und Prozessqualität sowie anschließend die Ergebnisqualität und die Patientenzufriedenheit auf der Ebene der Qualitätsunterdimensionen dargestellt.

Bei der Struktur- und Prozessqualität wird der Erfüllungsgrad in Prozent für die Indexeinrichtung angegeben sowie der Durchschnitt der Referenzeinrichtungen und deren Interquartile Range. Bei der Ergebnisqualität und Patientenzufriedenheit werden für jede Qualitätsunterdimension die Anteile der Referenzeinrichtungen aufgeführt, die im Durchschnitt liegen, die signifikant über und die signifikant unter dem Durchschnitt liegen und in welcher Klasse das Ergebnis der Indexeinrichtung liegt.

Anschließend folgen die Zuweisungssteuerungskriterien aus der Strukturhebung. Es wird angezeigt, ob sie erfüllt sind oder nicht.

Nach der Qualitätssynopse werden die Ergebnisse der Ergebnisqualität, der Patientenzufriedenheit und der Prozessqualität auf Qualitätsunterdimensionsebene dargestellt:

Die Darstellung der Ergebnisse der *Ergebnisqualität* wird mit einem Diagramm mit Box-and-Whisker-Plots der risikoadjustierten Einrichtungsergebnisse eröffnet. Dieser Übersichtsdarstellung folgt eine weitere Übersichtsdarstellung, nämlich zu dem Profil der z-standardisierten Einrichtungsergebnisse. Daran anschließend werden die Ergebnisse der einzelnen Qualitätsunterdimensionen in der Form dargestellt, dass die beobachteten (O) und die erwarteten (E) Mittelwerte sowie deren Differenz (O - E) für die Index- und die Referenzeinrichtungen angegeben werden. Dem folgt ein Benchmarkdiagramm aller risikoadjustierten O - E-Einrichtungsergebnisse, bei dem das Ergebnis der Indexeinrichtung farblich kenntlich gemacht wird. Mit der graphischen Gegenüberstellung der beobachteten

Mittelwerte der Index- und der Referenzeinrichtungen in einem Balkendiagramm endet die Ergebnisdarstellung der einzelnen Qualitätsunterdimensionen.

Die Darstellung der Ergebnisse der Ergebnisqualität wird durch die Übersichtsdarstellung der Effektstärken der einzelnen Qualitätsunterdimensionen abgeschlossen, bei denen der Patientenstatus auf Intervallskalenniveau bzw. annähernd Intervallskalenniveau gemessen wird. Die Einrichtungsergebnisse werden als Balkendiagramm der Indexeinrichtung und der Referenzeinrichtungen dargestellt.

Die Ergebnisse der *Patientenzufriedenheit* werden in gleicher Weise wie die Ergebnisqualität zuvor im Ergebnisbericht dargestellt. Bei den Balkendiagrammen am Ende der Ergebnisdarstellung werden jedoch für alle einzelnen Fragen, die zur Aggregation des Scores der betreffenden Qualitätsunterdimension einfließen, die Erfüllungsgrade der Index- und der Referenzeinrichtungen im Vergleich nicht risikoadjustiert dargestellt. Zudem werden keine Effektstärken dargestellt.

Die Ergebnisse der *Prozessqualität* werden wie die der Patientenzufriedenheit zuvor dargestellt. Box-and-Whisker-Plots, Profil der z-standardisierten Einrichtungsergebnisse und abschließend die Ergebnisse der einzelnen Qualitätsunterdimensionen mit den Balkendiagrammen für alle einfließenden Fragen.

Dem folgen die Ergebnisdarstellungen für die Patientenzufriedenheit und Prozessqualität auf Item-Ebene, *Patientenzufriedenheit Detailübersicht* und *Prozessqualität Detailübersicht*, die den Abschluss im Bericht bilden.

Für jedes Item aller Qualitätsunterdimensionen werden für die Antwortstufen „ja“, „teils, teils“ und „nein“ (bei der Patientenzufriedenheit) bzw. „ja“ und „nein“ (bei der Prozessqualität) und ggf. „betrifft mich nicht“ sowie für fehlende Angaben („keine Angaben“) die Häufigkeiten absolut und relativ für die Indexeinrichtung und die Referenzeinrichtungen getrennt angegeben. Anschließend werden die Anzahlen der teilnehmenden Patienten („Anzahl Patienten“) und der valide beantworteten Fragen („Anzahl berücksichtigter Antworten“) entsprechend dargestellt. Zusätzlich wird der Zustimmunganteil in Prozent ausgewiesen („Anteil erreichter Punktzahl an maximal erreichbarer Punktzahl“).

Die Darstellung der erhaltenen Antworten zur Patientenzufriedenheit unterscheiden sich nur durch die zusätzliche Antwortstufe „teils, teils“ von der Prozessqualität. Ansonsten sind sie identisch.

8 Ausblick

Nachfolgend werden einige Punkte aufgeführt, die für das QS-Verfahren zukünftig bedacht, berücksichtigt oder entschieden werden sollten.

8.1 Risikoadjustierung

Die Ergebnisse aller Qualitätsunterdimensionen der Ergebnisqualität und der Patientenzufriedenheit werden im Ergebnisbericht risikoadjustiert ausgewiesen. Dazu werden hinreichend prädiktive multiple lineare Regressionsmodelle gesucht, die auf Basis der in den Behandlerbögen dokumentierten Patientenparameter als Prädiktoren die unterschiedlichen Qualitätsunterdimensionen-Scores als Kriterium möglichst gut beschreiben. Auf dieser Grundlage werden für alle Patienten O – E-Werte errechnet und einrichtungsbezogen dargestellt.

Bisher wurden mit den Daten des Pretests und der Pilotphase nur beispielhaft Risikoadjustierungsmodelle gerechnet. Damit konnte prinzipiell gezeigt werden, dass sich mit den erhobenen Confoundern und den berechneten RZK-Scores tragfähige Modelle finden lassen. Für den Routinebetrieb sind dann 41 Modelle neu zu bestimmen, für 33 RZK und die acht Qualitätsunterdimensionen der Patientenzufriedenheit.

Die Datenquelle, Patient oder Bezugsperson, wird bei der Bestimmung von Risikoadjustierungsmodellen der RZK immer in den Ausgangskonstellationen als Prädiktorvariable mit einbezogen, um abzusichern, dass die unterschiedlichen Datenquellen - wie intendiert und konzeptionell entsprechend realisiert - keinen bedeutsamen systematischen Einfluss auf die Ergebnis-Scores ausüben.

Außerdem ist bei der Entwicklung der Risikoadjustierungsmodelle für die RZK 1.1 „Schmerzen“ sowie für die RZK 1.3 „Defäkation“ die Schmerzmitteleinnahme zu t0 und zu t2 (RZK 1.1 Frage a) bzw. die Medikamenteneinnahme zu Häufigkeit und Konsistenz des Stuhlganges zu beiden Messzeitpunkten (RZK 1.3 Fragen b und d) in die Ausgangskonstellationen als Prädiktorvariablen mit einzubeziehen.

