

**Qualitätssicherung in der Physiotherapie**

**Evaluation von Performance-Tests**

Dissertation zur  
Erlangung des akademischen Grades eines  
Doktors der Philosophie  
im Fachbereich Gesellschaftswissenschaften  
der Universität Kassel

**vorgelegt von**  
**WOLFGANG WEINHOLD**

**Tag der Disputation: 20. Februar 2008**

- 1. Gutachter: Prof. Dr. Manfred Wegner**
- 2. Gutachter: Prof. Dr. Volker Scheid**

**Kassel, August 2008**

## ERKLÄRUNG

Hiermit versichere ich, dass ich die vorliegende Dissertation selbstständig und ohne unerlaubte Hilfe angefertigt und andere als die in der Dissertation angegebenen Hilfsmittel nicht benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen sind, habe ich als solche kenntlich gemacht. Kein Teil dieser Arbeit ist in einem anderen Promotions- oder Habilitationsverfahren verwendet worden.

Kassel, den 15.8.08

---

Wolfgang Weinhold

## DANKSAGUNG

*Wer sich in der Mitte eines nichtwissenschaftlichen beruflichen Lebens wieder den Wissenschaften zuwendet, weiß um die Schwierigkeiten universitäre Anbindungen zu knüpfen. Mein größter Dank gilt daher meinem Doktorvater Herrn Professor Dr. Manfred Wegner – Lehrstuhl für Sportpsychologie an der Universität Kassel – dem ich für die bereitwillige Übernahme des Promotionsvorhabens, seine Unterstützung und sein persönliches Engagement danke. Ich danke ihm für die Selbstständigkeit, die er mir für die Erstellung dieser Arbeit gewährte und für die gemeinsamen Diskussionen in den verschiedenen Stadien der Entstehung, in denen ich immer wieder wichtige inhaltliche Impulse erhalten habe. Sehr angenehm ist seine Fähigkeit Kritik so zum Ausdruck zu bringen, dass immer noch ein Lob daraus zu entnehmen ist, was mir Mut machte, wenn ich zweifelte.*

*Ein besonderer Dank gilt auch Frank Naeve, dem Leiter des Ambulanten Reha-zentrums Kiel GmbH und dem Geschäftsführer des Lubinus Clinicums in Kiel Herrn Manfred Schmid, die mir die Untersuchungsdurchführung zur Dissertation im Reha-zentrum ermöglichten und organisatorische Veränderungen mittrugen, die von den Mitarbeiterinnen Maria Ehmke und Brigitte Seidel umgesetzt wurden.*

*Dank an meine therapeutischen Kollegen Michael Drauschke, Okke Duit, Bente Misch und Annika Roth für die zuverlässige und konstruktive Mitarbeit. Ohne sie als Therapeuten wäre eine praktische Durchführung nicht denkbar gewesen.*

*Mein besonderer Dank gilt dem Verein zur Förderung der Rehabilitationsforschung in Hamburg, Mecklenburg-Vorpommern und Schleswig-Holstein e.V., der die finanziellen Mittel zur Durchführung des Projektes bereitstellte. Insbesondere Frau Dr. Ruth Deck möchte ich einen ganz herzlichen Dank aussprechen. Sie war immer für mich ansprechbar und gab mir fundierte methodische und statistische Beratung und wurde nie müde, meine Entwürfe zu lesen und zu kommentieren.*

*Nichts von all dem ist ohne soziale Unterstützung denkbar. Nur so gelingt es, die krisenhaften Phasen aufzufangen und gut zu überdauern. Familie und Freunde gaben mir immer wieder Anlass den Schreibtisch zu verlassen und sorgten für angenehme Unterbrechungen, um den damit auch notwendigen Abstand zur Arbeit zu finden.*

*Dank an meine Familie.*

# INHALTSVERZEICHNIS

<b>1</b>	<b>EINLEITUNG</b>	<b>1</b>
1.1	Fragestellung.....	4
1.2	Ziel der Arbeit.....	4
<b>2</b>	<b>THEORETISCHER BEZUGSRAHMEN</b>	<b>7</b>
2.1	Salutogenese und biomedizinische Sichtweise.....	7
2.2	Internationale Klassifikation der Funktionsfähigkeit.....	9
2.3	Konzept der funktionalen Gesundheit.....	11
2.4	Forschungsstand in Physiotherapie und Rehabilitation.....	15
2.4.1	Forschungsstand in der Physiotherapie.....	15
2.4.2	Theoriengenerierung in der Physiotherapie.....	19
2.4.3	Forschungsstand in der Rehabilitation.....	20
2.4.4	Theoriengenerierung in der Rehabilitation.....	23
2.4.4.1	Theoriemodell der Rehabilitation.....	23
2.4.4.2	Bewältigungsprozesse.....	25
2.5	Therapieziele in Physiotherapie und Rehabilitation.....	27
2.6	Tests und Assessmentverfahren.....	32
2.6.1	Die Begriffe Test und Assessment.....	32
2.6.2	Testgütekriterien.....	34
2.6.2.1	Hauptgütekriterien.....	34
2.6.2.2	Nebengütekriterien.....	36
2.6.3	Vorzüge von Tests und Einflüsse auf Ergebnisse.....	37
2.6.4	Gliederung von Testverfahren.....	39
2.6.4.1	Instrumente der Selbsteinschätzung.....	39
2.6.4.2	Instrumente der Fremdbeurteilung.....	44
2.6.4.3	Kombinierte Testverfahren.....	45
2.6.5	Testverfahren im Kontext von Qualitätssicherung.....	46
2.6.6	Testverfahren im Kontext von Leitlinien.....	49
2.6.7	Zur Bedeutung von Performance-Tests.....	51
2.6.8	Zur Systematik von Performance-Tests.....	52

	2.6.9 und Forschungsstand und Forschungsdefizit.....	53
<b>2.7</b>	<b>Ableitung der Items für den Polla.....</b>	<b>58</b>
<b>2.8</b>	<b>Herleitung der Hypothesen.....</b>	<b>61</b>
<b>3</b>	<b>METHODE.....</b>	<b>63</b>
<b>3.1</b>	<b>Pilotstudie.....</b>	<b>63</b>
	3.1.1 Erprobung und Ergebnisse.....	63
	3.1.2 Kritische Bewertung und Konstruktion des Polla.....	65
<b>3.2</b>	<b>Hauptstudie.....</b>	<b>66</b>
	3.2.1 Studiendesign.....	66
	3.2.2 Patientenauswahl und Parallelisierung.....	67
<b>3.3</b>	<b>Erhebungsinstrumente.....</b>	<b>68</b>
	3.3.1 Der Performance-Test (Polla).....	68
	3.3.2 Gesundheitsbezogene Lebensqualität (SF-36).....	70
	3.3.3 Selbsteinschätzung von Aktivitäten.....	70
	3.3.4 Befundaufnahme und spezifische Zielkriterien.....	70
<b>4</b>	<b>DURCHFÜHRUNG DER STUDIE.....</b>	<b>72</b>
<b>4.1</b>	<b>Patientenrekrutierung.....</b>	<b>72</b>
<b>4.2</b>	<b>Erhebungsplan.....</b>	<b>72</b>
<b>4.3</b>	<b>Durchführung zur Interrater- und Intrarater Reliabilität.....</b>	<b>73</b>
<b>4.4</b>	<b>Kontrolle von Störgrößen und Qualitätssicherung.....</b>	<b>73</b>
<b>4.5</b>	<b>Behandlungsinhalte und Verantwortlichkeiten.....</b>	<b>74</b>
<b>4.6</b>	<b>Merkmale der Patientenstichprobe und Drop-out.....</b>	<b>75</b>
<b>4.7</b>	<b>Merkmale der parallelisierten Stichproben.....</b>	<b>76</b>
<b>4.8</b>	<b>Diagnosen und die Behandlungsstatistik.....</b>	<b>77</b>
<b>4.9</b>	<b>Statistische Auswertestrategie.....</b>	<b>78</b>
<b>5</b>	<b>AUSWERTUNGEN.....</b>	<b>80</b>
<b>5.1</b>	<b>Entscheidungsgrundlage für die Skalenbildung.....</b>	<b>80</b>
	5.1.1 Itemstatistik zu Polla und Selbsteinschätzung.....	80
	5.1.2 Faktorenanalyse zum Polla.....	82

5.1.3	Faktorenanalyse zur Selbsteinschätzung.....	83
5.1.4	Konsequenzen für die Summenskalenbildung.....	84
<b>5.2</b>	<b>Datenaufbereitung.....</b>	<b>85</b>
<b>6</b>	<b>ERGEBNISSE.....</b>	<b>87</b>
<b>6.1</b>	<b>Deskriptive Statistik zu Therapiebeginn.....</b>	<b>87</b>
6.1.1	Deskriptive Statistik zum Polla.....	87
6.1.2	Deskriptive Statistik zu den Summenskalen.....	90
6.1.3	Ergebnisse der spezifischen Zielkriterien.....	91
<b>6.2</b>	<b>Deskriptive Statistik zu Therapieende.....</b>	<b>92</b>
6.2.1	Deskriptive Statistik zum Polla.....	92
6.2.2	Deskriptive Statistik zu den Summenskalen.....	95
6.2.3	Ergebnisse der spezifischen Zielkriterien.....	96
6.2.4	Deskriptive Statistik zur Vergleichsgruppe.....	96
6.2.5	Zusammenhänge der spezifischen Zielkriterien.....	97
6.2.6	Nebenwirkungen von Tests.....	98
<b>6.3</b>	<b>Schließende Statistik.....</b>	<b>99</b>
6.3.1	Überprüfungen auf Normalverteilung.....	99
6.3.2	Veränderungen der Polla Items.....	99
6.3.3	Veränderungen der SF-36-Skalen.....	101
6.3.4	Analyse von Unterschieden in Altersklassen.....	103
6.3.4.1	Ergebnisse der Varianzanalysen.....	103
6.3.5	Analyse von Unterschieden in Diagnosegruppen.....	104
6.3.5.1	Merkmale der Stichprobe.....	105
6.3.5.2	Ergebnisse der Varianzanalysen.....	105
6.3.6	Veränderungen der spezifischen Zielkriterien.....	106
<b>6.4</b>	<b>Überprüfung von Testgütekriterien zum Polla zu t1 und t2.....</b>	<b>107</b>
6.4.1	Objektivität (Interrater-Reliabilität).....	107
6.4.2	Reliabilität.....	108
6.4.2.1	Intrarater-Reliabilität.....	108
6.4.2.2	Konsistenzanalyse.....	109
6.4.3	Konstruktvalidität.....	109
6.4.4	Kriteriumsvalidität (Korrelationshypothese).....	110
6.4.5	Unterschiede in der Parallelisierung.....	112
6.4.6	Änderungssensitivität (Unterschiedshypothese).....	113
<b>6.5</b>	<b>Auswertungen zum Follow-up mit 26 Patienten.....</b>	<b>113</b>
6.5.1	Deskriptive Statistik zum Follow-up.....	114

6.5.1.1	Itemstatistik und Reliabilität.....	115
6.5.2	Schließende Statistik zum Follow-up.....	116
6.5.2.1	Verteilungen und Gruppenunterschiede.....	116
6.5.2.2	Veränderungen der Items zum Polla.....	117
6.5.3	Testgütekriterien zum Polla im Follow-up.....	119
6.5.3.1	Konstruktvalidität.....	119
6.5.3.2	Kriteriumsvalidität.....	120
6.5.3.3	Unterschiede zw. Follow-up und Gesunden.....	120
6.5.3.4	Änderungssensitivität.....	121
<b>6.6</b>	<b>Zusammenfassende Darstellung der Testgütekriterien.....</b>	<b>123</b>
<b>7</b>	<b>DISKUSSION</b>	<b>125</b>
7.1	Diskussion der Kriteriumsvalidität.....	125
7.2	Diskussion der Änderungssensitivität.....	127
7.3	Diskussion der Konstruktvalidität.....	128
7.4	Diskussion zu Gruppenunterschieden.....	129
7.5	Diskussion zu spezifischen Zielkriterien.....	130
<b>8</b>	<b>SCHLUSSFOLGERUNGEN UND EMPFEHLUNGEN</b>	<b>132</b>
<b>9</b>	<b>ZUSAMMENFASSUNG</b>	<b>133</b>
	<b>LITERATUR</b>	<b>136</b>
ANHANG 1	Polla.....	148
ANHANG 2	Gesundheitsbezogene Lebensqualität (SF-36).....	149
ANHANG 3	Umkodierte und kalibrierte Items des SF-36.....	152
ANHANG 4	Selbsteinschätzung von Aktivitäten.....	153
ANHANG 5	Befundaufnahme und spezifische Zielkriterien.....	154
ANHANG 6	Patienteninformation und Einwilligungserklärung .....	155
ANHANG 7	Polla modifiziert.....	157
ANHANG 8	Testanweisung Polla modifiziert.....	158

## ABBILDUNGSVERZEICHNIS

Abb. 1:	Das bio-psycho-soziale Modell der ICF (WHO, S. 3, 2006).....	12
Abb. 2:	Theoriemodell der Rehabilitation (Gerdes & Weis, 2000, S. 48)	24
Abb. 3:	Schema zum Ablauf von Zielformulierung, Indikatoren, Outcome und Outcomemessung.....	31
Abb. 4:	Messzeitpunkte in der Längsschnittstudie.....	67
Abb. 5:	Postoperativer Therapiebeginn in Tagen. Waagrecht: Fallzahlen; Vertikal: Tage nach Operation.....	78
Abb. 6:	Die Verteilungen zum 1-Minute-Up & Go-Test zu $t_1$ (N=81).....	88
Abb. 7:	Die Verteilungen zum 2-Minuten-Gehtest zu $t_1$ (N=81).....	88
Abb. 8:	Die Verteilungen zum Treppenlaufen-Test zu $t_1$ (N=81).....	89
Abb. 9:	Die Mittelwerte der Summenskalen zu $t_1$ (N=81).....	90
Abb. 10:	Die Verteilungen zum 1-Minute-Up & Go-Test zu $t_2$ (N=81).....	93
Abb. 11:	Die Verteilungen zum 2-Minuten-Gehtest zu $t_2$ (N=81).....	93
Abb. 12:	Die Verteilungen zum Treppenlaufen-Test zu $t_2$ (N=81).....	94
Abb. 13:	Die Mittelwerte der Summenskalen zu $t_2$ (N=81).....	95
Abb. 14:	Die Mittelwerte der Summenskalen zu $t_1$ und $t_2$ (N=81) im Vergleich zur Gruppe von Gesunden (Ges; N=30).....	101
Abb. 15:	Die Mittelwerte der SF-36 Skalen zu $t_1$ und $t_2$ (Angaben in Prozent). KÖFU: Körperliche Funktionsfähigkeit; KÖRO: Körperliche Rollenfunktion; SCHM: Schmerz; AGES: Allgemeiner Gesundheitszustand; VITA: Vitalität; SOFU: Soziale Funktionsfähigkeit; EMRO: Emotionale Rollenfunktion; PSYC: Psychisches Wohlbefinden (N=81).....	102
Abb. 16:	Die Mittelwerte der Summenskalen zum Follow-up.....	115
Abb. 17:	Die Veränderungen der 0-100 transformierten Tests von $t_1$ bis zum Follow-up.....	119

## TABELLENVERZEICHNIS

Tab. 1:	Unterschiede zwischen der ICF und der ICDH (Schuntermann, 2002, S. 3).....	10
Tab. 2:	Gegenüberstellung von Krankheitsverarbeitung und Abwehr (Gerdes & Weis, 2000, S. 59).....	26
Tab. 3:	Beurteilung von Objektivitätskoeffizienten (Clarke, 1976, S. 27, aus Bös, 2001, S. 546).....	34
Tab. 4:	Darstellung von Faktoren, die Muskelkraftmessungen beeinflussen (Simmonds 1997, aus Cabri, 2001, S. 211).....	38
Tab. 5:	Generische Instrumente für Gesundheitszustand und intervenierende Merkmale (Biefang & Schuntermann, 2000, S. 110, Auszüge).....	40
Tab. 6:	Itemzuordnung zu den Summenskalen des SF-36 (Bullinger & Kirchberger, 1998, S. 66).....	42
Tab. 7:	Spezifische Instrumente für Gesundheitszustand und intervenierende Merkmale (Biefang & Schuntermann, 2000, S. 110, Auszüge).....	43
Tab. 8:	Verfahren der Qualitätssicherung in den einzelnen Stadien des Qualitätssicherungsprozesses (Dorenburg & Tiefensee, 2000, S. 205).....	48
Tab. 9:	Performance-Tests für die obere/untere Extremität (Ziel: Gleichgewicht, Funktionszustand).....	54
Tab. 10:	Performance-Tests für die untere Extremität (Ziel: Gleichgewicht, Funktionszustand).....	55
Tab. 11:	Performance-Tests zur Beurteilung der Kniefunktion nach Knieverletzungen (Bös, 2001, S. 354 f).....	56
Tab. 12:	Zuordnung von ICF-Kategorien zu Tests der Literatur und Ableitung von Items für den Polla.....	59
Tab. 13:	Ergebnisse der Pilotphase: Tests in der Reihenfolge deren	