Als weiterer Einflussfaktor sollte das Bundesland, in dem die Einrichtung liegt, auf einen möglichen systematischen Einfluss auf die Scores überprüft werden. Um kleine Zellenbesetzungen zu vermeiden, können Länder mit vergleichbaren Ergebnissen zu Gruppen zusammengefasst werden.

8.2 Revision des Behandlerbogens

Der Behandlerbogen wurde im Zuge der Erhebungen bisher nur dahingehend überarbeitet, als dass die Abfrage der „Basis-Assessments bei Aufnahme“ nach dem Pretest vereinheitlicht wurde und die Vorgabe, bestimmte Messinstrumente einheitlich durchzuführen und deren Ergebnisse zu dokumentieren, zurückgenommen wurde. Die Anzahl der mit dem Behandlerbogen erfassten Confounder wurde auf Grundlage der Ergebnisse aus dem Pretest und der Pilotphase dabei nicht verändert.

Die Reduktion der Confounder kann erst in der Routine erfolgen. Denn erst auf Grundlage der Daten aus dem Routinebetrieb können die erforderlichen Risikoadjustierungsmodelle ermittelt werden. Der Behandlerbogen kann dann auf die Patientenparameter reduziert werden, die für diese Modelle und für die Drop-Out-Analyse im Ergebnisbericht benötigt werden.

8.3 Überprüfung der RZK und der RZK-Systematik

Die Validierung der RZK bezog sich im Pretest und in der Pilotphase auf die am häufigsten ausgewählten RZK. Damit steht faktisch die empirische Validierung für eine Vielzahl von RZK noch aus. Für die noch nicht überprüften RZK lässt sich der Einsatz im Routinebetrieb jedoch damit rechtfertigen, dass das Konzept der Operationalisierungen durchgängig gleich ist und in den überprüften Fällen valide und reliabel Ergebnisse lieferte. Die Validierung der bisher nicht überprüften RZK sollte auf Basis der in der Routine gewonnenen Daten erfolgen.

Außerdem lässt sich auf Grundlage der Daten aus dem Routinebetrieb überprüfen, ob sich innerhalb der 33 Rehabilitations-Zielkategorien Strukturen empirisch entdecken lassen, die ggf. von der inhaltlich-formal begründeten ICF-Systematik der ICF-Domänen abweichen. Denkbar wäre hierfür eine explorative Faktorenanalyse, um mit Einrichtungsdaten zu rechnen. Im positiven Fall könnten die Ergebnisse der 33 RZK auf wenige, ihnen zugrunde liegende, wechselseitig unabhängige Faktoren verdichtet werden, was einer Vereinfachung der Ergebnisstruktur gleichkäme. Damit würden die Einrichtungsergebnisse ggf. an Aussagekraft gewinnen, als dass die Ergebnisse mehrerer RZK dann nicht mehr als unabhängige Resultate verstanden würden, sondern als unterschiedliche Facetten eines übergeordneten Qualitätsaspektes. Zudem könnte sich die Risikoadjustierung der Ergebnisqualitäts-Scores vereinfacht darstellen, als dann nur noch Modelle für die gefundenen RZK-Faktoren ermittelt werden müssten. Die Modelle der subsummierten RZK dürften sich untereinander ähneln.

8.4 Aufwände zur Realisierung des Verfahrens

Das Verfahren im Routinebetrieb durchzuführen ist mit spezifischen Aufwänden verbunden, die zum einen bei den teilnehmenden Rehabilitations-Einrichtungen liegen und zum anderen bei der Auswertungsstelle nach § 299 Abs. 3 SGB V anfallen. Grundsätzlich lassen sich bei beiden Akteuren zwei Komponenten unterscheiden: Das sind die Aufwände für die erstmalige Teilnahme bzw. Einrichtung des Qualitätssicherungsverfahrens und die sich davon abhebenden Aufwände für den Routinebetrieb.

8.4.1 Rehabilitations-Einrichtungen

Im Nachgang des Pretests wurden die Koordinatoren der daran teilnehmenden Einrichtungen zur ihren Erfahrungen befragt, die sie bei der praktischen Umsetzung der Erhebung gesammelt haben. Die Befragung erfolgte strukturiert unter Verwendung eines Leitfadens, der sieben Teilaspekte der praktischen Umsetzung differenzierte, die mit jeweils zwei Fragen, der Aufwand für die Einrichtungen mit drei Fragen, insgesamt also mit 15 Fragen angesprochen wurden (s. Anlage 18). Anfang Dezember 2016 konnten insgesamt acht Telefoninterviews durchgeführt werden. Detaillierte Ergebnisse sind den anliegenden Unterlagen zu entnehmen (s. Anlage 19).

In den Einrichtungen fallen mit der Erhebung verbundene Aufwände an, die sich danach unterscheiden, ob sie im Vorfeld oder während der Erhebung selbst entstehen. Im Vorfeld der Erhebung besteht der Aufwand u.a. darin, die Koordinatoren der Einrichtungen zu schulen und das Verfahren, die Abläufe, die Instrumente und die IT-Strukturen kennen zu lernen und in den Einrichtungen das erforderliche Wissen an die damit befassten Kollegen zu vermitteln. Bei der

Erhebung selbst steht die Akquise der Patienten und ggf. der Bezugspersonen, deren Information, das Verteilen und Einholen der Einwilligungserklärungen sowie der Fragebögen sowie, die Dokumentation der einbezogenen Fälle im Behandlerbogen an. Für die Nachbefragung sind die Fragebögen zu versenden, deren Rücklauf bei der Datenerfassungsstelle online zu beobachten und ggf. Nachfassaktionen durchzuführen. Schließlich erfolgt der Versand der gesammelten ausgefüllten t0-Fragebögen an das BQS-Institut.

Der mit der Erhebung verbundene Aufwand wurde von den Befragten in zwei Fällen als „zu aufwändig“ bewertet, in vier Fällen wurde an der Stelle das Potenzial des Verfahrens betont. Der berichtete Zeitaufwand für die unterschiedlichen Professionen fiel zwischen den Einrichtungen unterschiedlich aus. Für die Koordinatoren lagen die Angaben z.B. zwischen einer bis 2,5 Stunden pro Patient, im Mittel bei 1,6 Stunden pro Patient⁵⁵. Die Angaben der Ärzte variierten zwischen keiner bis einer Stunde pro Patient, im Mittel lagen sie bei 0,5 Stunden pro Patient. Zur Durchführung der Pretesterhebung liegt der Gesamtaufwand pro Patient zwischen 2,25 und 4 Stunden und im Mittel bei 3,1 Stunden.

Es ist darauf hinzuweisen, dass sich der Aufwand in der Routine im Vergleich zu den berichteten Aufwänden pro Patient deutlich verringern dürfte, weil die Angehörigenermittlung für die Parallelerhebung von Patienten und Bezugspersonen, wie im Pretest gefordert, wegfällt und die Akquisezeiträume länger sind und dadurch zunehmend effektivere Abläufe gefunden werden. Die zyklische Wiederholung der Befragungen wird zudem eine Steigerung der Effektivität bewirken. Die Dokumentation der Drop-Out-Fälle sowie die Nachfassaktionen, um die Mindestanzahl vollständig dokumentierter Fälle zu erreichen, sind in den erfassten Zeitangaben nicht enthalten, weil sie im Pretest nicht vorgesehen waren. Außerdem sind die Aufwände für die Schulungen der Koordinatoren zusätzlich zu berücksichtigen.