	Durchführung nicht möglich (n.m.) oder möglich (m.) war. Daneben die Anzahl überprüfter Tests (Test), Keine Angaben (K.A.) sowie das Gesamttestaufkommen (Ges).....	64
Tab. 14:	Deskriptive Ergebnisse der Pilotphase.....	64
Tab. 15:	Die Beurteilungskategorien zu den Polla Items.....	68
Tab. 16:	Die Aktivitäten im Polla (1-14) inklusive der gemessenen Tests (11-14).....	69
Tab. 17:	Die Beurteilungskategorien zu den Items der Selbsteinschät- zung zu Aktivitäten.....	70
Tab. 18:	Die Erfassung der spezifischen Zielkriterien in ihren Merk- malsausprägungen.....	71
Tab. 19:	Der Erhebungsplan im zeitlichen Verlauf der Untersuchung.....	72
Tab. 20:	Verteilung von Alter, Geschlecht und Body-Mass-Index (BMI) im Patientenkollektiv.....	75
Tab. 21:	Drop-out nach Geschlecht, Alter, Behandlungsanzahl und Ur- sachen des Abbruchs.....	75
Tab. 22:	Gruppenstatistiken: Verteilung von Alter, Geschlecht und BMI in den parallelisierten Gruppen „Patient“ und „Gesunde“ (n=30).....	76
Tab. 23:	Test bei unabhängigen Stichproben zu Unterschieden in Alter und BMI zwischen den parallelisierten Gruppen „Patient“ und „Gesunde“ (N=30).....	76
Tab. 24:	Deskriptive Statistik zu Alter und BMI in den parallelisierten Gruppen „Patient“ und „Gesunde“.....	77
Tab. 25:	Die Häufigkeiten der aufgenommenen Diagnosen (N=81).....	77
Tab. 26:	Der Schwierigkeitsindex und die Trennschärfen zu den Polla Items zu $t_1$ und $t_2$ .....	80
Tab. 27:	Der Schwierigkeitsindex und die Trennschärfen zur Selbstein- schätzung zu $t_1$ und $t_2$ .....	81
Tab. 28:	Erklärte Gesamtvarianz nach der Extraktionsmethode der	

	Hauptkomponentenanalyse ( $t_1$ ).....	82
Tab. 29:	Extraktionsmethode der Hauptkomponentenanalyse: Komponentenmatrix ( $t_1$ ).....	82
Tab. 30:	Hauptkomponentenanalyse mit Oblimin-Rotation ( $t_1$ ).....	83
Tab. 31:	Die Erklärte Gesamtvarianz nach der Extraktionsmethode der Hauptkomponentenanalyse.....	83
Tab. 32:	Extraktionsmethode nach der Hauptkomponentenanalyse mit Oblimin-Rotation zu $t_1$ .....	84
Tab. 33:	Die Häufigkeiten zu den Polla Items zu $t_1$ .....	87
Tab. 34:	Die Deskriptive Statistik zu den gemessenen Tests zum Messzeitpunkt $t_1$ .....	90
Tab. 35:	Mittelwerte und Standardabweichungen zu den Summenskalen (0-100) zu $t_1$ .....	91
Tab. 36:	Die deskriptive Statistik zur Schmerzintensität, Flexion und Kniebeuge zu $t_1$ .....	92
Tab. 37:	Die Häufigkeiten zu den Polla Items zu $t_2$ .....	92
Tab. 38:	Die deskriptive Statistik zu den gemessenen Tests ( $t_2$ ).....	94
Tab. 39:	Die Mittelwerte und Standardabweichungen der Summenskalen ( $t_2$ ).....	95
Tab. 40:	Die deskriptive Statistik zur Schmerzintensität, Flexion und Kniebeuge zu $t_2$ .....	96
Tab. 41:	Deskriptive Statistik zu den gemessenen Tests der Vergleichsgruppe.....	96
Tab. 42:	Die Zusammenhänge zu spezifischen Merkmalen nach Spearman-Rho.....	97
Tab. 43:	Korrelationen zwischen der aktiven Flexion, der schmerzhaften Kniebeuge und den SF-36 Dimensionen.....	98
Tab. 44:	Nebenwirkungen in Form von Schmerz, Druck oder Ziehen im Anschluss an Testdurchführung.....	98

Tab. 45:	Die Veränderungen der Polla Items $t_2$ gegenüber $t_1$ .....	99
Tab. 46:	Die deskriptive Statistik der gemessenen Tests zu den Zeitpunkten $t_1$ und $t_2$ .....	100
Tab. 47:	Die Ergebnisse des T-Tests auf Stichprobenunterschiede bei gepaarten Tests $t_1/t_2$ .....	100
Tab. 48:	Die deskriptive Statistik zu den SF-36 Skalen.....	102
Tab. 49:	Die deskriptive Statistik in den Altersklassen und Verteilungen der Diagnosen.....	103
Tab. 50:	Ergebnisse der Varianzanalysen zu Unterschieden zwischen Altersklassen (95% KI).....	104
Tab. 51:	Die deskriptive Statistik zu Alter und Geschlecht in den Diagnosegruppen.....	105
Tab. 52:	Ergebnisse der Varianzanalysen zu Unterschieden zwischen Diagnosegruppen (95% KI).....	105
Tab. 53:	Die Veränderungen der spezifischen Zielkriterien $t_1$ gegenüber $t_2$ nach dem T-Test.....	106
Tab. 54:	Die Veränderungen der spezifischen Zielkriterien $t_2$ gegenüber $t_1$ nach Wilcoxon.....	107
Tab. 55:	Das Kappa-Maß zu den drei Interratern (it) zu den Items.....	107
Tab. 56:	Korrelationskoeffizient in Klassen ( $t_2$ /Retest; durchschnittliche Maße).....	108
Tab. 57:	Interne Konsistenz zum Polla und zur Selbsteinschätzung.....	109
Tab. 58:	Korrelationen nach Spearman-Rho im Patientenkollektiv (n=81)	110
Tab. 59:	Die Korrelationen zwischen den Messinstrumenten zu Therapiebeginn.....	111
Tab. 60:	Die Unterschiede in den Gruppen „Gesunde“ und „Patienten“ (Varianzanalyse, 95% KI).....	112

Tab. 61:	Die Änderungssensitivität über die SRM in der Gesamtstichprobe.....	113
Tab. 62:	Die Häufigkeiten der Polla Items (Follow-up).....	114
Tab. 63:	Mittelwerte und Standardabweichungen zu den Summenskalen (Follow-up).....	114
Tab. 64:	Schwierigkeitsindex zu den Polla Items zu $t_1$ , $t_2$ und zum Follow-up.....	116
Tab. 65:	Veränderungen der Polla Items zum Follow-up gegenüber dem Therapieende.....	117
Tab. 66:	Die deskriptive Statistik zu den Messzeitpunkten $t_1$ , $t_2$ und zum Follow-up.....	118
Tab. 67:	Ergebnisse des T-Tests auf Stichprobenunterschiede bei gepaarten Tests ( $t_2$ /Follow-up).....	118
Tab. 68:	Die deskriptive Statistik bei gepaarten Stichproben.....	118
Tab. 69:	Korrelationen nach Spearman-Rho zum Follow-up ( $n=26$ ).....	119
Tab. 70:	Korrelationen zwischen den Erhebungsinstrumenten.....	120
Tab. 71:	Deskriptive Gruppenstatistiken.....	120
Tab. 72:	Unterschiede zwischen den Gruppen „Gesunde“ und „Follow-up“ (Varianzanalyse, 95% KI).....	121
Tab. 73:	Die Änderungssensitivität in der Gruppe mit Follow-up zu $t_1$ , $t_2$ und zum Follow-up.....	122
Tab. 74:	Überblick über die Testgütekriterien zum Polla und den gemessenen Tests zu $t_1$ und $t_2$ .....	123

## ABKÜRZUNGEN UND ERLÄUTERUNGEN

Polla:	<i>Performance-Test of lower limb activities:</i> Er umfasst die Überprüfung von 14 Beinaktivitäten durch die Beurteilungskriterien „uneingeschränkt“ (2), „eingeschränkt“ (1) oder „nicht möglich“ (0) (ANHANG 1).
Gemessene Tests:	Die Tests umfassen die Messung des 1-Minute-Up & Go-Test nach der Anzahl, des 2-Minuten-Gehtests nach der Gehstrecke in Metern, des Treppenlaufen-Tests in Sekunden und des 2-Minuten-Lauftests nach der Laufstrecke in Metern.
Polla modifiziert:	Die nach der Validierung für den Praxiseinsatz modifizierte und empfohlene Version des Polla (ANHANG 7).
Selbsteinschätzung:	Der Fragebogen zur Selbsteinschätzung von Einschränkungen zu 27 Aktivitäten an der unteren Extremität (ANHANG 4).
SF-36-Fragebogen:	Die Gesundheitsbezogene Lebensqualität mit 36 Fragen (ANHANG 2).
KÖFU:	Körperliche Funktionsfähigkeit
KÖRO:	Körperliche Rollenfunktion
SCHM:	Schmerz
AGES:	Allgemeine Gesundheitswahrnehmung
VITA:	Vitalität
SOFU:	Soziale Funktionsfähigkeit
EMRO:	Emotionale Rollenfunktion
PSYC:	Psychisches Wohlbefinden
Medikation:	Die Art und Häufigkeit der Medikamenteneinnahme.
Hauptproblem:	Das subjektiv vom Patienten angegebene Hauptproblem zu $t_1$ und $t_2$ : 1 = ständig; 8 = nie.
NRS:	Die vom Patienten beschriebene Schmerzintensität

nach der Numeric-Rating-Scale zu  $t_1$  und  $t_2$ :

0 = kein Schmerz; 10 = maximal vorstellbarer Schmerz

Schmerzdauer: Die vom Patienten subjektiv angegebene Schmerzdauer zu  $t_1$  und  $t_2$ :

1 = ständiger Schmerz; 6 = kein Schmerz.

Giving-Way Phänomen: Das vom Patienten angegebene spontane Wegsacken im Kniegelenk zu  $t_1$  und  $t_2$ :

1 = mehrmals täglich; 5 = kein Giving-Way.

Aktive Flexion: Aktive Beweglichkeitsmessung nach der Neutral-Null-Methode (Debrunner, 1971) am betroffenen Gelenk (Hüfte, Knie, oberes Sprunggelenk) in Rückenlage in Grad gemessen zu  $t_1$  und  $t_2$ .

Schmerzfreie Kniebeuge: Wie weit kann der Patient in die Knie gehen, ohne dass Schmerzen im Knie auftreten. Gemessen wird mit dem Winkelmesser in Grad am Kniegelenk zu  $t_1$  und  $t_2$ .

# 1 EINLEITUNG

Unter dem Druck knapper werdender Ressourcen im Gesundheitssystem gewinnen Überlegungen zur Erhöhung der Effizienz bisher vernachlässigter Arbeits- und Forschungsbereiche im gesundheitlichen Versorgungssystem an Akzeptanz. Ursachen dafür sind neben einer demografischen Entwicklung hin zu einer Bevölkerung höheren Alters die Zunahme chronischer Erkrankungen (Koch & Buschmann-Steinhage, 2004) ebenso wie die Einführung kostenintensiver apparativer Medizin und neuer Behandlungsmethoden mit den Folgen von Kostensteigerungen. In diesem Zusammenhang sind die Begrifflichkeiten der *Evidenzbasierten Medizin* (EBM, Sackett & Kunz, 1999) oder der *Qualitätssicherung* auch in der Physiotherapie nicht mehr wegzudenken (z.B. Cabri, 2001; Scherfer, 2001; Vandenboorn, 2000; Hallmann, 1999).

Unter *Evidenzbasierter Praxis* (EBP) wird dabei der bewusste, explizite und urteilsfähige Einsatz der derzeit besten Beweise in der klinischen Entscheidungsfindung für individuelle Patienten verstanden (Sackett & Kunz, 1999). Dabei handelt es sich um einen Lösungsansatz für das Problem, das in der Forschung entstandene Wissen in die Praxis zu transferieren. EBP an sich ist keine Wissenschaft, sondern eine Fähigkeit, mit wissenschaftlichen Ergebnissen in der eigenen therapeutischen Praxis systematisch angewandt umgehen zu können. Dies sollte jeder Physiotherapeut<sup>1</sup> in Zukunft beherrschen (Fransen & de Bruin, 2000). EBP leistet im Hinblick auf die Evaluation effektiver Behandlungsmethoden durch *randomisierte Studien* wichtige Hinweise (Fransen & de Bruin, 2000, de Bie, 1998a/b). Mit der Operationalisierung von Forschungsfragen auf Teilaspekte – wie sie in randomisierten Studien obligat sind – können und wollen selbst hochwertige Studienergebnisse keine umfassenden Antworten auf die vielen Fragen geben, welche die Komplexität der physiotherapeutischen Behandlungslandschaft mit ihren unterschiedlichsten Krankheitsbildern und Symptomenkomplexen sowie den ethischen Hindernissen bei der Planung und Durchführung von Untersuchungen ausmachen. Das Aufdecken von Effektivität in der Breite kann daher in absehbarer Zeit nicht geleistet werden.

Umso wichtiger erscheint der Aspekt der *Ergebnisevaluation* von Therapie, um die Sinnhaftigkeit der konkreten Behandlungspraxis zu überprüfen und zu dokumentieren. Entsprechend empfahl die Gesundheitsministerkonferenz 1999, dass alle Einrichtungen des Gesundheitswesens bis zum 1. Januar 2005 ein an dem Stand der Wissenschaft und Technik orientiertes *Qualitätsmanagement* einführen sollten.

<sup>1</sup> Aus Gründen der besseren Lesbarkeit wird in dieser Arbeit auf eine geschlechtsbezogene Differenzierung verzichtet. Mit diesem Vorgehen ist keine Wertung verbunden, die weiblichen Personen sind immer gleichfalls gemeint.

Per Gesetz im §135a SGB V festgeschrieben sind Leistungserbringer zur Sicherung und Weiterentwicklung der Qualität und der von ihnen erbrachten Leistungen verpflichtet, was mit einer *Dokumentationspflicht* einherging, die in der Physiotherapie aus den nachfolgend genannten Gründen nicht unproblematisch erscheint:

1. Zum einen zeigt sich eine gewisse Orientierungslosigkeit, aber auch Unwilligkeit zur Beschäftigung mit Dokumentation und Berichterstattung. *Orientierungslosigkeit*, weil verbindliche Standards für die Beurteilung des Erfolges fehlen, nur unzureichend bekannt sind oder die systematische Sammlung, Auswertung und eine für alle an der Behandlungskette Beteiligten anschauliche Präsentation der Ergebnisse nicht praktiziert wird. *Unwilligkeit*, weil aufwändiger Dokumentation auf therapeutischer Seite häufig mit dem Argument begegnet wird, die Zeit solle besser dem Patienten zu Gute kommen (Broda & Beckmann, 2000).
2. Zudem sprechen viele Therapieformen in der Dokumentation unterschiedliche Sprachen, so dass sie selbst von Physiotherapeuten nicht immer verstanden werden. Darüber hinaus hängt das was als Ergebnis dokumentiert wird im Wesentlichen von dem einzelnen Therapeuten ab und kann bei der Beurteilung ein und desselben Patienten durch verschiedene Therapeuten zu unterschiedlichen Ergebnissen und Empfehlungen führen (Wirtz, 2004). So sind folgende Formulierungen als Fazit von Therapie nicht selten:
  - a. „Die Schwellung des Gelenkes ist noch stark, die Therapie sollte zur Reduktion derselben fortgesetzt werden“, oder  
die Äußerung eines zweiten Therapeuten über den selben Patienten:
  - b. „Die Schwellung ist noch stark, sie ist jedoch weder therapeutisch beeinflussbar noch behindert sie den Patienten im Alltag; die Therapie kann daher abgeschlossen werden“.
3. Schließlich beurteilen Physiotherapeuten häufig immer noch übermäßig Maße der Körperstruktur, wie z.B. Schwellung, Rötung, Atrophie oder Beweglichkeit und lösen sich zu wenig von der strukturellen Ebene. Diese bilden zwar eine wichtige Grundlage, geben aber auf die Frage nach verbleibenden Behinderungen in Alltag und Beruf nur unzureichende Antworten.