Aus den Rückmeldungen einer stationären Einrichtung war zu entnehmen, dass dort die Abläufe zu Rehabilitations-Beginn grundlegend umorganisiert wurden, um die Auswahl der patientenspezifischen RZK möglichst frühzeitig abzuschließen, damit die t0-Messungen möglichst frühzeitig durchgeführt werden können und damit der Zeitraum möglichst groß wird für die Delta-Bildung, also für die Maßnahmen, die einen Therapieeffekt erzielen, möglichst groß wird. An der Stelle bleibt jedoch offen, ob die Umstellung nur für die Zeit der Befragung oder ob sie von Dauer war und ob sie damit eine nachhaltige Veränderung zu einer effektiveren Behandlung der Patienten bewirkt hat.

8.4.2 Auswertungsstelle nach § 299 Abs. 3 SGB V

Auch bei der Auswertungsstelle, die mit der Durchführung der Befragungen in der Routine beauftragt ist, lassen sich grundsätzlich zwei Komponenten des Aufwandes unterscheiden, zum einen die systematische und inhaltliche Implementierung des Verfahrens und zum anderen die anschließende Durchführung.

Der technisch-inhaltliche Aufbau des Verfahrens erfolgt einmalig, solange das Verfahren unverändert bleibt. Folgende Bereiche gehören zur Implementierung:

- **Metadatenrepository**
 - Zentrale Verwaltung und Versionierung aller Variablen, die in den einzelnen Bereichen online oder über einen Papierfragebogen erhoben werden und am Ende gemeinsam ausgewertet werden

- **Schnittstelle zur teilnehmenden Einrichtung (Extranet)**
 - Registrierung der teilnehmenden und nicht teilnehmenden Einrichtungen mit deren Fachabteilungen, mit der Option Vorgaben der GKV zu ändern und zu ergänzen
 - Benutzerverwaltung mit Rechtestrukturen für Koordinator, Stellvertreter, Behandler und Visitor
 - Bereitstellung von News des Verfahrens mit Ankündigung per Email
 - Bereitstellung von indikations- und rollenspezifischen Downloads
 - Supportplattform als zentrale transparente Schnittstelle für die Kommunikation zwischen der Auswertungsstelle und den beteiligten Personen und Einrichtungen (Telefon, Email, Chat, Post)

⁵⁵ Die Zeitangaben beziehen sich auf die Aufwände in den Einrichtungen im Pretest.

- Sichere Bereitstellung der Ergebnisberichte Teil 1 und Teil 2 als Download für die teilnehmenden Einrichtungen (nur für den Koordinator)
 - Fachabteilungsspezifische elektronische Erhebung der Strukturdaten durch den Koordinator oder Stellvertreter. Umfangreiche Validierung, Abschluss, Erhebung druckbar in DIN A4, vollständige Datenhistorie
 - Fachabteilungsspezifische elektronische Erhebung der Behandlerdaten jedes teilnehmenden Patienten oder dessen Bezugsperson (jeweils spezifisch) durch den Koordinator, Stellvertreter oder Behandler. Umfangreiche Validierung, Abschluss, Erhebung druckbar in DIN A4, vollständige Datenhistorie
 - Fachabteilungsspezifisch und auf Basis der dokumentierten RZK-Auswahl in den dokumentierten Behandlerdaten individuell zusammengestellter druckbarer Fragebogen für teilnehmende Patienten oder Bezugspersonen zum Behandlungsbeginn und zur Nachbefragung inklusive Teilnehmerkennung zur Datenzusammenführung
 - Fachabteilungsspezifische elektronische Überprüfung der Strukturdaten nach Visitation durch den Visitor. Umfangreiche Validierung, Abschluss, Erhebung druckbar in DIN A4, vollständige Datenhistorie
- Scanningsystem
 - Verarbeitung von individualisiert zusammengestellten Fragebögen
 - Menschliche Validierung der Scannergebnisse als Teil der Qualitätssicherung
 - Erhebung von Freitextangaben
 - Automatisierter Export in den Datenpool des Verfahrens
 - Schnittstelle zur Darstellung des Rücklaufs inklusive des Verarbeitungsprozesses einzelner Fragebögen in das Teilnehmer-Extranet in Echtzeit für alle Indikationen
- Auswertungssystem
 - Tool zur Datenzusammenführung, Aufbereitung und Bereitstellung
 - Übernahme aller Metadaten aus dem Metadatenrepository
 - Definition der Kennzahlen pro Bericht
 - Integration der Risikoadjustierung
 - Einbindung von speziellen statistischen Methoden aus R
 - Erstellung des Layouts pro Bericht (Tabellen, Texte, Grafiken)
 - Produktion der Auswertungsberichte Teil 1 (Strukturerhebung) und Teil 2 (Patientenbefragung) als PDF und XLSX
 - Darstellung der Ergebnisse der teilnehmenden Fachabteilung einer Einrichtung gegen das Gesamtergebnis aller teilnehmenden Einrichtungen dieser Fachabteilung
 - Exportschnittstelle zur elektronischen Bereitstellung aller Berichte selektiv
- DBMS - Datenbankmanagementsystem
 - Verarbeitung von Tabellen mit mehr als 1500 Feldern
- XML-Exportschnittstelle - Ergebnisdaten und Berichte an ITSG⁵⁶ (im Auftrag der GKV)
 - Export bestimmter Kennzahlen in einem definierten Exportformat (XML) - Teil 1
 - Export bestimmter Kennzahlen in einem definierten Exportformat (XML) - Teil 2
 - Export der Ergebnisberichte aller Teilnehmer für Teil 1 und Teil 2
 - Export der Teilnehmerliste im Exportformat XML
 - Übertragung an Test- und Produktivsystem, asymmetrisch verschlüsselt und mit SFTP-Übertragen auf einen Server der ITSG

Da die Berechnungen der Scores komplex sind, müssen zur Vermeidung von Berichtsfehlern Maßnahmen zur Qualitätssicherung vorgesehen werden. Diese wären z.B. die Auswertungen zur Überprüfung der Angaben in den Ergebnisberichten in einem anderen Format parallel zu programmieren, um bei Abweichungen Hinweise auf mögliche Fehler zu erhalten. Schließlich sind 41 Risikoadjustierungsmodelle für diverse Scores zu entwickeln.

Die Durchführung des QS-Verfahrens in der Routine erfolgt regelmäßig im Dreijahreszyklus. Bestandteile der Durchführung der Befragungen sind

- die Einrichtungsakquise

⁵⁶ Informationstechnische Servicestelle der gesetzlichen Krankenversicherung

- die Schulung der Koordinatoren
- der inhaltliche und technische Support
- die Datenerfassung der Fragebögen inklusive deren Prüfung
- die Datenhaltung
- die Datenbereitstellung
- die Produktion der Ergebnisberichte
- die Überprüfung der Angaben in den Ergebnisberichten
- die Überprüfung und Anpassung von 41 Risikoadjustierungsmodellen
- die inhaltlich oder technisch bedingte Umsetzung einzelner Anpassungserfordernisse.

Der Aufwand für die Implementierung hängt wesentlich von den, beim Auftragnehmer bereits vorhandenen Tools ab, um die o.g. Prozesse zu realisieren. Ohne Vorhandensein entsprechender Werkzeuge bedarf die Entwicklung mehrere Mannjahre. Als wesentliche Komplexität ist dabei die Programmierung der risikoadjustierten Auswertungen im Bereich der RZK zu sehen, zu denen es in dieser Form, nach Kenntnisstand des Berichterstellers, noch keine Routinetools gibt.

Der Aufwand für die sich regelmäßig wiederholende Durchführung im Routinebetrieb ist in der Größenordnung mit den übrigen Verfahren des QS-Reha[®]-Verfahrens zu vergleichen. Die Schätzungen beziehen sich auf den Teil 2 der Erhebung, da die Strukturhebung nicht Gegenstand des Forschungsprojektes war.

8.5 Nutzung von Daten nach § 299 SGB V für das beschlossene Verfahren

§ 299 Abs. 1 SGB V beinhaltet die Befugnis, Daten zum Zwecke der Qualitätssicherung zu erheben. Bei der Entwicklung von Qualitätssicherungsverfahren für die medizinische Versorgung von gesetzlich versicherten Patienten ist stets zu prüfen, ob der Aufwand für die dazu erforderliche Datengewinnung reduziert werden kann, indem bereits vorliegende Daten genutzt und einbezogen werden. Dazu lässt sich jedoch feststellen, dass die QS-Reha-Daten anonymisiert vorliegen. Dies bedeutet, dass die darin enthaltenen Patientendaten anderen Daten aus anderen Qualitätssicherungsverfahren nicht individuell zugeordnet werden können. Eine Verwendung von Daten aus anderen Qualitätssicherungsverfahren scheidet somit aus.

Zudem ist zu prüfen, ob auf Grundlage von § 299 Abs. 1a SGB V auch „Routinedaten“ der Kassen genutzt werden können. Grundlage hierfür ist, dass § 299 Abs. 1a SGB V auf § 284 SGB V verweist, in dem u.a. Abrechnungsdaten angeführt (§ 284 Abs. 1 Nr. 8 SGB V) werden. Hierbei wäre zu prüfen, ob auf Grundlage von Routinedaten relevante Parameter genutzt werden können, die den Gesundheitsstatus vor und nach den Rehabilitations-Maßnahmen erfassen, also den Behandlungsprozess über Sektorengrenzen hinweg beschreiben und die Wirksamkeit und Nachhaltigkeit der Rehabilitations-Maßnahmen damit objektiv dargestellt werden kann.

§ 299 Abs. 3 Satz 3 SGB V besagt, dass Daten, die zum Zwecke der Qualitätssicherung verarbeitet werden, nicht mit für andere Zwecke als die Qualitätssicherung erhobenen Datenbeständen zusammengeführt und ausgewertet werden dürfen. Das heisst nach heutigem rechtlichen Stand dürften Abrechnungsdaten und sonstige versichertenbezogene Leistungsdaten nicht mit QS-Daten zusammengeführt werden.

Im vorliegenden Fall besteht ein weiteres praktisches Hindernis darin, dass der Datenaustausch zwischen den Rehabilitations-Einrichtungen und den Kassen für die Abrechnung der erbrachten Leistungen derzeit umgestellt wird von einem papiergebundenen Verfahren auf einen elektronischen Datenaustausch. Aktuell hat eine Vielzahl von Rehabilitations-Einrichtungen diese Umstellung noch nicht vollzogen. Zudem gibt es kein definiertes Zeitziel, zu dem die Umstellung vollständig abgeschlossen sein soll. Das Verfahren für den elektronischen Datenaustausch ist im Detail definiert, nur noch nicht flächendeckend umgesetzt. Die Routinephase hat also noch nicht begonnen. Damit liegen die Routinedaten der Rehabilitations-Einrichtungen derzeit nicht für alle behandelten geriatrischen Patienten in elektronischer Form vor. Der Einbezug von Routinedaten in das hier entwickelte Qualitätssicherungsverfahren ist wegen ihrer mangelnden Vollständigkeit z.Z. nicht möglich.

Wenn diese rechtlichen und tatsächlichen Hindernisse nicht bestehen würden, wären zudem weitere theoretische Überlegungen zur Sinnhaftigkeit der Nutzung von Routinedaten notwendig. Diese werden im Folgenden thematisiert.

Mit dem Behandlerbogen werden überwiegend Patientenmerkmale und einzelne Prozessparameter erfasst. Nur bei sehr wenigen davon gibt es Überschneidungen mit den Routinedaten. Dazu zählen Alter, Geschlecht, Aufnahmedatum, Entlassungsdatum, Diagnosen und Verlängerungsdaten des Patienten. Im QS-Verfahren auf diese Daten zuzugreifen, anstatt sie in den Einrichtungen eingeben zu lassen, scheint keinen nennenswerten Vorteil zu liefern. Die Messung der Ergebnis- und der Prozessqualität sowie der Patientenzufriedenheit erfolgen auf Basis von Befragungen. Hier ist eine Substitution durch Routinedaten grundsätzlich ausgeschlossen.

Es bleibt weiterhin zu überlegen, inwieweit der Einbezug von Routinedaten dazu genutzt werden kann, die Auswirkungen der Rehabilitations-Maßnahmen zu erfassen und darzustellen.

Aus methodischer Sicht lässt sich die Wirksamkeit von Rehabilitations-Maßnahmen grundsätzlich in Querschnitts- und in Längsschnittstudien zeigen. Dazu zählen u.a. das Fall-Kontroll-Design und der Vorher-Nachher-Vergleich. Das Fall-Kontroll-Design scheidet als mögliches Evaluationsdesign aus ethischen Gründen aus, weil zur in der geriatrischen Rehabilitation behandelten Patientengruppe keine vergleichbare Personengruppe gefunden werden kann, bei der zwar ein vergleichbarer Rehabilitations-Bedarf besteht, dem aber nicht oder in systematisch anderer Form nachgekommen wird.

Bei einem Vorher-Nachher-Vergleich werden Indikatoren (State-Merkmale) erforderlich, die den Gesundheits- bzw. Krankheitsstatus im betrachteten Patientenkollektiv valide und reliabel erfassen und eine angemessene Änderungssensitivität aufweisen. Als Indikatoren, die mit ihren Veränderungen auch die langfristige Wirksamkeit von Rehabilitations-Maßnahmen im Vorher-Nachher-Vergleich abbilden, sind z.B. die Verordnung von Hilfsmitteln (Art und Frequenz), der Bezug von Medikamenten (differenziert nach Wirkstoff und Dosis) sowie die Durchführung von therapeutischen Maßnahmen (Art und Frequenz) denkbar. Die damit gewonnenen Ergebnisse, können auf globaler Ebene interpretiert und zur Evaluation des Verfahrens herangezogen werden. Auswertungen und Ergebnisse auf Einrichtungsebene werden aufgrund der besonderen Inhomogenität des geriatrischen Patienten-klientels erst nach einer Standardisierung der Einrichtungsergebnisse interpretierbar. Denkbar ist auch die Betrachtung von Subgruppen-Ergebnissen, was einer Stratifizierung der Studienpopulation entspricht. Analysen auf Patientenebene sind zulässige Einzelfallbetrachtungen, die jedoch nicht verallgemeinerbar sind.