In diesem Zusammenhang fällt der Blick auf die Diskussion und Anwendung von *standardisierten Tests* und *Assessmentverfahren* zur Überprüfung von Ergebnissen als zentralem Aspekt einer Qualitätssicherung von Therapie.

Die in den neunziger Jahren zunächst in der medizinischen Rehabilitation aufgekommene Diskussion zu Assessmentverfahren hat seit wenigen Jahren auch die

Physiotherapie erreicht (ZVK, 2006a). Ein im Jahr 2006 vom Zentralverband für Physiotherapeuten (ZVK) der Öffentlichkeit vorgestelltes Handbuch zu Testverfahren hält zahlreiche Verfahren für die Überprüfung von Therapieergebnissen bereit.

Weitere Empfehlungen zur Messung von Therapieergebnissen sind vielfältig (z.B. Biefang & Schuntermann, 2000; Biefang, Potthoff & Schliehe, 1999; Krämer & Maichl, 1993; IQPTR, 2006; IQPR, 2006). Neben einfachen Parametern zur Beurteilung von Strukturen und Funktionen wie beispielsweise der Beweglichkeitsmessung von Gelenken, der Umfangmessung von Extremitäten oder Muskelfunktionsüberprüfungen, gibt es Verfahren wie z.B. isokinetische Krafttestungen, elektronische Geräte zur Messung der Beweglichkeit (z.B. der Wirbelsäule), Messplatten zur Beurteilung der Stand- und Sprungkoordination oder EDV-gestützte Kraftmessanlagen an Geräten der medizinischen Trainingstherapie. Sowohl der finanzielle als auch der räumliche Aufwand für diese Testverfahren – von vielen physiotherapeutischen Einrichtungen nicht leistbar – steht in keinem Verhältnis zu dem geringen Informationsgehalt der oft isolierten Ergebnisse. Kann nicht beispielsweise ein gut validierter Fragebogen, ein einfacher standardisierter Rückenkrafttest oder ein Funktionstest zur Überprüfung von Aktivitäten gleichfalls gute oder sogar inhaltlich breiter angelegte Ergebnisse als eine kostenintensive EDV-gestützte Messanlage hervorbringen? In der Praxis der Physiotherapie haben meist Ökonomie und Praktikabilität von Testverfahren Priorität.

Defizite hinsichtlich einfacher und kostengünstiger Testverfahren zeigen sich insbesondere im Zusammenhang mit Funktions- oder Leistungstests, denen auch so genannte *Performance-Tests* zuzuordnen sind (Bührlen, Gerdes & Jäckel; 2002). Bührlen et al. (2002, S. 77) sind davon überzeugt, „dass ein praktikabler Performance-Test ein wichtiges Instrument für die Erfassung eines der wichtigsten Therapieziele in der Rehabilitation, nämlich die Funktionsfähigkeit im Alltag, darstellt“, gegenwärtig aber nicht vorhanden ist.

Andere in der Literatur aufgeführte Verfahren sind häufig zu aufwändig, umfassen den gesamten Bewegungsapparat und benötigen zwei Tage zur Überprüfung (Kaiser, Kersting & Schian, 2000; Isernhagen, 1988) oder zielen auf altersmäßig eng begrenzte Zielgruppen, wie es nicht dem Klientel einer ambulanten physiotherapeutischen Einrichtung entspricht.

Die Aufgabe der vorliegenden Arbeit besteht daher in der Konstruktion und Validierung eines Performance-Tests für die untere Extremität.

Gefördert wurde das Projekt aus den Mitteln des Norddeutschen Verbunds für Rehabilitationsforschung. Der Performance-Test – folgend als *Performance-Test of lower limb activities* (Polla) bezeichnet – soll zu Beginn einer physiotherapeutischen Routinebehandlung, gegen Therapieende nach sechs Wochen und zu einem Follow-up eingesetzt werden, um den Funktionsstand von Patienten zu überprüfen.

Die sich daraus ableitenden Fragestellungen, das Ziel und die weitere Vorgehensweise werden im Folgenden dargelegt.

## 1.1 Fragestellung

Der praktisch arbeitende Physiotherapeut erwartet im Verlauf einer Behandlung Veränderungen, die seinen Patienten einem gemeinsam ausgehandelten Therapieziel näher bringen. Zu erwarten ist, dass sich Einschränkungen bei Alltagsaktivitäten nach sechs Wochen physiotherapeutischer Behandlung verringern und der Patient aus seiner subjektiven Sicht wieder besser in der Lage ist, seinen regulären Alltag zu bewältigen. Aus Therapeutensicht sollten die Veränderungen zur Verlaufskontrolle dokumentierbar sein, um sowohl Begründungen des klinischen Handelns (Clinical-Reasoning) im Hinblick auf die weitere Behandlungsstrategie, als auch zur Rechtfertigung gegenüber dem behandelnden Arzt deutlich zu machen. Ein wichtiger Aspekt bei der Evaluation von Therapie-Outcome beschäftigt sich daher mit der Frage von Veränderungen zwischen verschiedenen Messzeitpunkten. Kann ein Instrument unter Einbezug schwieriger Krankheitsbilder die mit kleineren Fortschritten im Genesungsprozess einhergehenden Veränderungen statistisch belegen?

Eine weitere Fragestellung zielt auf die möglichen Zusammenhänge zwischen der objektiven Überprüfung von Aktivitäten durch den Polla und der subjektiven Selbsteinschätzung der Lebensqualität. Geht also beispielsweise nach einem bestimmten Behandlungszeitraum mit der Verlängerung der Gehstrecke oder der Verkürzung der Zeit beim Treppensteigen eine Verbesserung des Allgemeinen Gesundheitszustands einher?

Weitere zentrale Fragestellungen der Arbeit beschäftigen sich mit den Testgütekriterien zum Polla und seiner Fähigkeit Veränderungen im Therapieverlauf zu verschiedenen Zeitpunkten abzubilden. Entsprechend den Fragestellungen resultiert daraus das nun folgende Ziel der Arbeit.

## 1.2 Ziel der Arbeit

Ziel der Arbeit ist die Konstruktion und Evaluation eines Performance-Tests. Im Hinblick auf die Anforderungen der Testentwicklung und seiner späterer Implementierung in die Arbeitsprozesse von Physiotherapeuten soll der Polla ohne großen finanziellen, materiellen und räumlichen Aufwand innerhalb einer Behandlungszeit

von maximal 20 Minuten durchführbar sein. Zielgruppe sind, so wie sie sich im physiotherapeutischen Alltag darstellen, unterschiedliche orthopädisch-traumatologische Indikationsgruppen.

Die Evaluation des Performance-Tests beschäftigt sich zum einen mit der Fragestellung von Zusammenhängen zwischen den Ergebnissen des Performance-Tests, der Selbsteinschätzung zu Aktivitäten und der gesundheitsbezogenen Lebensqualität sowie der Klärung der Frage, ob der Polla in der Lage ist, Veränderungen zwischen verschiedenen Messzeitpunkten aufzuzeigen. Darüber hinaus hat die Studie die Überprüfung von Testgütekriterien zum Ziel.

Der Polla kann als standardisiertes Therapieurteil die Grundlage für leicht handhabbare und verständliche Berichte im praktischen Arbeitsalltag sein. Alle beteiligten Personen im Gesundheitswesen könnten im Sinne einer gemeinsamen Sprache den Entwicklungsstand des Patienten in Anlehnung an Einschränkungen von Aktivitäten im Alltag nachvollziehen. Für Therapeuten können die Ergebnisse bei der Formulierung von Behandlungszielen im klinischen Handeln behilflich sein.

Der weitere Aufbau der vorliegenden Arbeit beinhaltet im Kapitel 2 die Darstellung des *theoretischen Bezugsrahmens*. Die *biomedizinische Sichtweise* wird gegenüber der *Salutogenese* als der ganzheitlichen Sichtweise abgegrenzt und mit der Internationalen Klassifikation der Funktionsfähigkeit und Behinderung der Weltgesundheitsorganisation (ICF) als Instrument der Beschreibung von Behinderung aus ganzheitlicher Sicht ergänzt. Die ICF bedarf als Basis eines theoretischen Ansatzes in der Rehabilitation einer umfassenden Betrachtung.

Aus dem gegenwärtigen *Stand der Forschung in Physiotherapie und Rehabilitation* wird ein *Theoriemodell der Rehabilitation* nach Gerdes & Weis (2000) vorgestellt, in dem sich neben der ICF die *Therapieziele in Physiotherapie und Rehabilitation* sowie deren Überprüfung durch Testverfahren oder Performance-Tests übergeordnet verankern lassen. An die Beschreibung von Zieldimensionen in der Rehabilitation schließt sich die Erläuterung von *Grundlagen zu Tests und Assessmentverfahren* an, mit der das Outcome von Therapie erfasst wird. Für die Konstruktion des Polla werden die *Bedeutung von Performance-Tests als Outcome-Parameter* sowie ihre *methodologische Problematik* erläutert. Anhand der aktuellen Literatur wird das Thema Performance-Tests erläutert.

In Kapitel 3 der Methode werden zunächst die Ergebnisse einer Pilotstudie und die Konsequenzen für die Entwicklung einer endgültigen Testbatterie dargelegt. Es werden das Forschungsdesign und das methodische Vorgehen erläutert sowie operationalisierte Ziele benannt. Die Auswahl des Patientenkollektivs und der Messinstrumente schließen sich an.

Kapitel 4 beschreibt die *praktische Untersuchungsdurchführung* mit dem *Erhebungsplan*, den *Merkmale der untersuchten Stichprobe* sowie die *statistische Auswertestrategie*.

Vor der *Datenaufbereitung* durch die Bildung von Summenskalen zu den Messinstrumenten in Kapitel 5 klären eine *Itemstatistik* und eine *Faktorenanalyse* zum Polla sowie zur Selbsteinschätzung zu Aktivitäten die *Itemaufgabenschwierigkeit* und eine mögliche inhaltliche Dimensionierung einzelner Tests.

Kapitel 6 präsentiert die *deskriptive Statistik zu Therapiebeginn und Therapieabschluss*. In der *schließenden Statistik* werden die Veränderungen der über die Messinstrumente erhobenen Ergebnisse zu Therapieende gegenüber dem Untersuchungsbeginn betrachtet.

Es schließt sich die *Überprüfung von Testgütekriterien zum Performance-Test* an. Geklärt wird, ob unterschiedliche Beurteiler bei zeitgleicher Überprüfung zu ein und demselben Ergebnis kommen und inwieweit dieselben Beurteiler bei Testwiederholungen innerhalb kurzer Zeitabstände ebenfalls gleiche Ergebnisse erzielen. Ferner werden die Messgenauigkeit (*Konsistenzanalyse*) und die statistische Validität über die *Konstruktvalidität* und *Kriteriumsvalidität* im Vergleich mit einem Außenkriterium (SF-36) bestimmt. Schließlich werden die Darstellungen zur *Änderungssensitivität* und *Unterscheidungsfähigkeit* der Tests zwischen Patienten und Gesunden betrachtet.

In Kapitel 7 werden die zentralen Hypothesen diskutiert. Einige Ergebnismerkmale können mit Befunden bereits vorliegender Ergebnisse anderer Untersuchungen abgeglichen werden.

Kapitel 8 werden Schlussfolgerungen gezogen und Empfehlungen für den klinischen Praxiseinsatz des Polla gegeben. Die Empfehlungen für den Praxiseinsatz beinhalten zur Verbesserung von Zeitökonomie und Praktikabilität eine modifizierte Form des evaluierten Polla. Kapitel 9 fasst die Ergebnisse zusammen.

## 2 THEORETISCHER BEZUGSRAHMEN

### 2.1 Salutogenese und biomedizinische Sichtweise

Zu Beginn des 19. Jahrhunderts entwickelte sich unter dem Einfluss naturwissenschaftlichen Denkens ein Krankheitsverständnis, das als biomedizinisches Krankheitsmodell bezeichnet wird (Faltermaier, 1994) und bis in die siebziger Jahre das Verständnis von Krankheit und Gesundheit prägte. Im Mittelpunkt standen die Beschwerden, Symptome oder Schmerzen des Patienten. Alle Anstrengungen des medizinischen Systems, der Ärzte und der Therapeuten richteten sich auf die Diagnose und das möglichst schnelle Beseitigen der Symptome und Beschwerden, wobei die Erwartungen des Patienten an die Möglichkeiten des medizinischen Versorgungssystems hoch waren. Trotz beeindruckender Erfolge in Diagnostik und Therapie vieler Erkrankungen wuchs in den letzten Jahren die Kritik an der so genannten Apparatedizin und der primären Orientierung an Symptomen. Unter dem Eindruck einer immer stärkeren Technisierung der Medizin wurde die Vernachlässigung der Person, also die Vernachlässigung der Ganzheitlichkeit beklagt. Die Kritik an unserem medizinischen Versorgungssystem geht einher mit einer Diskussion um den Gesundheits- und Krankheitsbegriff und mit einem komplex gewordenen Verständnis davon, was Krankheit verursacht und eine Heilung oder Linderung ermöglicht.

Es gibt eine Vielzahl von Ansätzen zur Definition von Gesundheit und Krankheit, die sich an unterschiedlichen Gesundheitsnormen orientieren (Bengel, 2001). Die World Health Organization (WHO, 1948) definiert „Gesundheit“ als eine Idealnorm in Bezug auf das vollkommene körperliche und psychische Wohlbefinden. In diesem Zusammenhang war es vor allem das Verdienst von Engel (1979), die biomedizinische Sichtweise um die der psychosozialen Komponente zu erweitern und Gesundheitsstörungen auf der Grundlage eines bio-psycho-sozialen Hintergrunds zu analysieren.

Der amerikanisch-israelische Medizinsoziologe Aaron Antonovsky (1923-1994) kritisierte eine rein pathogenetisch kurative Sichtweise und stellte dieser ein Konzept gegenüber, das nicht die Frage nach den Ursachen von Krankheiten in den Vordergrund rückt, sondern die Frage formulierte, warum und wie Personen trotz verschiedener krankheitserregender Bedingungen gesund bleiben. Antonovskys Vorstellung über die Entstehung von Gesundheit ist von systemtheoretischen Überlegungen beeinflusst: Gesundheit ist kein normaler, passiver Gleichgewichtszustand, sondern ein labiles, aktives und sich dynamisch regulierendes Geschehen, das in ständiger Auseinandersetzung mit krank machenden Einflüssen immer wieder neu aufgebaut wird. Das Grundprinzip menschlicher Existenz ist nicht

Gleichgewicht und Gesundheit, sondern Ungleichgewicht, Krankheit und Leiden. Unordnung und die Tendenz zu mehr Entropie sind allgegenwärtig: „Der menschliche Organismus ist ein System und wie alle Systeme der Kraft der Entropie ausgeliefert“ (Antonovsky, 1993, S. 7). Die Auffassung von den Konstrukten „Gesundheit“ und „Krankheit“ wird von Antonovsky (1987, S. 16f.) in seinem Kerngedanken als Kohärenzgefühl (sense of coherence, SOC) mit seinen drei Komponenten formuliert:

1. *„Verstehbarkeit“*: Das Ausmaß, in dem man die aus der internen und externen Umgebung stammenden Reize, mit denen man konfrontiert ist, als kognitiv sinnvoll sowie als geordnete, konsistente und strukturierte Information wahrnimmt.
2. *Machbarkeit*: Das Ausmaß, in dem man wahrnimmt, dass die einem zur Verfügung stehenden Ressourcen geeignet sind, den Anforderungen durch die einströmenden Reize zu entsprechen.
3. *Sinnhaftigkeit*: Das Ausmaß, in dem man das Gefühl hat, dass das Leben einen emotionalen Sinn hat, dass zumindest einige Probleme und Anforderungen, die das Leben einem auferlegt, Herausforderungen sind, die es Wert sind Energie einzusetzen, sich zu verpflichten und zu engagieren.“

Antonovsky (1997) betrachtet Gesundheit als ein Kontinuum zwischen zwei Polen und ist sich dabei im Klaren, dass äußere Faktoren wie Krieg, Hunger oder schlechte hygienische Verhältnisse die Gesundheit gefährden (Bengel, 2001).