Empirische Wirksamkeitsuntersuchungen auf Basis von Routinedaten sind jedoch nicht Bestandteil des Projektes. Diese können im Rahmen gesondert beauftragter Projekte durchgeführt werden.

8.6 Mobile geriatrische Rehabilitation

Im Jahr 2016 wurden bundesweit 1.242 Patienten in 12 Einrichtungen mit einer mobilen geriatrischen Rehabilitation versorgt (Pippel et al., 2017). Die Indikation für mobile Rehabilitations-Maßnahmen verlangt gemäß den Rahmenempfehlungen zur mobilen geriatrischen Rehabilitation⁵⁷ (RE MoGeRe) neben der Rehabilitations-begründenden Diagnose sowie Rehabilitationsbedürftigkeit, Rehabilitationsfähigkeit, alltagsrelevanten Rehabilitationszielen und positiver Rehabilitationsprognose, dass mindestens eine der folgenden vier erheblichen Schädigungen vorliegt, um die Durchführung im vertrauten Wohnumfeld zu begründen:

- erhebliche Schädigung mentaler Funktionen kognitionsbezogener Art
- erhebliche Schädigung mentaler Funktionen verhaltensbezogener Art
- erhebliche Schädigung der Stimm-, Sprech- oder Sprachfunktion mit vorbestehenden Kompensations-/Ersatzstrategien im ständiger Wohnumfeld
- erhebliche Schädigung der Seh- und Hörfunktion mit vorbestehenden spezifischen Anpassungen im ständiger Umfeld

Mehrfachnennungen dieser vier Indikationskriterien sind zulässig.

8.6.1 Regelmäßige Erhebung von Basis- und Strukturdaten

Seit 2008 ist das Kompetenz Centrum Geriatrie (KCG) durch die Verbände der Krankenkassen auf Bundesebene mit der einrichtungübergreifenden Erhebung und Auswertung der Basisdokumentation Mobile Geriatrische Rehabilitation beauftragt.

Die Basisdokumentation dient der Beschreibung des Leistungsgeschehens und damit als Entscheidungsgrundlage für künftige Versorgungsentscheidungen und ggf. der Weiterentwicklung der mobilen geriatrischen Rehabilitation. Sie ersetzt kein externes Qualitätssicherungsverfahren. Die Teilnahme der Einrichtungen an der Basisdokumentation ist vertraglich geregelt. Die Basis-

⁵⁷ Siehe unten: Spitzenverbände der gesetzlichen Krankenkassen

dokumentation ist als Vollerhebung aller GKV-Fälle konzipiert und damit die größte Datenbasis zur mobilen geriatrischen Rehabilitation in Deutschland (Meinck et al., 2017).

Mit der Basisdokumentation werden von den teilnehmenden Einrichtungen für jeden im Kalenderjahr behandelten Patienten Basisdaten wie Alter, Geschlecht, Wohnort, Pflegestufe und Art der Pflegeleistung, medizinische Befunde wie Hauptdiagnose und Indikationskriterien und Assessmentbefunde zur Selbstversorgungsfähigkeit gemäß Barthel-Index erhoben. Zusätzlich werden prozessseitige Basisdaten wie Zuweisungspfad, Aufnahme- und Entlassungsdatum, Behandlungstage, Rehabilitations-Verlängerung, Behandlungsunterbrechung, Grund für das Behandlungsende sowie Anzahl und Art der Therapieeinheiten erfasst.

Parallel zur Basisdokumentation erfolgt ebenfalls im Jahresrhythmus, zum Stichtag 31.12., eine Strukturerhebung. Der dazu verwendete Strukturfragebogen erfragt Art und Stellenumfang der in der Rehabilitationseinrichtung tätigen Berufsgruppen. Für die Professionen: Arzt, Pflegekraft, Sozialarbeiter, Diätassistent/Ökotrophologe, Physiotherapeut, Ergotherapeut, Stimm-/Sprech-/Sprachtherapeut und klinischer (Neuro-) Psychologe wird der Stellenumfang für das Verfahrensjahr in Vollzeitäquivalenten erfasst. Außerdem wird nach weiteren geriatrischen Versorgungsangeboten am gleichen Standort sowie nach der Anzahl durchgeführter mobiler Rehabilitationsmaßnahmen gefragt.

8.6.2 Jahresbericht

Das KCG wertet die Daten der Basisdokumentation sowie der Strukturerhebung aus und stellt sie in einem Jahresbericht für den GKV-Spitzenverband und die Verbände der Krankenkassen auf Bundesebene zusammen (Pippel et al., 2017). Der Jahresbericht ist nicht öffentlich zugänglich, die Berichtshoheit liegt beim GKV-Spitzenverband.

Patientenseitige Basisdaten

Der Bericht enthält einrichtungsübergreifende und einrichtungsbezogene Auswertungen der patientenseitigen Basisdaten.

Einrichtungsübergreifende Auswertungen der Basisdaten dienen z.B. zur Beschreibung des behandelten Patientenkollektivs. Dabei werden den Ergebnissen des betrachteten Verfahrensjahres die Vorjahresangaben gegenübergestellt. Im Einzelnen berichtet werden:

- Alters- und Geschlechtsverteilung,
- Verteilung der Pflegestufen und der Wohnsituation,
- Art und Anzahl der vorliegenden Indikationskriterien,
- die zehn häufigsten Hauptdiagnosen,
- Barthel-Index zu Rehabilitations-Beginn und zu Rehabilitations-Ende

Zusätzlich werden einrichtungsbezogene Auswertungen der Basisdaten dargestellt zu:

- Fallzahlen,
- Zuweisungspfaden,
- Verteilungen der 5-Jahres-Altersklassen,
- Pflegestufen,
- Anteilen der Pflegeheimbewohner,
- Anzahlen der Indikationskriterien,
- Barthel-Indices zu Rehabilitations-Beginn und den Differenzen der Barthel-Indices zwischen Rehabilitations-Beginn und Rehabilitations-Ende,

so dass die Daten der einzelnen teilnehmenden Rehabilitations-Einrichtungen identifiziert und miteinander verglichen werden können.

Prozesseitige Basisdaten

Prozessparameter, die einrichtungsübergreifend dargestellt werden, sind:

- Rehabilitationsdauer,
- erbrachte Therapie- und Leistungseinheiten nach Berufsgruppen getrennt,
- Behandlungsdichte,
- Entlassungsgründe.

Einrichtungsbezogen dargestellt werden die Prozessparameter:

- Anzahl erbrachte Therapie- und Leistungseinheiten nach Berufsgruppen getrennt,
- Anteil der Berufsgruppen in den Rehabilitations-Maßnahmen,

Behandlungsdichte pro Behandlungstag für regulär entlassene Patienten, in Kurzzeitpflege erbrachte Therapie- und Leistungseinheiten.