Dennoch gibt es auch unter gleichen äußeren Bedingungen Unterschiede im Gesundheitszustand verschiedener Menschen. Wenn die äußeren Bedingungen vergleichbar sind, dann wird es nach Auffassung Antonovskys von der Ausprägung der individuellen, sowohl kognitiven als auch affektiv-motivationalen Grundeinstellung abhängen, wie gut Menschen in der Lage sind, vorhandene Ressourcen zum Erhalt ihrer Gesundheit und ihres Wohlbefindens zu nutzen. Je ausgeprägter das Kohärenzgefühl einer Person ist, desto gesünder sollte sie sein bzw. desto schneller sollte sie gesund werden und bleiben (Bengel, 2001).

Eine wie oben beschriebene und über die biomedizinische hinausgehende ganzheitliche bio-psycho-soziale Betrachtungsweise des Menschen erfordert ein Beschreibungsinstrument zur Erfassung der komplexen Zusammenhänge von Gesundheit und Krankheit des Menschen in seiner jeweilig verflochtenen kulturellen Umgebung mit den ihn umgebenden Umweltfaktoren. Die Weltgesundheitsorganisation (WHO) hat sich dieses mit der Entwicklung der Internationalen Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF) zur Aufgabe gemacht.

## 2.2 Internationale Klassifikation der Funktionsfähigkeit

Die Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF) wurde im Jahr 2001 von der Vollversammlung der Weltgesundheitsorganisation (WHO) verabschiedet (Schuntermann, 2006). Das wichtigste Ziel ist es,

„eine gemeinsame Sprache für die Beschreibung der funktionalen Gesundheit zur Verfügung zu stellen, um die Kommunikation zwischen Fachleuten im Gesundheits- und Sozialwesen, insbesondere in der Rehabilitation, sowie den Menschen mit Beeinträchtigungen ihrer Funktionsfähigkeit zu verbessern. Darüber hinaus stellt sie ein systematisches Verschlüsselungssystem für Gesundheitsinformationssysteme bereit und sie ermöglicht Datenvergleiche zwischen Ländern, Disziplinen im Gesundheitswesen, Gesundheitsdiensten sowie im Zeitverlauf“ (Schuntermann, 2006, S. 9)

Mit dem zentralen Aspekt der funktionalen Gesundheit greift sie die Beschreibung bio-psycho-sozialer Komponenten von Gesundheit im Sinne einer Ressourcenorientierung auf. Danach gilt eine Person als funktional gesund, wenn vor dem gesamten Lebenshintergrund (Schuntermann, 2006, S. 2):

1. „ihre körperlichen Funktionen (einschließlich des geistigen und seelischen Bereichs) und ihre Körperstrukturen allgemein anerkannten (statistischen) Normen entsprechen (Konzept der Körperfunktionen und -strukturen),
2. sie all das tut oder tun kann, was von einem Menschen ohne Gesundheitsproblem erwartet wird (Konzept der Aktivitäten), und
3. sie ihr Dasein in allen Lebensbereichen, die ihr wichtig sind, in der Weise und dem Umfang entfalten kann, wie es von einem Menschen ohne Beeinträchtigung der Körperfunktionen oder -strukturen oder der Aktivitäten erwartet wird (Konzept der Teilhabe an Lebensbereichen.“

Die ICF geht inhaltlich über die Aussage ihrer Vorgängerin, der ICIDH (International Classification of Impairments, Disabilities and Handicaps) von 1980 weit hinaus. Ihre Beschreibungen beschränken sich auf die von Schäden, Fähigkeitsstörungen und Handicaps (Tab. 1). Die ICF stellt ein übergreifendes Konzept der funktionalen Gesundheit dar, welches im Grundmodell kein Krankheitsfolgenmodell, sondern die Komponenten von Gesundheit in Form von Behinderungen nicht nur defizitorientiert, sondern darüber hinaus auch ressourcenorientiert klassifiziert. So lassen sich gleichzeitig positive und negative Bilder der Funktionsfähigkeit erstellen.

Der größte Unterschied zur ICIDH besteht in den formalen Begrifflichkeiten zu Beeinträchtigungen der Funktionsfähigkeit unter expliziter Bezugnahme auf Kontextfaktoren. Unter *Kontextfaktoren* werden *Umwelt-* und *personbezogene Faktoren* (Kap. 2.3) verstanden, wobei der grundlegende Aspekt nicht die Schädigung, sondern der Störungsbegriff der Beeinträchtigung

- mit dem Konzept der Körperfunktionen und -strukturen und seinem Störungsbegriff der Schädigung (Funktionsstörung, Strukturschaden),

- mit dem Konzept der Aktivitäten und seinem Störungsbegriff der Beeinträchtigung einer Aktivität
- sowie mit dem Konzept der Partizipation oder Teilhabe mit dem Störungsbegriff der Beeinträchtigung der Partizipation (Teilhabe, Schuntermann, 2002).

Einen zusammenfassenden Überblick über die Unterschiede zwischen ICF und ICIDH gibt Tabelle 1.

Tab. 1. Unterschiede zwischen der ICF und der ICIDH (Schuntermann, 2002, S. 3).

	ICIDH	ICF
<b>Konzept:</b>	kein übergreifendes Konzept	Konzept der funktionalen Gesundheit (Funktionsfähigkeit)
<b>Orientierung:</b>	Defizitorientiert: Es werden Behinderungen klassifiziert.	Ressourcen- und defizitorientiert: Es werden Bereiche klassifiziert, in denen Behinderungen auftreten können. Es können unmittelbar positive und negative Bilder der Funktionsfähigkeit erstellt werden.
<b>Behinderung:</b>	formaler Oberbegriff zu Schädigungen, Fähigkeitsstörungen und (sozialen) Beeinträchtigungen. Keine explizite Bezugnahme auf Kontextfaktoren.	formaler Oberbegriff zu Beeinträchtigungen der Funktionsfähigkeit unter expliziter Bezugnahme auf Kontextfaktoren.
<b>grundlegende Aspekte:</b>	<ul style="list-style-type: none"> <li>– Schädigung</li> <li>– Fähigkeitsstörung</li> <li>– (soziale) Beeinträchtigung</li> </ul>	<ul style="list-style-type: none"> <li>– Körperfunktionen/-strukturen. Störungsbegriff: Schädigung (Funktionsstörung, Struktur-schaden)</li> <li>– Aktivitäten. Störungsbegriff: Beeinträchtigung der Aktivitäten</li> <li>– Partizipation [Teilhabe]. Störungsbegriff: Beeinträchtigung der Partizipation [Teilhabe]</li> </ul>
<b>soziale Beeinträchtigung:</b>	Attribut einer Person	Partizipation [Teilhabe] und deren Beeinträchtigung definiert als Wechselwirkung zwischen dem gesundheitlichen Problem (ICD) einer Person und ihren Umweltfaktoren.
<b>Umweltfaktoren:</b>	bleiben unberücksichtigt	Umweltfaktoren sind integraler Bestandteil des Konzept und werden klassifiziert
<b>personbezogene (persönliche) Faktoren:</b>	werden höchstens implizit berücksichtigt.	werden explizit erwähnt, aber nicht klassifiziert.
<b>Anwendungsbereich:</b>	nur im gesundheitlichen Kontext	

Die Beeinträchtigung der funktionalen Gesundheit einer Person ist immer das Ergebnis von Wechselwirkungen zwischen dem Gesundheitsproblem einer Person und seinen Kontextfaktoren und wird nach ICF mit dem Begriff der „Behinderung“ belegt.

Der für die Validierung eines Performance-Tests relevante Bereich der Konzepte bezieht sich insbesondere auf die gemeinsame Klassifikation der Aktivitäten und der Teilhabe und bedarf im folgenden Kapitel einer näheren Betrachtung.

## 2.3 Konzept der funktionalen Gesundheit

Auf der Basis der Konstruktion eines Performance-Tests für die untere Extremität bedarf das Konzept der funktionalen Gesundheit mit der Beschreibung seiner Aspekte zu *Körperstrukturen* und *-funktionen*, zu *Aktivitäten* sowie zur *Teilhabe* einer näheren Betrachtung.

Beeinträchtigungen sind immer vor dem Lebenshintergrund betroffener Personen zu sehen und mit dem Begriff der *Kontextfaktoren – umwelt- und personbezogene Faktoren* – belegt. „*Umweltfaktoren* bilden die materielle, soziale und einstellungsbezogene Umwelt ab, in der Menschen leben und ihr Dasein entfalten“ (WHO, 2006, S.3). Die Umweltfaktoren beschreiben:

Kapitel 1: Produkte und Technologien

Kapitel 2: Natürliche und vom Menschen veränderte Umwelt

Kapitel 3: Unterstützung und Beziehungen

Kapitel 5: Dienste, Systeme und Handlungsgrundsätze (WHO, 2006).

*Personbezogene Faktoren* einer Person (ihre Eigenschaften und Attribute) umfassen Gegebenheiten des Individuums, die nicht Teil ihres Gesundheitsproblems oder -zustands sind (WHO, 2006). Sind bestimmte Faktoren Teil des Gesundheitsproblems, wie z.B. mangelnder Handlungswille bei Depression, dann gehören sie nicht zu den personbezogenen Faktoren. Im Beispiel liegt eine Funktionsstörung vor (WHO, 2006, S. 3).

Die funktionale Gesundheit variiert mit dem Gesundheitsproblem und seinen gegenseitigen Abhängigkeiten durch die Konzepte mit den dazugehörigen Kontextfaktoren. Jedes dieser Elemente kommt dabei als Ausgangspunkt für eine mögliche Änderung der funktionalen Gesundheit in Betracht (WHO, 2006). Die Zusammenhänge zwischen den Aspekten mit ihren Wechselwirkungen wie sie in der Abbildung 1 dargestellt sind, seien an einem fiktiven Beispiel der klinischen Praxis verdeutlicht:

*Beispiel:* Ein Fliesenleger arbeitet tagsüber vornehmlich kniend (Aspekt der Strukturen) und klagt seit kurzer Zeit zunehmend über Schmerzen in beiden Kniegelenken (Aspekt der Funktionen). Eine klinische Untersuchung liefert Hinweise auf einen möglicherweise defekten Meniskus (Aspekt der Strukturen) und zeigt eine Schwellung des Kniegelenkes (Aspekt der Funktionen). Die Schmerzen (Konzept der Funktionen) führen unter Bewegungseinschränkungen (Aspekt der Funktionen) dazu, dass der Patient Beeinträchtigungen beim Knien (Aspekt der Aktivitäten) entwickelt, die insgesamt zu einer Behinderung der Berufsausübung führen (Aspekt der Teilhabe). Er ist sogar zeitweise arbeitsunfähig. Eine Veränderung umweltbezogener Faktoren, z.B. durch die Nutzung eines Kniekissens (Produkte und Technologien für die Erwerbstätigkeit), könnte u.U. die Reduktion der Problematik herbeiführen. Allerdings müssen nicht zwangsläufig Behinderungen im Bereich von Aktivitäten auftreten. So ist vorstellbar, dass der gleiche Patient bei Ausübung einer sitzenden Berufstätigkeit weder Beeinträchtigungen auf der Funktionsebene durch eine Schwellung, noch Behinderungen auf Aktivitätsebene im Hinblick auf das Knien aufweisen würde.

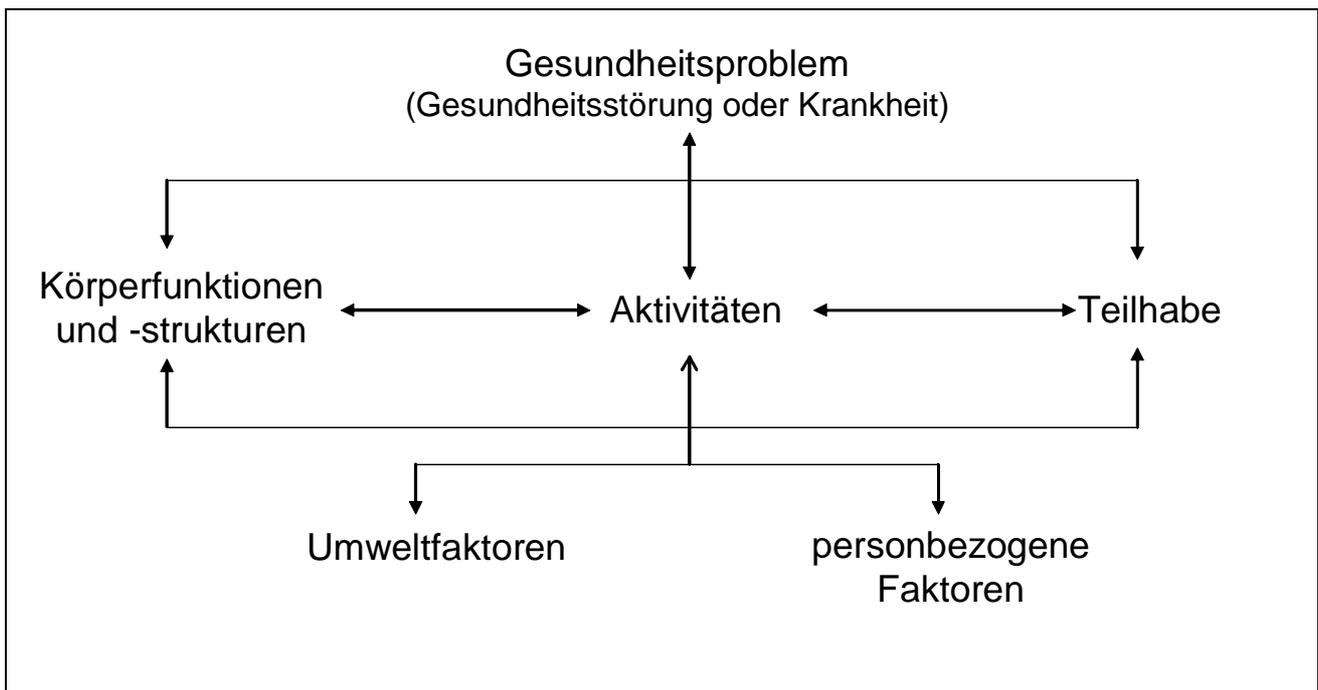


Abb. 1. Das bio-psycho-soziales Modell der ICF (WHO, S. 3, 2006).

Der Aspekt der *Körperstrukturen* beschreibt anatomische Teile des Körpers, wie Organe, Gliedmaßen, Knochen, Weichteile oder Strukturen innerer Organe und

neuronalen Systeme. *Körperfunktionen* beschreiben die physiologischen Funktionen von Körpersystemen (einschließlich psychologische Funktionen, WHO, 2006). Zusätzlich zu den biomedizinischen Aspekten der Körperstrukturen und -funktionen auf der Ebene des Organismus, werden Aspekte des Menschen als handelndes Subjekt im Sinne von *Aktivitäten* und als gleichberechtigtes Subjekt in Gesellschaft und Umwelt (*Teilhabe*) integriert und sind damit von zentraler Bedeutung für die Rehabilitation.

*Aktivitäten* sind von der *Teilhabe* nicht trennbar, sondern bilden in vielen Fällen die Voraussetzung, um in den Alltag wieder eingebunden werden zu können. So ist ein Patient nach einer traumatischen Schulterluxation nicht in der Lage – seinem Hobby – dem Handballspiel nachzukommen. Insofern präsentieren sich in der Klassifikation *Aktivitäten* und *Teilhabe* gemeinsam und umfassen nach WHO (2002) neun Kapitel:

Kapitel 1: Lernen und Wissensanwendung

Kapitel 2: Allgemeine Aufgaben und Anforderungen

Kapitel 3: Kommunikation

Kapitel 4: Mobilität

Kapitel 5: Selbstversorgung

Kapitel 6: Häusliches Leben

Kapitel 7: Interpersonelle Interaktionen und Beziehungen

Kapitel 8: Bedeutende Lebensbereiche

Kapitel 9: Gemeinschafts-, soziales und staatsbürgerliches Leben

Im Zusammenhang mit der unteren Extremität werden unter *Aktivitäten/Teilhabe* im Kapitel 4 und 5 die Domänen *Mobilität* und *Selbstversorgung* klassifiziert (WHO, 2002). Sie stellen die zentralen Aktivitäten für den Polla bereit, die ohne organisatorischen Aufwand überprüfbar sind. Sie werden nachfolgend ausgeführt.

#### **Kapitel 4: Mobilität**

Die Körperposition ändern und aufrechterhalten (d410-d429)

*d410 Eine elementare Körperposition wechseln:* In eine und aus einer Körperposition zu gelangen und sich von einem Ort zu einem anderen zu bewegen, wie von einem Stuhl aufstehen, um sich in ein Bett zu legen, in eine und aus einer knienden oder hockenden Position gelangen...

*d415 In einer Körperposition verbleiben:* In derselben erforderlichen Körperposition zu verbleiben, wie sitzen bleiben oder bei der Arbeit bzw. in der Schule stehen bleiben. Inklusive: In liegender, hockender, kniender, sitzender oder stehender Position verbleiben...

*d420 Sich verlagern:* Sich von einer Oberfläche auf eine andere zu bewegen, wie auf einer Bank entlang gleiten oder sich ohne Änderung der Körperposition aus dem Bett auf einen Stuhl bewegen. Inklusive: Sich während des Sitzens oder Liegens verlagern...

**Gegenstände tragen, bewegen und handhaben (d430-d449)**

*d430 Gegenstände anheben und tragen:* Einen Gegenstand anzuheben oder etwas von einem Platz zu einem anderen zu tragen, wie eine Tasse anheben oder ein Kind von einem Zimmer in ein anderes tragen...

**Gehen und sich fortbewegen (d450-d469)**

*d450 Gehen:* Sich zu Fuß auf einer Oberfläche Schritt für Schritt so fortzubewegen, dass stets wenigstens ein Fuß den Boden berührt, wie beim Spazieren, Schlendern, Vorwärts-, Rückwärts- oder Seitwärtsgehen...

*d455 Sich auf andere Weise fortbewegen:* Sich auf andere Weise als gehend von einem Ort zu einem anderen fortzubewegen, wie über einen Fels klettern oder eine Straße entlang rennen, springen, spurten, hüpfen, einen Purzelbaum schlagen oder um Hindernisse rennen...

*d460 Sich in verschiedenen Umgebungen fortbewegen:* In verschiedenen Orten und Situationen zu gehen und sich fortzubewegen, wie in einem Haus oder Gebäude von einem Raum in einen anderen gehen oder auf einer Straße einer Stadt gehen. Inklusiv: Sich in seiner Wohnung umherbewegen, in der Wohnung krabbeln oder (Treppen) steigen, in anderen Gebäuden als zu Hause bzw. außerhalb seiner Wohnung oder anderen Gebäuden gehen oder sich fortbewegen...

**Kapitel 5: Selbstversorgung**

*d540 Sich kleiden:* Die koordinierten Handlungen und Aufgaben durchzuführen, welche das An- und Ausziehen von Kleidung und Schuhwerk in Abfolge und entsprechend den sozialen und klimatischen Bedingungen betreffen, wie Hemden, Röcke, Blusen, Hosen, Unterwäsche, Saris, Kimonos, Strumpfhosen, Hüte, Handschuhe, Mäntel, Schuhe, Stiefel, Sandalen oder Slipper anziehen, ordnen und ausziehen... (DIMDI, 2005)

Weitere Kategorien zu *Mobilität* und *Selbstversorgung* beinhalten Aktivitäten, die sich einer objektiven Überprüfung unter der Therapie entziehen. Darunter fallen z.B. das „*sich in verschiedenen Umgebungen fortbewegen*“ (d460), das „*sich unter Verwendung von Geräten/Ausrüstung fortbewegen*“ (d465) oder das „*sich mit Transportmitteln fortbewegen*“ (d470-d489). Weiterhin finden sich Aktivitäten, die nicht im unmittelbaren Zusammenhang mit der unteren Extremität stehen und daher als nicht relevant erachtet werden. Sie beinhalten beispielsweise den „*feinmotorischen Handgebrauch*“ (d440) oder den „*Hand- und Armgebrauch*“ (d445).

Die ICF-Aktivitäten sind nur im Zusammenhang mit einer Beurteilung vollständig. Das *erste Beurteilungsmerkmal* beschreibt den Schweregrad eines Problems (DIMDI, 2005). Entsprechend kann ein Problem als

- „nicht vorhanden“,
- „leicht ausgeprägt“,
- „mäßig ausgeprägt“,
- erheblich ausgeprägt“,
- „voll ausgeprägt“,

„nicht spezifiziert“ oder als „nicht anwendbar“ beschrieben werden. Insofern stellt die ICF kein Assessmentverfahren dar. Sie bietet aber eine Grundlage für die Entwicklung oder Weiterentwicklung von Messinstrumenten (Schuntermann, 2006). Instrumente mit anderen Standards von Quantifizierungen, die sich häufig auch inhaltlich nicht den Dimensionen der ICF eindeutig zuordnen lassen, wären insbesondere im Hinblick auf die Zuweisung zu den *Beurteilungsmerkmalen* weiter zu entwickeln. Entsprechend bleibt es zukünftiger Forschung vorbehalten Verfahren zu entwickeln, welche systematisch die Items der ICF operationalisieren (Farin, Fleitz & Follert, 2006; Biefang & Schuntermann, 2000; Biefang, Potthoff & Schliehe, 1999).

Auch wenn die Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit in der praktischen Handhabung im klinischen Alltag noch Defizite aufzeigt, machen die folgenden Ausführungen zum Forschungsstand und zur Theoriengenerierung deutlich, dass die ICF sich sowohl in der Physiotherapie wie auch in der Rehabilitation zunehmend als theoretisches Grundgerüst durchsetzt.

## 2.4. Forschungsstand in Physiotherapie und Rehabilitation

### 2.4.1 Forschungsstand in der Physiotherapie

Seit der Eröffnung der ersten staatlich genehmigten Lehranstalt für Heilgymnastik, deren Gründung auf den Kieler Arzt J. H. *Lubinus* im Jahr 1900 zurückgeht, hat sich die Physiotherapie in ihrer nunmehr über 100-jährigen Geschichte erst in der jüngeren Vergangenheit der Akademisierung zugewandt. Ohne Akademisierung ist eine Forschungsdisziplin nicht vorstellbar. Einige Meilensteine der Entwicklung zu einer Therapiewissenschaft der Physiotherapie in Deutschland, sollen an dieser Stelle aufgezeigt werden (vgl. Klemme, Geuter & Willimczik, 2007).

- Mit der *Akademisierung* ist die *Physiotherapie* bemüht, die bisher vorherrschenden Strukturen des Status eines Hilfsheilberufes aufzuweichen und Wege in eine eigenständige Wissenschaftsdisziplin zu finden. Sie hinkt mit ihrer Entwicklung im internationalen Vergleich (England, Skandinavien, Holland, Belgien) um etwa 15-20 Jahre hinterher (Schämann, 2005). Seit dem Start des ersten Bachelorstudiengangs für Physiotherapie im Jahr 2001 sind weitere 12 Bachelor- und drei Masterstudiengänge etabliert worden (Werner, 2007). Die ersten Bachelor-abgänger haben im Rahmen des Abkommens von Bologna als „reflektierte Praktiker“ die Fachhochschulen im Jahr 2002 verlassen – seit 2005 sind die ersten Masterabschlüsse erreicht.
- Unter der Prämisse einer wissenschaftlichen Fundierung und Qualitätssicherung des berufspraktischen Handelns (Hochschulverbund der Gesundheitsfach-

berufe, HVG e.V., 2007), haben sich die deutschen Hochschulen mit den Studiengängen Physiotherapie, Ergotherapie und Logopädie im Jahr 2007 zu dem HVG zusammen geschlossen.

- Gegenwärtig beschäftigt sich die Physiotherapie mit *Fragen* nach dem, was *Professionalisierung und Akademisierung* ausmachen (vgl. z.B. Zimmermann, 2007; Bäumer, 2007; Schämnn, 2005) und mit einer Standortbestimmung zu ihren Bezugswissenschaften Medizin, Pädagogik und Psychologie, Sozial- und Gesundheitswissenschaften sowie zu den Rehabilitationswissenschaften. Zudem wird nicht zuletzt mit der *Zukunftsinitiative* (ZIPT) nach den neuen Tätigkeitsfeldern von Physiotherapie gefragt (ZIPT, 2007).
- Die Organisationen des *Bundesverbands selbstständiger Physiotherapeuten* (IFK, 2003) und des *Zentralverbands für Physiotherapeuten* (ZVK) haben die Förderung von Forschung durch die Einrichtung von Wissenschaftspreisen aufgenommen.
- Der *Zentralverband für Physiotherapeuten* mit der Gründung des Bildungswerkes der *Physio-Akademie* versteht sich (ZVK, 2006b)
  - „als Bildungsanbieter im Dienst von Physiotherapeutinnen und Physiotherapeuten, als Brücke zwischen herkömmlichen und akademischen Fort- und Weiterbildungen und als Förderer von Forschung und Evaluation in der Physiotherapie in einem internationalen Kontext. Sie strebt an, durch ihr Engagement im Bereich der Aus-, Fort- und Weiterbildung einerseits und im Bereich von Forschung und Evaluation andererseits zur Weiterentwicklung der deutschen Physiotherapie einen spürbaren Beitrag zu leisten.“
- Die Hilfestellungen reichen von einer Website zum Thema Forschung und Evaluation über eine ZVK-Stiftung (Stiftung zur Förderung von Forschung und Evaluation in der Physiotherapie) bis hin zu Beitragsreihen in den einschlägigen Fachzeitschriften zur Klärung grundlegender forschungsmethodologischer Fragen.
- Bezüglich der Fachzeitschriften für Physiotherapie ist ebenfalls eine intensive Entwicklung zu beobachten. In Deutschland kommt vor allem der *Zeitschrift für Physiotherapeuten* im 59. Erscheinungsjahr als auflagenstärkste Publikation eine zentrale Rolle in der Kommunikation von Entwicklungen in der Physiotherapie zu. Die Veränderungen in der Art ihrer Beiträge - hin zu mehr Wissenschaftlichkeit - sind nicht zuletzt als Reaktion auf die in jüngerer Zeit regelmäßigen Veröffentlichungen *Physiopraxis* und *physioscience* zu verstehen und steigern die Qualität wissenschaftlicher Veröffentlichungen in der Breite. „*physioscience*“ ist das deutschsprachige Forum für wissenschaftlich interessierte Physiotherapeuten ([www.thieme.de/physioonline/physioscience/index.html](http://www.thieme.de/physioonline/physioscience/index.html)) und bietet Original- und Übersichtsarbeiten. Die Veränderungen in den Publikationen finden ihren Ausdruck,

1. in einer insgesamt höheren wissenschaftlichen Qualität der Beiträge,
2. in vermehrten Beiträgen zur evidenzbasierten Praxis, z.B. durch Präsentation und Diskussion internationaler Studienergebnisse zu physiotherapeutischen Fragestellungen oder durch vereinzelte systematische Reviews,
3. in Beitragsreihen zu grundlegendem forschungsmethodologischen Handwerkszeug für physiotherapeutische Fragestellungen und
4. in einer breiten Diskussion über Akademisierung und Professionalisierung der Physiotherapie.

Damit werden insgesamt die Voraussetzungen geschaffen eine breitere Öffentlichkeit zu erreichen und die Erweiterung physiotherapeutisch-wissenschaftlichen Denkens und Handelns zu fördern.

*Physiotherapeutische Forschung* in Deutschland ist, was auch nicht verwundern mag, daher „eher die Ausnahme als die Regel“ (Witte, 2005, S. 58). Sie ist weitgehend noch unorganisiert und wird gegenwärtig thematisch von den Nachbardisziplinen vereinnahmt. Forschende Physiotherapeuten, strukturell überwiegend an Kliniken eingebunden, liefern häufig Teilergebnisse zu komplexeren Themenstellungen, deren Konzeptionen von Medizinern, Politologen, Soziologen, Sozialwissenschaftlern, Psychologen, Sportwissenschaftlern, Ökonomen oder Rehabilitationswissenschaftlern entwickelt werden.

Für eine künftig *eigenständige Wissenschaftsdisziplin* der Physiotherapie ist die Erfüllung bestimmter Kriterien erforderlich. Sie werden in der allgemeinen Wissenschaftstheorie traditionell seit Aristoteles und verstärkt von Kant bis heute gefordert und umfassen nach Klemme, Geuter und Willimczik (2007, S. 82)

- „Das Vorliegen eines eigenständigen, strukturierten Gegenstands;
- Das Vorhandensein von spezifischen Forschungsmethoden;
- Eine Systematik von Erkenntnissen (Theorien).

Als Nebenwirkungen gelten:

- Eine wissenschaftliche Fachsprache;
- Eine bedeutende Geschichte;
- Institutionalisierte Einrichtungen.“

Eine *bundesweite Vernetzung* wissenschaftlich arbeitender Physiotherapeuten im Sinne eines Forschungsverbundes wird gefordert (Schämann, 2003) und sollte sich mit dem Aufbau von Forschungsförderung und der Formulierung und Bearbeitung zentraler Forschungsschwerpunkte beschäftigen und an einer pragmatischen Imp-

lementierung von Ergebnissen in den Arbeitsalltag interessiert sein. Nur so gelingt langfristig eine Standortsicherung der Physiotherapie im Geflecht lange etablierter Wissenschaftsdisziplinen im Gesundheitswesen.

Ein übergreifender Blick auf die Nachbardisziplin der Sportwissenschaft, die mit der Gründung der *Deutschen Vereinigung für Sportwissenschaft (dvs, 2006)* bereits vor 30 Jahren ihre wissenschaftliche Entwicklung einleitete, zeigt Zielformulierungen, die auch der Physiotherapie zu wünschen wären. Die Vereinigung hat sich zum Ziel gesetzt,

1. die Forschung anzuregen und zu unterstützen,
2. gute wissenschaftliche Praxis auf der Basis berufsethischer Grundsätze zu sichern,
3. die Kommunikation zwischen verschiedenen Disziplinen zu verbessern,
4. die Lehre zu vertiefen und Beratung zu leisten,
5. zu Fragen von Studium und Prüfung Stellung zu nehmen,
6. den Nachwuchs zu fördern,
7. regionale Einrichtungen bei der Strukturentwicklung zu unterstützen,
8. die Personalstruktur wissenschafts- und zeitgerecht weiter zu entwickeln,
9. die Belange der Sportwissenschaft im nationalen und internationalen Bereich zu vertreten.

Wichtige *Fragestellungen in der Zukunft* sollten sich unter anderem mit den Entwicklungsmöglichkeiten einer künftigen Physiotherapie im Rahmen traditioneller und möglicherweise neuer Tätigkeitsfelder in Abhängigkeit eingeschränkter finanzieller Spielräume des Gesundheitssystems, demographischer Entwicklungen, epidemiologischer Analysen, rehabilitationswissenschaftlicher Ergebnisse und Entwicklungstrends im bewegungsorientierten, sich immer mehr verzahnenden Gesundheits-, Freizeit- und Fitnessmarkt beschäftigen.

Zu der Entwicklung einer physiotherapeutischen Wissenschaftsdisziplin gehört auch die Entwicklung theoriegeleiteten Handelns. Die Entwicklungstendenzen werden im nächsten Kapitel beschrieben.

## 2.4.2 Theoriengenerierung in der Physiotherapie

Vergleichbar dem im letzten Kapitel gezeichneten Stand als Wissenschaftsdisziplin befindet sich die Entwicklung theoriegeleiteten Handelns in der Physiotherapie noch in den „Kinderschuhen“. Eine von allen Seiten akzeptierte Theorie ist gegenwärtig nicht in Sicht, verschiedene Betrachtungsweisen sind aber in Ansätzen vorhanden.

Es fehlt ein Bezugsrahmen, der alle Physiotherapeuten in einer Sprache miteinander verbindet. Die Spezialisierung und Ausdifferenzierung des Wissens und der einzelnen therapeutischen Sprachen hat in der Physiotherapie zu einer so weiten Zersplitterung geführt, dass der gemeinsame Nenner des professionellen Selbstverständnisses sehr klein geworden ist (Klemme et al., 2007; ZIPT, 2003).