Strukturdaten

Die Auswertungen der Strukturdaten werden für jede einzelne teilnehmende Einrichtung gesondert dargestellt. Für jede Einrichtung werden die Vollzeitäquivalente nach Berufsgruppen getrennt aufgelistet, wobei diese noch einmal danach differenziert werden, wie viele davon für Therapie- und Leistungseinheiten aufgewendet wurden bzw. für Therapie- und Leistungseinheiten inklusive Wegezeiten.

Gespräche mit Einrichtungen aufgrund fachlich auffälliger Befunde

Auf Grundlage in den Jahresberichten festgestellter auffälliger Befunde werden seit 2012 im Auftrag des GKV-Spitzenverbandes und der Verbände der Krankenkassen auf Bundesebene punktuell Gespräche mit ausgewählten Einrichtungen geführt. Die Gespräche erfolgen vor Ort oder telefonisch durch einen Fachgutachter des KCG. Die Ergebnisse werden in den Jahresberichten dargelegt und ggf. ergänzend mit den Auftraggebern besprochen.

Diese Gespräche dienen dazu, die Hintergründe unerwarteter oder auffällig von anderen Einrichtungen abweichender Ergebnisse in der Basisdokumentation auf potenzielle Weiterentwicklungsbedarfe indikativer, prozessualer oder sonstiger Aspekte der RE MoGeRe zu prüfen.

8.6.3 Qualitätssicherung in der mobilen geriatrischen Rehabilitation

Die durch das KCG jährlich erhobene Basisdokumentation inklusive des Strukturbogens sind nicht zur Qualitätssicherung konzipiert. Die erhobenen Daten werden zwar einrichtungsvergleichend dargestellt, es gibt jedoch keine Referenzbereiche, die gute Qualität definieren und keine Sanktionsmechanismen. Die Ergebnisdarstellung bleibt deskriptiv und beruht ausschließlich auf den von den Einrichtungen selbst generierten und übermittelten Falldaten.

Das im Rahmen des Projektes entwickelte und in diesem Abschlussbericht ausführlich dargestellte Verfahren zur individualisierten *Messung der Ergebnisqualität* in der ambulanten und stationären geriatrischen Rehabilitation ist konzeptionell so angelegt, dass es ohne Anpassung unmittelbar auf die mobile geriatrische Rehabilitation angewendet werden könnte. Der Ansatz mit den individuellen Rehabilitations-Zielen, der Auswahl von drei bis zehn aus 33 Rehabilitations-Zielkategorien, den Operationalisierungen der Rehabilitations-Zielkategorien sowie der Patienten- respektive der Alternativbefragung von Bezugspersonen bei Ausfall der Patienten könnten identisch übernommen werden. Außerdem könnte der Behandlerbogen zur Datengewinnung für die Risikoadjustierung übernommen werden.

Zur *Messung der Prozessqualität, der Patientenzufriedenheit und der Strukturqualität* wäre wegen der deutlich anderen Strukturen und Prozesse der Leistungserbringung in der MoGeRe eine zumindest partielle Neuentwicklung der spezifischen Instrumente erforderlich. Hierzu wären wieder alle Schritte zur Exploration des Themenfeldes und die zweistufige empirische Überprüfung der Fragebögen durchzuführen, um Instrumente mit vergleichbarer Güte zu bekommen. Das Verfahren zur Beauftragung der Entwicklung und die Entwicklung der Instrumente selbst dürften einige Zeit bis zur Einsatzreife benötigen.

Aus methodischer Sicht ist ein vergleichendes Qualitätssicherungsverfahren auf Grundlage der teilnehmenden Einrichtungen und den realisierten Behandlungsvolumina grundsätzlich möglich. Auch hier leistet das Benchmarkverfahren seine orientierende Funktion für die einzelnen Einrichtungen hinsichtlich der erbrachten Versorgungsqualität. Für die Patienten ist eine vergleichende Ergebnisdarstellung eher von nachgeordneter Relevanz, da für sie kaum realistische Auswahloptionen bestehen.

Die vergleichsweise geringe Anzahl an Einrichtungen ist methodisch unproblematisch. Die Fallzahlen liegen im Durchschnitt bei ca. 100 Fällen pro Einrichtung und Jahr. Rund 20% der Einrichtungen dokumentieren derzeit weniger als 20 Fälle p.a. Solche Einrichtungen können, wie im QS-Reha®-Verfahren ab weniger als 50 Fälle vorgesehen, eine Auswertung ohne Einrichtungsvergleich erhalten.

8.6.4 Fazit

Mit der Basisdokumentation und dem Strukturbogen werden seit Jahren kontinuierlich grundlegende Versorgungsdaten der mobilen geriatrischen Rehabilitation flächendeckend erhoben.

Der Behandlerbogen sowie das Verfahren zur Messung der Ergebnisqualität für die ambulante und stationäre geriatrische Rehabilitation könnten ohne nennenswerten Entwicklungsaufwand kurzfristig in der mobilen geriatrischen Rehabilitation eingesetzt werden. Die für die mobile geriatrische Rehabilitation spezifischen Instrumente zur Messung der Prozessqualität und der Patientenzufriedenheit müssten zumindest teilweise neu entwickelt werden. Entsprechend der im QS-Reha[®]-Verfahren verwendeten Bögen zur Messung der Strukturqualität der ambulanten und der stationären Rehabilitation müsste zudem ein spezifisch an die Verhältnisse in der mobilen geriatrischen Rehabilitation angepasster Bogen entwickelt werden.

Mit den beiden empirischen Ansätzen werden unterschiedliche Ziele verfolgt. Mit der Basisdokumentation und der Strukturhebung werden unterschiedliche epidemiologische und formale Aspekte der mobilen geriatrischen rehabilitativen Versorgung flächendeckend erhoben, um ggf. Änderungsbedarfe bei den Rahmenempfehlungen der MoGeRe zu erkennen. Mit dem neu entwickelten und in diesem Abschlussbericht dargestellten Qualitätssicherungsverfahren wird die Versorgungsqualität der Rehabilitation von ambulant und stationär behandelten GKV-Patienten gemessen und die Ergebnisse einrichtungsbezogen den beteiligten Akteuren zur Steuerung vorgelegt.

Hinsichtlich der Ein- und/oder Fortführung eines oder beider dieser Ansätze sind folgende Punkte zu berücksichtigen:

- Mit einem Verzicht auf die Basisdokumentation und die Strukturhebung geht eine jahrelange Kontinuität der Ergebnisse dieser Erhebung und ihrer Verlaufsentwicklung verloren.
- Zugleich würden Änderungsbedarfe für die Rahmenempfehlungen der MoGeRe nicht mehr systematisch sondiert werden.
- Eine Implementierung des neuen Qualitätssicherungsverfahrens und seine Anpassung an die mobile geriatrische Rehabilitation ist mit erheblichem Mehraufwand verbunden, u.a. wegen der erforderlichen Neuentwicklungen für die Bereiche Prozessqualität, Patientenzufriedenheit und Strukturqualität.
- Durch die Einführung des neuen Qualitätssicherungsverfahrens werden Ergebnisse wichtiger Qualitätsdimensionen kurz- bis mittelfristig verfügbar. Insbesondere die Ergebnisqualität dürfte über das interne QM der Einrichtungen zum Nutzen der Patienten beitragen.
- Aus methodischer Sicht gibt es keine zwingenden Gründe, die mobile geriatrische Rehabilitation nicht ins QS-Reha[®]-Verfahren aufzunehmen.