Pionierarbeit zu einer *Theorie der Physiotherapie* in Deutschland leistete Hüter-Becker seit 1997 zunächst mit der Entwicklung des *neuen Denkmodells zur Physiotherapie* und dem darauf aufbauenden Modell der *Integrativen Physiotherapie* (Hüter-Becker, 2002). Ausgangspunkt der Entwicklung war die starke Bindung an medizinische Fachgebiete, die dazu führte, „dass die Physiotherapie als eigenständiges Fach konturlos geblieben ist“ (Hüter-Becker, 1997, S. 565). Das *neue Denkmodell der Physiotherapie*, das über die klinischen Fachdisziplinen hinaus die eigenständige Position der Physiotherapie betont, entfaltet seine Wirkorte physiotherapeutischen Handelns an den vernetzten Funktionskreisen

- des Bewegungssystems,
- der Inneren Organe,
- der Bewegungsentwicklung und -kontrolle sowie
- dem Verhalten und Erleben.

In einer konsequenten Weiterentwicklung zur *Integrativen Physiotherapie* wird die Perspektive vernetzter Funktionskreise um die Integration des Patienten in das Alltagsleben mit seinen sozialen Bezügen erweitert und ist terminologisch wie inhaltlich an den salutogenetischen Denk- und Handlungsansatz und die *Internationale Klassifikation der Funktionsfähigkeit* (ICF) angelehnt (Hüter-Becker, 2002). Mit der Eingliederung in einen größeren Rahmen soll der Weg zur internationalen Rehabilitationswissenschaft aufgezeigt werden. *Neuere Veröffentlichungen zu Theorieansätzen* in der Physiotherapie beschäftigen sich:

1. mit der Begründung einer Therapiewissenschaft Physiotherapie auf der Grundlage wissenschaftstheoretischer Überlegungen und einer vorläufigen Klärung der Frage nach dem „*lebensweltlichen Gegenstand*“ von Physiotherapie (Klemme et al., 2007, S. 80). Dieser „lebensweltliche Gegenstand“ ist gekenn-

zeichnet einerseits durch *Wirkorte* (vgl. Hüter-Becker, 2002) und andererseits durch *Merkmale*, die Aktivitäten von Patienten (z.B. gezielte Übungen oder Training), Aktivitäten von Therapeuten (z.B. hands-on-Zeiten) sowie Aktivitäten im Rahmen der Interaktion zwischen beiden umfassen. Einen „*wissenschaftlichen Gegenstand der Physiotherapie*“ gilt es nach Klemme et al. (2007) noch zu bearbeiten.

2. mit der Bedeutung von Gesundheitsförderung und Prävention für eine künftige Physiotherapie. Höppner & Borgetto (2007, S. 676) äußern dazu:

„Für die Weiterentwicklung der Physiotherapie wird entscheidend sein, wie schnell und mit welchen Qualifikationen sich eine große Zahl von Physiotherapeuten diesem Bereich [Gesundheitsförderung und Prävention; Anmerkung des Autors] zuwenden wird.“

3. mit dem, was Physiotherapie heute bewegt und in Zukunft bewegen wird. Bäumer (2006; S. 317) stellt dazu „Kriterien einer künftigen Berufsausübung“ und „Kriterien eines klaren Berufsbildes“ heraus und formuliert für die Akademisierung der Physiotherapie auf Ebene der Fachhochschulen zwei Hauptziele, die sich mit der „Vorbereitung und Aneignung einer qualifizierten, wissenschaftlich fundierten“ Praxis der Physiotherapie und der „Grundlegung einer eigenständigen Wissenschaftlichkeit (Ebene der theoretischen Kompetenz)“ (Bäumer, 2006; S. 321) beschäftigen.
4. mit der Frage: „Was ist Physiotherapie? – Unser Selbstverständnis“ (ZIPT 2005). ZIPT stellt dabei das Individuum *Mensch* in der Therapeuten-Klienten-Beziehung sowie die *Bewegung* als zentralem Arbeitsmittel der Physiotherapie in den Vordergrund. An den Orten, wo Physiotherapie wirkt, lehnt sich ZIPT an Hüter-Becker (2002) mit dem neuen Denkmodell der Physiotherapie und ihren Funktionskreisen an. ZIPT manifestiert das Selbstverständnis der Physiotherapie mit der subjektiven Untersuchung, der Funktionsuntersuchung, dem Therapieplan, der Intervention und der Evaluation im eigentlichen *Physiotherapeutischen Prozess*.

Die Ansatzpunkte für theoriegeleitetes Handeln sind also mehrdimensional. Ein Blick auf die Entwicklungen in den Rehabilitationswissenschaften – einer weiteren Nachbardisziplin neben den Sportwissenschaften – könnten zur Findung eines wissenschaftlichen Selbstverständnisses der Physiotherapie beitragen.

### **2.4.3 Forschungsstand in der Rehabilitation**

Die Rehabilitationswissenschaften – eine ebenfalls noch junge Wissenschaftsdisziplin – haben im Vergleich zur Physiotherapie eine starke Vernetzung von

Wissenschaftlern unterschiedlichster Wissenschaftsdisziplinen mit verschiedenen rehabilitationswissenschaftlichen Forschungsschwerpunkten an universitären Einrichtungen.

Die Rehabilitationswissenschaften unterlagen in den vergangenen 15 Jahren einem starken Wandel (Schliehe & Jäckel, 2004), wobei nach Meinung von Experten einhellig eine erhebliche Diskrepanz zwischen den Aufwendungen für Rehabilitationsmaßnahmen und den Bemühungen ihrer wissenschaftlichen Fundierung bestand (Koch & Buschmann-Steinhage, 2004). Ein Entwicklungsschub geht auf die frühen neunziger Jahre zurück. Zunächst wurden rehabilitationswissenschaftliche Fragestellungen in angrenzenden Fachbereichen der Sozialmedizin, der Psychologie oder der Physikalischen Medizin bearbeitet (Koch & Bengel, 2000). Eigene Spezialabteilungen gab es bis auf wenige Ausnahmen nicht. Die *Deutsche Sporthochschule Köln* (DSHS) bildete mit einer Professur für Rehabilitation hier die Ausnahme. Ab Mitte der neunziger Jahre wurde mit der Einrichtung von Stiftungsprofessuren an verschiedenen Orten begonnen.

Das gewachsene Interesse an Rehabilitationsforschung (Koch, Schliehe & Aufderheide, 1998; Haaf & Schliehe, 2000) geht auf die veränderte Interessenslage seitens der Gesundheitspolitik, der Rentenversicherungsträger und der Rehabilitationskliniken infolge knapper werdender Ressourcen und dem erheblich wachsenden Anteil chronischer Erkrankungen mit langfristig progredienten Verläufen zurück (Koch & Buschmann-Steinhage, 2004). Koch und Buschmann-Steinhage (2004) benennen strukturelle Defizite, die gegen Ende der achtziger Jahre zunehmend in die Kritik kamen. Dazu gehörten z.B. ungeklärte Bedarfsorientierung, Schnittstellenprobleme sowie eine starke Betonung von stationären Leistungen bei gleichzeitigem Fehlen ambulanter Maßnahmen. Relevante Forschungsthemen in jüngerer Zeit ergeben sich nach Schliehe und Jäckel (2004) im Hinblick auf:

- die Entwicklung kostensparsamer Angebotsformen (z.B. Disease-Management-Programme, so genannte Chronikerprogramme) in Verbindung mit einer verbesserten Steuerung der Inanspruchnahme,
- neue Möglichkeiten in der Integrierten Versorgung unabhängig vom Sicherstellungsauftrag der kassenärztlichen Vereinigungen,
- eine stärkere Dienstleistungsorientierung mit neuen Fragen zur Selbstbestimmung von Patienten, insbesondere von Behinderten,
- die *Diagnosis Related Groups* (Fallpauschalen),
- Bemühungen um die Verstetigung des in der stationären oder ambulanten Rehabilitation erreichten Erfolges durch entsprechende Nachsorgeprogramme (Koch, Lehmann & Morfeld, 2007).

Entsprechend vielfältig waren die Maßnahmen zur Stärkung der Rehabilitationswissenschaften seit 1994. Als wichtigste strukturelle Maßnahmen sind zwei größere Programminitiativen zu erwähnen. Zum einen das 1994 eingeführte Qualitätssicherungsprogramm der Rentenversicherung der *Bundesversicherungsanstalt für Angestellte* (BfA) und zum anderen das rehabilitationswissenschaftliche Verbundforschungsprogramm des Bundesministeriums für Bildung und Forschung (BMBF) und des Verbands Deutscher Rentenversicherungsträger (*vdr*). Acht regionale Forschungsverbände werden in diesem Rahmen mit insgesamt 41 Millionen Euro gefördert.

Hauptziel der Rehabilitationsforschung ist die anwendungsbezogene Forschung, wobei die Grundlagenforschung in den jeweiligen Fachgebieten realisiert wird (Löschmann & Bengel, 2000). Als umfassendes Ziel der Rehabilitation gilt die soziale Integration des Behinderten (Koch & Buschmann-Steinhage, 2004). Dabei erfolgt die Rehabilitationsforschung in einem Spannungsfeld institutioneller und organisatorischer Besonderheiten, die sich aus ihren unterschiedlichen Interessengruppen wie der Gesundheitspolitik, den Kosten-, Leistungsträgern sowie Forschern verschiedener Fachdisziplinen wie Politologen, Soziologen, Sozialwissenschaftlern, Psychologen, Ökonomen, Rehabilitationswissenschaftlern oder Gesundheitswissenschaftlern und dem vielfältigen therapeutischen Personal ergeben (Koch & Buschmann-Steinhage, 2004). Die Vernetzung zu Themen der menschlichen Gesundheit und die damit einhergehenden Kommunikationsprozesse zwischen verschiedenen Spezialisten haben sich weitaus schwieriger gestaltet als rein experimentelle Forschung. Dieses Spannungsfeld hat der jungen Rehabilitationswissenschaft noch keinen eigenständigen Platz eingeräumt und so befindet sie sich auf der Suche nach ihrer Identität (Löschmann & Bengel, 2000).

In der theoretischen Diskussion überwiegt ein Pluralismus mit einer Methodenvielfalt, die je nach Gegenstand eher sozialwissenschaftlicher oder naturwissenschaftlicher, gelegentlich auch geisteswissenschaftlicher Betrachtung unterzogen wird (Koch & Bengel, 2000). Entsprechend werden vorhandene Forschungszugänge aus anderen Fachdisziplinen auf rehabilitationsspezifische Fragestellungen angepasst, deren Analyse vor dem Hintergrund des bio-psycho-sozialen Gesundheits- und Krankheitsbegriffs eine individuums-, institutions-, programm-, bevölkerungs- und umweltbezogene Perspektive erfordern (Koch & Bengel, 2000). Eigenständige spezifische Methodenansätze in der Rehabilitation gibt es derzeit jedoch nicht.

Koch und Bengel (2000) fordern vor dem Hintergrund einer zunehmenden Bedeutung der Rehabilitation durch die demografische Entwicklung und die Zunahme chronischer Erkrankungen einen eigenständigen Wissenschaftsbereich der Rehabilitationswissenschaften. Die weitere Entwicklung in Richtung Selbstständigkeit der Rehabilitationswissenschaften ist nach Koch und Bengel (2000) von der Stabilisie-

rung und Erweiterung bestehender Forschungsnetzwerke, von der Sicherstellung von Kooperationen zu Trägern und Leistungserbringern sowie von der Aquise weiterer Fördermittel abhängig. Zudem hängt eine Professionalisierung in Weiter- und Fortbildung in Richtung „Rehabilitationswesen“ von einer weiteren Identitätsbildung durch wissenschaftliche Kommunikationsorgane wie

1. dem Rehabilitationswissenschaftlichen Kolloquium,
2. der im Jahr 2000 gegründeten Deutschen Gesellschaft für Rehabilitationswissenschaften (DGRW)
3. oder Fachzeitschriften wie der Rehabilitation ab.

Wie in Kapitel 2.4.2 ausgeführt, braucht eine Wissenschaftsdisziplin eine theoretische Verankerung. In diesem Zusammenhang soll im Folgenden ein in den Rehabilitationswissenschaften zurzeit diskutiertes Modell vorgestellt werden.

## **2.4.4 Theoriengenerierung in der Rehabilitation**

### *2.4.4.1 Theoriemodell der Rehabilitation*

Vor dem Hintergrund einer zunehmenden Bedeutung von Rehabilitation fordern Gerdes und Weis (2000) ein Theoriekonzept. Grundlage der Entwicklung sei die Unterscheidung rehabilitationsspezifischer Aspekte von den anderen Aspekten des gesundheitlichen Versorgungssystems und das Überführen dieser Aspekte in ein theoretisches Modell. Dabei sollten sowohl die Standpunkte, unter denen es erstellt wird, als auch die dazugehörigen Interventionen offen gelegt, und die eigentlichen Fragen nach der Art der Probleme, die den Gegenstand Rehabilitation ausmachen, geklärt werden (Gerdes & Weis, 2000).

Die eigentliche Frage danach, was Rehabilitation ausmacht, beantwortet sich indes durch die Veränderung der Zielgruppe von Rehabilitanden in den vergangenen 20 Jahren im Hinblick auf die Altersstruktur und die Zunahme chronischer Erkrankungen. Damit sind Gesundheitsschäden (-störungen) nicht kurzfristig „abzustellen“ und das in der Rehabilitation allgemein hin akzeptierte übergeordnete Ziel einer dauerhaften Wiedereingliederung von Patienten in das alltägliche Leben – Verbesserung der Teilhabe – muss unter einem Paradigmenwechsel langfristig angestrebt werden (vgl. Koch & Buschmann-Steinhage, 2004; Gerdes & Weis, 2000; Bullinger & Ravens-Sieberer, 2000; Nübling & Schmidt, 2000).

Nach Gerdes und Weis (2000) hat die Dimension der Gesundheitsstörung (bzw. des Gesundheitsschadens) als Ausgangspunkt einer Theorie der Rehabilitation die Funktion, die Rehabilitation als Teil des Gesundheitssystems – und nicht irgend-

welcher sozialer Unterstützungssysteme – zu definieren (Abb. 1). Mit der Anlehnung an die ICF wird für eine Theorie der Bezugsrahmen im Sinne einer gemeinsamen Sprache geschaffen. Diese gemeinsame Sprache steht allen am Rehabilitationsprozess Beteiligten zur Verfügung und lässt die einzelnen Akteure und Teilprozesse zum Erreichen eines übergeordneten Ziels beitragen, auf das sich alle Beteiligten beziehen können (Gerdes & Weis, 2000). Entsprechend werden Gesundheitsstörungen dort beschrieben, wo Aktivitäts- und Teilhabestörungen vorliegen, um damit gleichzeitig Zielgruppeneingrenzungen auf Personen vorzunehmen, die für eine medizinische Rehabilitation in Frage kommen oder nicht (Abb. 2).

Bengel und Koch (2000, S. 11) fordern, neben dem Konzept der Gesundheitsschädigungen (ICF) für die Theoriengewinnung, „Aussagen zur Diagnostik, zur Behandlung und zu den Ergebnissen, aber auch zu den organisationalen und strukturellen Rahmenbedingungen“. Darüber hinaus bietet eine Modellbildung die Grundlage für die Entwicklung von Assessmentverfahren zur systematischen Verankerung der zentralen Zieldimensionen in der Rehabilitation (Bührlen et al. 2002).

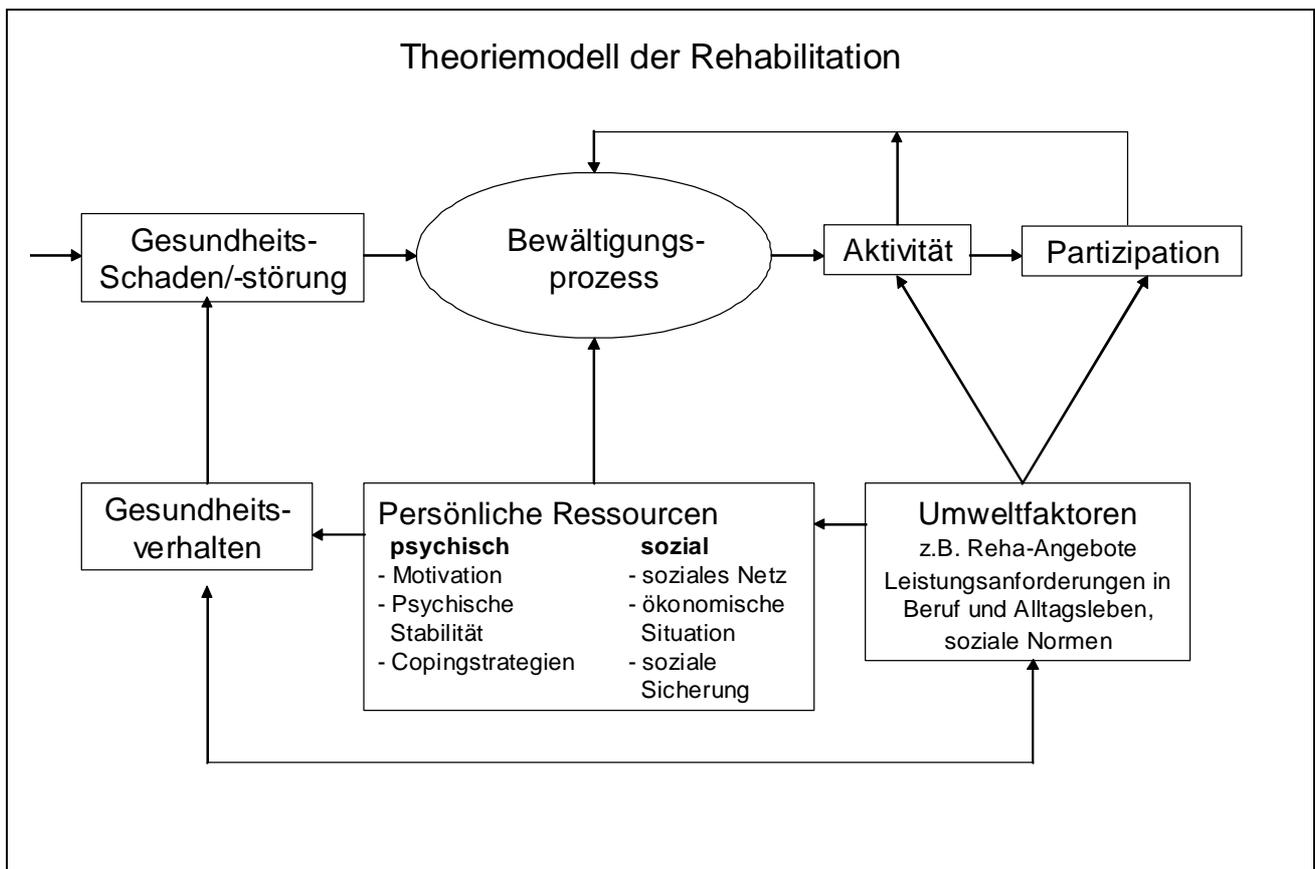


Abb. 2. Theoriemodell der Rehabilitation (Gerdes & Weis, 2000, S. 48).

Mit der Behandlung von Patienten führt eine Beseitigung von Aktivitätsstörungen – insbesondere unter dem Blickwinkel Älterer und chronisch Erkrankter – allerdings

nicht zwangsläufig dazu, dass die Betroffenen den Alltagsanforderungen gewachsen sind. Gerdes und Weis (2000) erweitern das Theorienmodell daher um den Bewältigungsprozess mit seinen psychischen und sozialen Ressourcen sowie das Gesundheitsverhalten auf der edukativen Ebene. Nach ihrer Auffassung werden die Auswirkungen von Gesundheitsstörungen auf Aktivitäts- und Teilhabeebene entscheidend vom Bewältigungsprozess beeinflusst. Dieser findet eben nicht wie in der Akutmedizin üblich durch Ärzte oder Therapeuten statt, sondern durch den Patienten oder den Rehabilitanden selbst. Nur dieser selbst kann mit seiner „subjektiven Wahrnehmung“ gesundheitliche Störungen beeinflussen oder objektiv messbare Veränderungen im Rehabilitationsverlauf für sich positiv deuten. Die zunächst „subjektive Wahrnehmung“ ist allerdings eine Ansammlung von psychischen und sozialen Ressourcen, die nach Gerdes und Weis (2000) eine zentrale Rolle im Rehabilitationsverlauf einnehmen und von daher einer genaueren Betrachtung bedürfen.

#### 2.4.4.2 *Bewältigungsprozesse*

Unter einem Bewältigungsprozess werden Anpassungsprozesse eines Individuums verstanden, die im Zusammenhang mit einer Gesundheitsschädigung und den damit wahrgenommenen Belastungen stehen (Wegner, 2001). Die Bewältigungsprozesse bei der Krankheitsverarbeitung gehen nach Heim (1988, S. 9) auf das Bemühen zurück, „bereits bestehende oder zu erwartende Belastungen durch die Krankheit mehr psychisch (emotional, kognitiv) oder durch zielgerichtetes Verhalten und Handeln zu reduzieren, auszugleichen oder zu verarbeiten“.

Krankheitsverarbeitung und Coping werden in der Literatur weitgehend als synonyme Begriffe verwendet (Gerdes & Weis, 2000). Historisch steht das Konzept der Krankheitsverarbeitung in der Tradition der psychophysiologischen Stresstheorie sowie der psychoanalytischen Ich-Psychologie und Abwehrlehre Anna Freuds (1959). Beide Theorien haben die Entwicklung der Bewältigungsforschung entscheidend beeinflusst und so beziehen sich neuere Ansätze vor allem auf das *Transaktionale Stressmodell* nach Lazarus und Folkman (1984), die wohl einflussreichste Stressbewältigungstheorie (Gerdes & Weis, 2000). Lazarus und Folkman (1984) interpretieren die Ergebnisse der Stressforschung dahingehend, dass nicht nur die Belastung als solche für die Bewältigung ausschlaggebend ist, sondern dass vor allem die subjektive Bewertung des Individuums einen entscheidenden Einfluss auf die Bewältigung hat (Gerdes & Weiß, 2000).

Trotz einer Vielzahl ähnlicher und konvergenter theoretischer Konzepte (Wegner, 2001) besteht weitgehend Einigkeit darüber, dass die Bewältigung einer Krankheit sowohl auf einer kognitiven, emotionalen wie auch auf einer handlungsbezogenen Ebene verläuft. Allen Ansätzen sind zwei Grundprinzipien der Bewältigung gemein-

sam, die einerseits „durch Begriffe wie Annäherung, Problemzentriertheit und aktive Auseinandersetzung, andererseits durch Vermeidung, Abwehr oder Verdrängung“ (Gerdes & Weis, 2000, S.58) gekennzeichnet sind.

In der psychoanalytischen Ich-Psychologie unterscheidet Haan (1977) das Coping, die Abwehr und Fragmentierungsprozesse. Wegner (2001, S.73) beschreibt dies so:

„Während Copingprozesse zielgerichtet, flexibel und realitätsnah sind und einen adäquaten Affektausdruck ermöglichen, sind Abwehrprozesse hingegen zwanghaft, rigide, die Realität verzerrend und mit einem versteckten Affektausdruck gekoppelt. Fragmentierungsprozesse sind die eigentlich pathologischen Prozesse, da sie automatisierte, ritualisierte und irrationale Prozesse umfassen.“

Die klinische Erfahrung zeigt allerdings, dass die starke Abgrenzung nicht haltbar ist, sondern Abwehr und Coping eng miteinander verknüpft sind (Gerdes & Weis, 2000). Tabelle 2 zeigt eine Gegenüberstellung von Aspekten der Krankheitsverarbeitung und Abwehr.

Tab. 2. Gegenüberstellung von Krankheitsverarbeitung und Abwehr (Gerdes & Weis, 2000, S. 59).

Coping	Abwehr
Problemorientiertheit (handlungs- und umweltbezogen)	Emotionszentriert (intrapsychische Emotionsregulierung)
Bewusst	Unbewusst
Realitätsorientiert	Irrationale Form der Verarbeitung
Realitätsangemessenheit	Realitätsverzerrung
Ziel: Umweltveränderung, kognitive Umbewertung, Selbstveränderung	Ziel: Aufrechterhaltung des Selbstwertgefühls
Bewusste Kontrolle unangenehmer Emotionen zum Erhalt der Handlungsfähigkeit	Abwehr von negativen Emotionen und Antrieben

Einflüsse auf die Krankheitsbewältigung aus der Sicht der sozialen Ressourcen ergeben sich nach Filipp und Aymanns (1987) durch positive soziale Unterstützung z.B. aus dem Familienkreis. Für den Familienkreis kann die Situation aber auch eine Belastung darstellen und unter Umständen auf den betroffenen Patienten zurückwirken.

Die Ziele der Krankheitsbewältigung unterscheiden sich je nach Betrachterperspektive (Heim, 1988). So kann

1. aus Sicht des Patienten die Wiedererlangung der körperlichen Leistungsfähigkeit und des subjektiven Wohlbefindens,
2. aus Sicht des sozialen Umfeldes die volle Wiedergewinnung der familiären Rollenfunktion und

3. aus Sicht der medizinischen Leistungserbringer z.B. die Compliance vorrangig beurteilt werden.

Entsprechend ist der Prozess einer Krankheitsbewältigung eng mit den Möglichkeiten oder dem Erkennen persönlicher Ressourcen wie z.B. der Problemlösefähigkeit, der Ich-Identität oder sozialer Unterstützungssysteme eines jeden Menschen verbunden. Hier wirkt wieder Antonovskys Kohärenzgefühl (Kap 2.1), das mit seiner globalen Orientierung dem Menschen ein generalisiertes, überdauerndes und dynamisches Gefühl des Vertrauens vermitteln will.

Für die Rehabilitation nimmt der Krankheitsbewältigungsprozess als Aspekt der persönlichen Ressource eine zentrale Bedeutung im Theoriemodell von Gerdes und Weis (2000) ein. Damit werden die wesentlichen Zielbereiche benannt und die zentralen Gebiete einer Ergebnismessung in der Rehabilitation darüber hinaus auf der somatischen (im Theoriemodell: Gesundheitsschaden bzw. -störung), der alltagsfunktionalen und beruflichen (Aktivität, Partizipation), der psychischen und sozialen (Persönliche Ressourcen) und auf der edukativen Ebene (Gesundheitsverhalten) vorgegeben. Diese Zieldimensionen sollen im Folgenden näher betrachtet werden.

## 2.5 Therapieziele in Physiotherapie und Rehabilitation

In der Krankengymnastik wird im Allgemeinen von *Nah- und Fernzielen* gesprochen. Die Ziele werden einerseits über Nahziele durch Gesichtspunkte (Maßnahmegruppen) und andererseits durch Maßnahmen oder Behandlungstechniken im Hinblick auf das allgemein anerkannte Fernziel der vollständigen Wiederherstellung des Patienten für seinen funktionellen Alltag repräsentiert (Kolster & Ebelt-Paprotny, 1996). Ein aus der klinischen Erfahrung des Autors konstruiertes Beispiel soll dies im Folgenden verdeutlichen.

*Beispiel:* Eine ältere Patientin äußert unmittelbar nach künstlichem Kniegelenkersatz als *Nahziel* „Treppensteigen“ und als *Fernziel* „in den Bergen wieder ein wenig walken“ zu können. *Gesichtspunkte* umfassen in der frühen postoperativen Phase z.B. die evidenzbasierten Maßnahmegruppen der Gelenkkühlung (Brosseau & Yonge, 2003), der Bewegungsschiene (Brosseau & Milne, 2004) und das Prinzip der Frühmobilisation überhaupt (Reitmann & Emerson, 2003; Creditor, 1993). Darunter sind zunächst Bewegungsübungen im Liegen, Sitzen sowie Übungen im Stand zu verstehen. Daran schließen sich die *Gesichtspunkte* zur Anpassung an die Anforderungen alltäglicher Belastungen an. Einzelne *Behandlungstechniken oder Maßnahmen* dazu könnten z.B. das Gehen an Unterarmgehstützen oder das Treppensteigen oder

bei besonderen Problemen im Patellagleitlager u. U. eine zusätzliche Mobilisation der Kniescheibe umfassen.

*Therapieziele* für den physiotherapeutischen Praxisalltag sind darüber hinaus im Heilmittelkatalog (Beyer, 2004) niedergelegt und vorrangig auf der Struktur- und Funktionsebene unmittelbar aus den in Verordnungen formulierten Leitsymptomen zu Diagnosen ableitbar.

Was dagegen in der medizinischen Rehabilitation ein Behandlungserfolg ist, ist immer noch schwer zu beantworten (Broda, 2000). Dies auch deshalb, da während des allgemeinen Krankheitsgeschehens mit einem natürlichen Heilungsverlauf zu rechnen ist, der nicht zwangsläufig auf die Einwirkungen der Therapie zurückzuführen ist. Dabei übersteigt die Anzahl der in der Literatur berichteten Maße zur Erfolgskontrolle die Zahl der vorhandenen Studien zu Interventionen um das Fünffache (Schulte, 1993).

*Rehabilitationsziele* werden in der Rehabilitation auf einen festgestellten individuellen Rehabilitationsbedarf hin mit definierten Maßnahmen belegt (Bullinger & Ravens-Sieberer, 2000). Nach Schliehe & Haaf (1996, S. 669) nimmt der Begriff des Rehabilitationserfolges Bezug auf die formulierten Ziele:

„So erbringt die Rentenversicherung medizinische berufsfördernde und ergänzende Leistungen zur Rehabilitation, um

1. den Auswirkungen einer Krankheit oder einer körperlichen, geistigen oder seelischen Behinderung auf die Erwerbsfähigkeit der Versicherten entgegen zu wirken oder sie zu überwinden und
2. dadurch Beeinträchtigungen der Erwerbsfähigkeit der Versicherten oder ihr vorzeitiges Ausscheiden aus dem Erwerbsleben zu verhindern oder sie möglichst dauerhaft in das Erwerbsleben wieder einzugliedern“.

Die Operationalisierungen der Ziele werden auch als *Indikatoren* oder *Zielkriterien* bezeichnet (Bullinger & Ravens-Sieberer, 2000), wobei das *Outcome* die daraus ableitbaren Informationen über den Grad der Therapiezielerreichung widerspiegelt. Beeinflusst werden die Ergebnisse dabei nicht nur durch die Behandlungsmaßnahmen, sondern häufig in entscheidendem Maße durch verschiedene *Prädiktoren*, wie soziodemographische Variablen, Komorbidität, Behandlungsmotivation und -erwartungen, soziale Unterstützung, Krankheitsverarbeitung (Kap. 2.4.4.2), Kontrollüberzeugungen, berufliche Belastungen sowie Gesundheitsverhalten, Angst und Depressivität (Biefang et al., 1997).

Nach Bullinger (2000, S. 308) ist das Outcome „die über Indikatoren erfasste Zielerreichung in Abhängigkeit der Zielformulierung und unter Berücksichtigung des relativen Beitrags von Prädiktoren“. Das Outcome umfasst demnach:

1. „Sozioökonomische Indikatoren (z.B. Erwerbstätigkeit)

2. Biomedizinische Indikatoren (z.B. Wiedererlangung der Bewegung)
3. Psychosoziale Indikatoren (z.B. Coping) oder
4. Gesundheitsökonomische Indikatoren (z.B. Kosten der Behandlung).“ (Bullinger, 2000, S. 308)

Biefang et al. (1997, S. 215) präzisieren und empfehlen prinzipiell die Unterscheidung von Outcomes auf den Ebenen

1. der „Patientensicht (Lebensqualität, Selbstständigkeit im Alltag, Gesundheitsverhalten, Beschwerden, Angst, Depressivität und Behandlungszufriedenheit)“,
2. der Ärzte- oder Therapeutesicht (biomedizinische Indikatoren),
3. der Versorgungsstrukturen (zur Beurteilung der Effizienz von Rehabilitationsmaßnahmen: sozialmedizinische Indikatoren, gesundheitsökonomische Indikatoren).