9 Bewertung des beschlossenen QS-Verfahrens

Abschließend soll betrachtet werden, in wieweit sich das in diesem Projekt entwickelte Verfahren zur Qualitätssicherung der Versorgung von Patienten der ambulanten und stationären geriatrischen Rehabilitation eignet.

Die Durchführung von zwei unabhängigen Erhebungen liefert eine solide Grundlage zur Beurteilung unterschiedlicher relevanter Aspekte des vorliegenden Verfahrens, die am Ende dazu beitragen, ein praktisch und messtechnisch überprüftes Verfahren vorlegen zu können, das in der Routine seinen Zweck erfüllt und zur Verbesserung der Qualität der Versorgung von Patienten in der geriatrischen Rehabilitation beitragen kann.

Die Ergebnisse aus dem Pretest und aus der Pilotphase zeigen eindeutig, dass die Messung der Ergebnisqualität, der Prozessqualität und der Patientenzufriedenheit durch Befragung von Patienten bzw. deren Bezugspersonen einen praktisch und messtheoretisch erfolgreichen Ansatz darstellt.

Die Messinstrumente werden von den Patienten und von den Bezugspersonen verstanden und angenommen. Sie werden von ihnen als inhaltlich angemessen betrachtet und werden entsprechend vollständig ausgefüllt. Die beobachteten Rücklaufquoten liegen noch deutlich unter den Werten, die in den anderen Indikationsbereichen des QS-Reha[®]-Verfahrens üblich sind. Es kann jedoch davon ausgegangen werden, dass auch in der geriatrischen Rehabilitation vergleichbare Werte in der Routine erreicht werden, wenn nämlich die Einrichtungen ihre Nachfassaktionen starten, auf die im Pretest und in der Pilotphase verzichtet wurde.

Bezogen auf die Messung der Ergebnisqualität hat sich gezeigt, dass das Konzept der ereignisorientierten Abfrage für Patienten funktioniert und die alternative Befragung von Bezugspersonen reliable Angaben beisteuern kann, wenn Patienten für die Befragung nicht zur Verfügung stehen. Zudem zeigt sich, dass das Prä-Post-Design mit der Nullpunktmessung zu Rehabilitations-Beginn und der Wiederholungsmessung zwei Wochen nach Rehabilitations-Ende für geriatrische Patienten eine geeignete Methode zur Veränderungsmessung darstellt. Die individualisierte Messung der Ergebnisqualität hat sich ebenfalls bewährt. Dazu zählen im Einzelnen: der Prozess der Vereinbarung und Festlegung des individuellen Rehabilitations-Ziels zwischen dem Behandler und dem Rehabilitanden, die Systematik der 33 RZK, der festgelegte Auswahlmodus sowie die einzelnen Operationalisierungen. Für die am häufigsten ausgewählten RZK konnten überwiegend positive empirische Belege für die Reliabilität der Messungen sowie deren kriteriumsbezogener Validität gefunden werden.

Die zur Messung der Prozessqualität und der Patientenzufriedenheit definierten Qualitätsunterdimensionen wurden in zwei unabhängigen Stichproben faktorenanalytisch identifiziert und bestätigt. D.h., die im Pretest gefundenen Faktoren konnten in der Pilotphase mit hoher Übereinstimmung repliziert werden. Das besagt, dass die Faktoren stabile Entitäten erfassen. Diese eignen sich im besonderen Maße dazu, Qualitätsunterdimensionen zu definieren. Der Vorteil des faktorenanalytischen Ansatzes ist, dass dadurch Urteilsdimensionen verwendet werden, die die Patienten implizit selbst verwenden, dass die Messergebnisse zudem wechselseitig unabhängig sind und die Unterdimensionen innerhalb der betrachteten Qualitätsdimensionen eine Einfachstruktur bilden. Diese Befunde belegen die Konstruktvalidität der mit den Instrumenten gemessenen Qualitätsunterdimensionen. Zudem konnte in beiden Erhebungen die Reliabilität dieser Qualitätsunterdimensionen gezeigt werden.

Die Messeigenschaften der entwickelten Instrumente sind empirisch stabil und belegen deren Güte. Auf der anderen Seite besagen die guten bzw. gut interpretierbaren Ergebnisse beider unabhängigen Erhebungen in der Summe, dass die zu Rehabilitations-Beginn und zwei Wochen nach Entlassung aus der Rehabilitations-Einrichtung befragten geriatrischen Patienten in der Lage sind, sinnvolle und in sich konsistente Auskünfte zu geben und dass ihre Antworten weit davon entfernt sind, zufällig oder schematisch und nicht belastbar zu sein.

Mit den vorliegenden überprüften Instrumenten lässt sich die Ergebnis- und die Prozessqualität sowie die Zufriedenheit der Patienten mit der in den Einrichtungen erbrachten Versorgung valide und reliabel erheben. Unter Berücksichtigung der Confounder lassen sich auf dieser Grundlage Einrichtungsvergleiche durchführen.

Die zur Datenerfassung und zur Kommunikation zwischen den Einrichtungen und dem BQS-Institut entwickelten IT-Strukturen wurden aus dem QS-Reha[®]-Verfahren übernommen und angepasst. Für das Geriatriemodul wurde jedoch zusätzlich eine Anwendung programmiert, die erlaubt, auf Grundlage der patientenspezifisch ausgewählten RZK alle für den Patienten erforderlichen individualisierten Fragebögen unmittelbar nach Bestätigung der Eingabe bereitzustellen. Damit können die Einrichtungen zeitnah mit den individualisierten Fragebögen versorgt werden, ohne dass sie selbst die Fragebögen vorhalten und zusammenstellen müssten. Diese Anwendung hat sich in der Praxis bewährt.

Die Erfahrungen aus beiden Erhebungen haben weiterhin gezeigt, dass die Komplexität des Verfahrens, der erwartete Aufwand für die Einrichtung wie für die Patienten sowie die internen Abläufe auf die Erfordernisse des Verfahrens einzustellen, die Einrichtungen anfangs irritiert. Nach anfänglicher Zurückhaltung gewinnt am Ende die Überzeugung die Oberhand, an einem ernstzunehmenden Qualitätssicherungsverfahren teilgenommen zu haben.

Der Pretest und die Pilotphase sind als Stichprobenerhebungen geplant und durchgeführt worden und sind daher beide weit von den Fallzahlen und Einrichtungszahlen entfernt, wie sie in der Routine erzielt werden. Darin liegt begründet, dass bestimmte Komponenten des Qualitätssicherungsverfahrens zum Projektabschluss noch ausstehen, für deren Entwicklung ein vollständiger Datensatz, wie er in der Routine gewonnen wird, notwendig vorliegen muss.

Hier ist die Entwicklung aller Risikoadjustierungsmodelle zu nennen sowie die Reduktion des Behandlerbogens, die erst nach der Bestimmung aller Risikoadjustierungsmodelle sinnvoll erfolgen kann. Auch die Validierung aller bisher nicht empirisch überprüften RZK ist eine Aufgabe, die auf Grundlage beider durchgeführten Erhebungen nicht umgesetzt werden konnte, weil durch die

individualisierten Ergebnisqualitätsmessungen, anders als in einem einheitlichen Standardverfahren, nicht von jedem teilnehmenden Patienten zu allen Qualitätsaspekten gleichermaßen Angaben vorlagen und damit häufig die empirische Basis zu deren Überprüfung fehlte. Diese Einschränkung war zur Zeit der Beauftragung nicht absehbar, weil die Idee der Individualisierung der Ergebnisqualitätsmessung als Lösungsansatz erst im Laufe des Projektes entwickelt wurde.

Diese Aufgaben können angegangen werden, wenn die Daten aus dem Routinebetrieb vollständig vorliegen.

10 Danksagung

Am Zustandekommen des in diesem Projekt entwickelten Qualitätssicherungsverfahrens für die geriatrische Rehabilitation war eine Vielzahl von Personen und Einrichtungen beteiligt, denen an dieser Stelle für ihren Beitrag gedankt werden soll.

Der Dank geht an die Rehabilitations-Einrichtungen, die uns bei der Exploration der Themenfelder unterstützt und die Diskussion mit Therapeuten, Ärzten sowie Mitarbeitern der Pflege und der Sozialdienste und die Durchführung der kognitiven Pretests ermöglicht haben. Den Rehabilitations-Einrichtungen, die am Pretest und an der Pilotphase teilgenommen haben, ist ebenfalls zu danken. Zudem sind die Patienten und Bezugspersonen zu erwähnen, die sich an den kognitiven Pretests sowie am Pretest und an der Pilotphase beteiligt und mit ihren Angaben zur Entwicklung des Verfahrens beigetragen haben.

Dank gilt ebenfalls den externen Experten für ihre Beiträge in den Experteninterviews und den Experten-Workshops sowie den Teilnehmern im Steuerungskreis und den Mitarbeitern des GKV-Spitzenverbandes für die kritische Begleitung der Entwicklung des Projektes.

Für die externe Unterstützung des BQS-Instituts möchten wir den Experten, Prof. Dr. med. Clemens Becker und Mitarbeitern (Robert-Bosch-Krankenhaus, Stuttgart) sowie Prof. Dr. Erik Farin-Glattacker (Institut für Medizinische Biometrie und Statistik, Freiburg), unseren besonderen Dank aussprechen.

Die ehemaligen Kollegen, F. - J. Grothaus und Dr. med. Dagmar Hertle, haben sich im besonderen Maße für die Entwicklung der Gesamtkonzeption sowie die Exploration und Operationalisierungen der zu messenden Qualitätsaspekte engagiert. Dafür danken wir ihnen herzlich. Schließlich ist auf die Kollegen der IT-Abteilung hinzuweisen, die einen wesentlichen Beitrag zur Realisierung des Projektes beigesteuert haben.

Die genannten Beteiligten haben in unterschiedlicher Form zum Gelingen des Projektes beigetragen und die Erstellung eines komplexen Instrumentariums unterstützt, das im Idealfall die Qualität der Versorgung von Patienten in der geriatrischen Rehabilitation zu verbessern hilft.

11 Literatur/Quellen

- Bortz J. Lehrbuch der empirischen Forschung für Sozialwissenschaftler. Heidelberg. Springer Verlag; 1984.
- Bortz J. Statistik für Human- und Sozialwissenschaftler. 6. Auflage, Heidelberg. Springer Verlag; 2005.
- Bortz J. und Döring N. Forschungsmethoden und Evaluation. 4. Auflage, Heidelberg. Springer Verlag; 2006.
- De Morton NA, Davidson M, Keating JL (2008) The de Morton Mobility Index (DEMMI): an essential health index for an ageing world. Health Qual Life Outcomes, 6: 63
- Fisseni H-J. Lehrbuch der psychologischen Diagnostik. Mit Hinweisen zur Intervention. Fisseni H-J, (Hrsg.). 3. Auflage, Göttingen; Bern; Toronto; Seattle; Oxford; Prag. Hogrefe; 2004.
- Folstein MF, Folstein SE, Mc Mugh PR. „Mini-mental state“: a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975; 12: 189-198
- Gassmann K. Geriatrische Rehabilitation vor der Gesundheitsreform - Beispiele aus Bayern: Qualitätssicherung von größtem Interesse. In: Füsgen IH (Hrsg.). Zukunftsforum Demenz Geriatrische Rehabilitation Vom Ermessen zur Pflicht - auch für den dementen Patienten, S. 27-35. Wiesbaden: Medical Tribune Verlagsgesellschaft mbH, 2007.
- Granger CV, Hamilton BB, Linacre JM, Heinemann AW, Wright BD. Performance profiles of the functional independence measure. In: Am J Phys Med Rehabil 1993; 72: 84-89.
- Häder M. Empirische Sozialforschung. Wiesbaden. VS Verlag für Sozialwissenschaften, 2006.
- Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI (1990) Clarifying confusion: The Confusion Assessment Method. A new method for detection of delirium. Ann Intern Med, 113: 941–948.
- Lachs MS, Feinstein AR, Cooney LM, Jr., Drickamer MA, Marottoli RA, Pannill FC et al. A simple procedure for general screening for functional disability in elderly patients. Ann Intern Med 1990;112: 699-706.
- Lienert GA, Raatz U. Testaufbau und Testanalyse. Lienert GA, Raatz U, (Hrsg.). 6. Auflage, Weinheim. Beltz Psychologie Verlags Union; 1998.
- Mahoney FI, Barthel DW. Functional Evaluation. The Barthel Index. MD State Med J 1965; 14: 61-65.
- Meinck M, Pippel K, Lübke N (2017) Mobile geriatrische Rehabilitation in der gesetzlichen Krankenversicherung: Konzeptionelle Ausrichtung und Ergebnisse der bundesweiten Basisdokumentation (Teil 1), Z Gerontol Geriat 50 (3): 226-232.
- Pippel K, Lübke N, Meinck M (2017) Basisdokumentation mobile geriatrische Rehabilitation - Jahresbericht 2016. Hamburg, Kompetenz-Centrum Geriatrie. Der Jahresbericht ist nicht öffentlich zugänglich, die Berichtshoheit liegt beim GKV-Spitzenverband.
- Spitzenverbände der gesetzlichen Krankenkassen (Hrsg) Rahmenempfehlungen zur mobilen geriatrischen Rehabilitation. https://www.vdek.com/vertragspartner/vorsorge-rehabilitation/mobile_reha.html.
- World Health Organization (WHO). ICF – Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit. Deutsches Institut für Medizinische Dokumentation und Information (DIMDI), WHO-Kooperationszentrum für das System Internationaler Klassifikationen (Hrsg.). Genf. 2005.
- Yesavage JA, Brink TL, Rose TL et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 1982-83; 39: 37-49.