Wichtig erscheint die Darstellung von Behandlungsergebnissen durch *objektive* und *subjektive Messmethoden* (Lüdtke, 2000; Wollmerstedt et al., 2006). So wird einerseits der Patient in seinem persönlichen Empfinden ernst genommen, andererseits kann sich der Therapeut von subjektiven Aussagen des Patienten distanzieren und in Verbindung mit seinem eigenen Urteil über die Therapie ein Gesamtbild machen. Darin wird deutlich, dass die Überprüfung des Therapiestandes über das in der biomedizinischen Sichtweise charakteristische „objektive Messen“ hinausgehen muss und Möglichkeiten der Selbstbeurteilung bereithält, die die Entwicklung eines Patienten in seinem Umfeld und daher aus ganzheitlicher, biopsychosozialer Sichtweise widerspiegeln (Scherfer, 2003). Auch Broda (2000) fordert ein multiples Erfolgskriterium, welches neben subjektiven Beurteilungen von Patienten auch die Beurteilung aus therapeutischer und sozialmedizinischer Sicht berücksichtigt. Neben diesen sind aber auch strukturelle Gegebenheiten von entscheidender Bedeutung. So wirkt sich sowohl die Qualität der interdisziplinären Zusammenarbeit zwischen Therapeuten, Ärzten und Kostenträgern ebenso wie die Vernetzung von Theorie, Forschung und Praxis auf den Rehabilitationserfolg aus (Weis & Koch, 1998).

*Rehabilitationsziele bei Erkrankungen des Haltungs- und Bewegungsapparates* zielen nach Pfingsten und Hildebrandt (1995) auf die Schmerzbefreiung und -reduktion sowie Rückgewinnung oder weitgehende Normalisierung motorischer Funktionen. Restliche Funktionsstörungen sollten abgebaut und Behinderungen durch spezielles Training um den Bereich Alltagsbewältigung erweitert werden. Informationen über Ätiologie und Verlauf der Erkrankung sowie Motivation zur Veränderung von Verhalten im Rahmen von Sekundärprävention sollten Anleitungen

zu Bewegung und Sport oder beispielsweise Inanspruchnahme von Selbsthilfegruppen beinhalten (Pfungsten & Hildebrandt, 1995).

Zur *Konkretisierung der Zielauswahl* wurden im Rahmen des Qualitätssicherungsprogramms der Rentenversicherung Listen von potentiell relevanten Therapiezielen erarbeitet, die sich in somatische, funktionale, psychosoziale und edukative Ziele gruppieren lassen. Als häufigste Ziele beim Bewegungsapparat werden Schmerzreduktion, Muskelkräftigung, -stabilität, und Beweglichkeit auf Funktionsebene genannt (Protz, Gerdes, Maier-Riehle & Jäckel, 1996). Aber ebenso wird der Verbesserung einer qualitativ hochwertigen Ausführung bestimmter Aktivitäten, der Steigerung der Belastungsfähigkeit in Alltag und Beruf sowie der Wiedererlangung von Arbeits- und Sportfähigkeit eine wichtige Bedeutung für die Zielkriterien aus Patientensicht beigemessen (Weinhold, 2008).

Als Ergebnis kann festgehalten werden, dass verschiedene Zieldimensionen über unterschiedliche Ergebnisparameter erfasst werden müssen. Dies vor allem vor dem Hintergrund der Evaluation von Effekten an größeren Stichproben. Hier könnten sich mögliche Effekte der Rehabilitation nivellieren, denn nicht alle Parameter sind für alle Patienten gleich relevant und für manche ist der eine oder andere Test am Anfang nicht notwendig (Gerdes, 2006). Insofern sind in der medizinischen Rehabilitation unterschiedliche Kategoriensysteme und Therapieziellisten gebräuchlich, die indikationsspezifisch auf die Besonderheiten der jeweils behandelten Patientengruppe zugeschnitten sind (Steffanowski, Lichtenberg, Schmidt, Huber, Wittmann & Nübling, 2004).

Die *Therapiezielfestlegung vor Behandlungsbeginn* bleibt, ausgehend vom individuellen Behandlungsauftrag, eine entscheidende Schnittstelle zwischen Prozess- und Ergebnisqualität und gilt als Herausforderung für die Qualitätssicherung (Steffanowski et al., 2004). Denn die Zielerreichung kann stark von der Erwartungshaltung des Patienten abhängen (Mahomed, Liang, Cook, Daltroy, Fortin, Fossel & Katz, 2006). So sind bei größerer Zielübereinstimmung zwischen Arzt, Physiotherapeut und Patient eine höhere Zufriedenheit mit der Rehabilitationsmaßnahme, ein höherer Grad der Zielerreichung aus Sicht aller Beteiligten am Ende der Rehabilitation sowie stärkere Verbesserungen in den verschiedenen Outcome-Variablen zu erwarten als bei geringer oder fehlender Zielübereinstimmung (Rietz, Josenhans, Höder & Arlt, 2003; Altenhöner, Leppin, Grande & Romppel, 2001).

Die Abbildung 3 macht zusammenfassend die Abläufe von der Aufnahme eines Patienten über die Zielformulierung auf den verschiedenen möglichen Ebenen

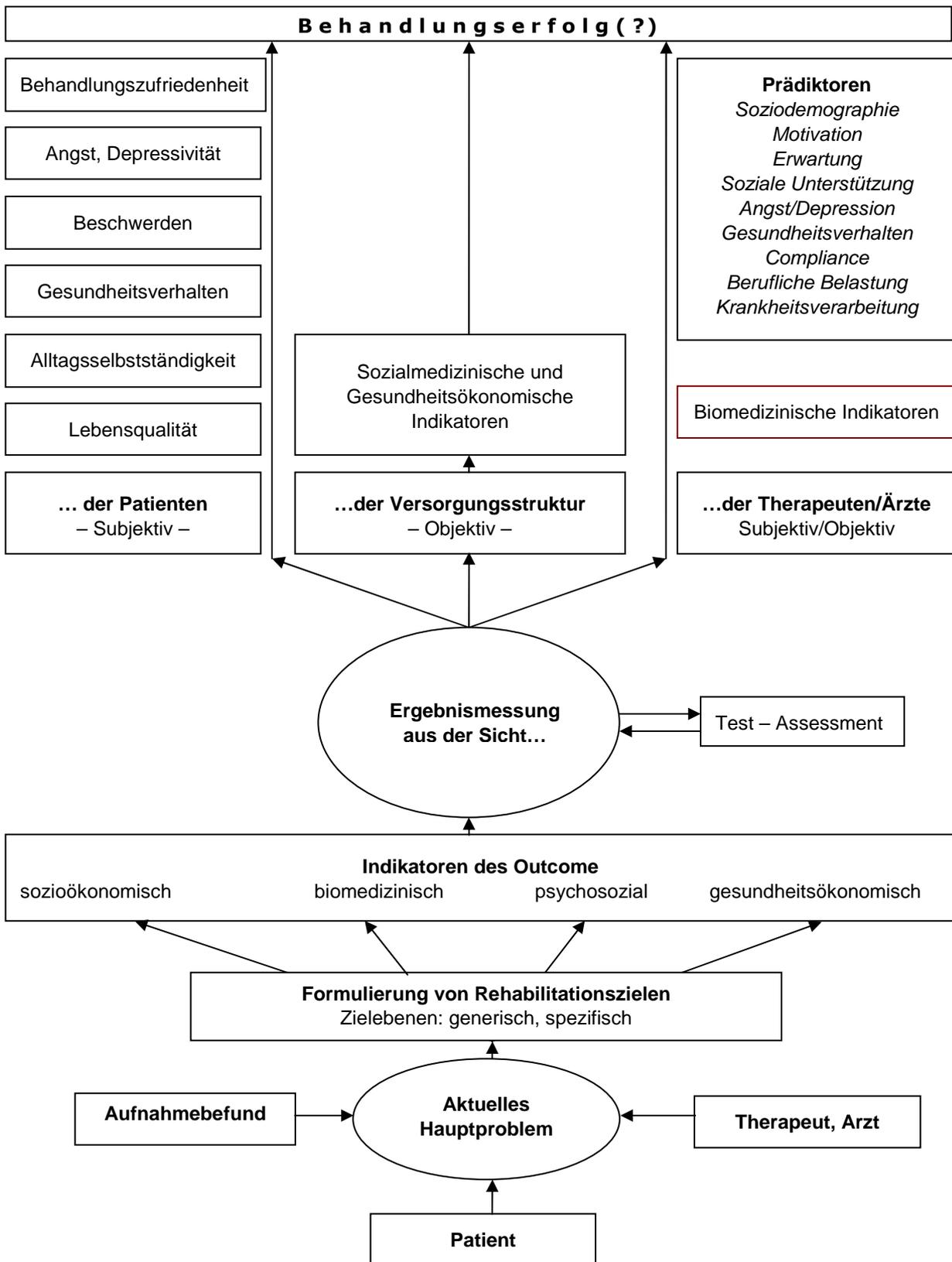


Abb. 3. Schema zum Ablauf von Zielformulierung, Indikatoren, Outcome und Outcomemessung.

bis hin zur Ergebnismessung aus unterschiedlicher Sicht anschaulich, wobei die Instrumente bestenfalls Behandlungserfolge herausstellen.

Die Vielfältigkeit von Zielkriterien lässt die Schwierigkeit bei der Auswahl geeigneter Instrumente verständlich werden. Vor einer Gliederung sowie einer Darstellung ausgewählter Tests und Assessmentverfahren sollen *Grundlagen zu Tests und Assessmentverfahren* dargestellt werden.

## 2.6 Grundlagen zu Tests und Assessmentverfahren

In den folgenden Kapiteln werden Grundlagen zu den Begriffen *Test* und *Assessment* aufgegriffen sowie Anforderungen an Tests im Hinblick auf ihre Gütekriterien erläutert. Testverfahren werden gegliedert und es wird die Bedeutung von Performance-Tests ebenso wie die Problematik zur Systematisierung dargelegt. Es schließt sich ein Überblick über den gegenwärtigen Literaturstand zum Thema Performance-Tests an und es wird auf die Bedeutung von Testverfahren im Rahmen von Qualitätssicherung und Leitlinien eingegangen.

### 2.6.1 Die Begriffe *Test* und *Assessment*

Als ein dem englischen entlehntem und mit „Probe“ umschriebenen Begriff, kommt dem *Test* nach Lienert und Raatz (1998) mehrfache Bedeutung zu. Neben der Untersuchung eines Persönlichkeitsmerkmals wird darunter sowohl die Gesamtheit der zur Durchführung notwendigen Materialien als auch der Vorgang der Durchführung selbst verstanden (Lienert & Raatz, 1998). Ferner werden Untersuchungen mit Stichprobencharakter und unter Anwendung statistischer Verfahren (z.B. t-test) oder auch Klassenarbeiten in der Schule mit einem Test in Zusammenhang gebracht. Zur Vereinheitlichung des Begriffes *Test* schlagen Lienert und Raatz (1998, S. 1) folgende Definition vor:

„Ein Test ist ein wissenschaftliches Routineverfahren zur Untersuchung eines oder mehrerer empirisch abgrenzbarer Persönlichkeitsmerkmale mit dem Ziel einer möglichst quantitativen Aussage über den relativen Grad der individuellen Merkmalsausprägung“.

Bös (2001, S. 533) ergänzt den Begriff des *Tests* um den eines *motorischen Tests* und bezieht den Gegenstandsbereich auf das „individuelle, allgemeine und spezielle motorische Fertigniveau“. *Motorische Tests* müssen nach Bös (2001, S. 533) „unter Standardbedingungen durchgeführt werden und den statistischen Gütekriterien des jeweiligen testtheoretischen Modells genügen“.

Ebenfalls dem angloamerikanischen Raum entlehnt ist das Wort *Assessment*, was soviel wie „Einschätzung, Auswertung, Beurteilung“ (Wydra, 2004, S. 9) bedeutet. Der Gebrauch des Begriffes im Deutschen geht auf die zunächst im englischsprachigen Raum aufgenommene Diskussion um Assessmentverfahren zurück und ergab sich einerseits aus einer Anpassung an den internationalen Sprachgebrauch und andererseits daraus, dass ein *Assessment* ein komplexeres Geschehen ausdrückt, als es der Begriff von *Diagnostik* vermag (Wydra, 2004). Abgrenzungen zwischen dem Begriff *Assessment* und *Test* sind in der Literatur nicht auszumachen. Die Begriffe sollen daher als Synonyme benutzt werden.

Wydra (2004, S.10) versteht unter einem Assessment

„einen multidimensionalen und interdisziplinären diagnostischen Prozess, mit dem Ziel, die medizinischen, psychosozialen und funktionellen Probleme und Ressourcen eines Patienten zu erfassen und einen umfassenden Behandlungs- und Betreuungsplan zu entwickeln“.

*Assessments* stellen als quantitative oder standardisierte Methoden die Beurteilung der Behandlungseffekte oder Outcomes auf eine möglichst objektive und überprüfbare Basis (Biefang et al., 1999).

Dabei sollte die Diagnostik im therapeutischen Bereich niemals Selbstzweck sein, sondern vielmehr unmittelbar in einem Verwertungszusammenhang mit der Behandlungsstrategie (Clinical-Reasoning) stehen (Wydra, 2004). So könnte die Frage aufgeworfen werden, ob sich schlechte Testergebnisse möglicherweise eher aus einem Defizit

1. im Struktur- oder Aktivitätsbereich (z.B. Bewegungseinschränkungen),
2. im koordinativen Bereich (z.B. Gleichgewicht),
3. durch mangelnde Übung oder
4. durch ängstliches Verhalten begründen,
5. oder sich eher aus hohen Anforderungen, die sich aus den Vorstellungen des jeweiligen Patienten in seiner Rollenfunktion im gesellschaftlichen Zusammenleben oder letztlich
6. aus der falschen Auswahl des Assessments ergeben.

Entsprechende Schlüsse könnten dann zu Maßnahmen in der Therapie führen.

## 2.6.2 Testgütekriterien

Die Frage der Anforderung, die ein Test erfüllen muss, um auf die tatsächliche Ausprägung des getesteten Merkmales schließen zu können, ist Gegenstand der Testtheorie (Bortz & Döring, 2006). Ihr liegen die drei zentralen Hauptgütekriterien *Objektivität*, *Reliabilität* und *Validität* zugrunde, die die Qualität eines Tests bestimmen. Bedingte Forderungen an Tests werden als Nebengütekriterien bezeichnet und umfassen die *Normierung*, *Vergleichbarkeit*, *Ökonomie* und die *Nützlichkeit* eines Tests (Lienert & Raatz, 1998). Zusätzliche Kriterien betreffen die Änderungssensitivität und die Fähigkeit eines Tests, zwischen Gesunden und Betroffenen zu unterscheiden (Biefang et al., 1999).

Die Validierung des Polla in der vorliegenden Untersuchung folgt mit der Bestimmung von Objektivität, Reliabilität und Validität den Empfehlungen zu Testgütekriterien von Lienert und Raatz (1998) und Bös (2001). In den folgenden Ausführungen wird nur auf die in der Untersuchung betrachteten Gütekriterien näher eingegangen.

### 2.6.2.1 Hauptgütekriterien

#### *Objektivität*

„Unter der Objektivität eines Tests verstehen wir den Grad, in dem die Ergebnisse eines Tests unabhängig vom Untersucher sind“ (Lienert & Raatz, 1998, S. 7). Anders formuliert kann gefragt werden, inwiefern unterschiedliche Untersucher bei dem selben Patienten zu identischen Ergebnissen kommen. Sie wird auch als *interpersonelle Übereinstimmung* oder *Interrater-Reliabilität* bezeichnet. Letztere soll begrifflich im Folgenden benutzt werden. Zu ihrer Überprüfung werden statistische Verfahren wie das Kappa-Maß oder durchschnittliche Korrelationskoeffizienten angewandt (Biefang et al., 1999), die quantitativ die Höhe der Korrelationen von Messwertreihen bei den verschiedenen Versuchsleitern beschreiben (Bös, 2001). Zur Beurteilung von Objektivitätskoeffizienten nennt Clarke (1976, S. 27) Werte, die in Tabelle 3 wiedergegeben sind.

Tab. 3. Beurteilung von Objektivitätskoeffizienten (Clarke, 1976, S. 27, aus Bös, 2001, S. 546).

Koeffizient	Beurteilung
0.95-0.99	very high; found among the best tests
0.90-0.94	high; acceptable
0.80-0.89	fairly adequate for individual measurement
0.70-0.79	adequate for group measurement, but not satisfactory for individual measurement
0.60-0.69	useful for group averages and school surveys, but entirely inadequate for individual measurement

Eine Testobjektivität von über 0,90 ist in Anlehnung an die Beurteilungskriterien nach Clarke (1976) zu fordern (Bös, 2001).

### *Reliabilität*

Die Genauigkeit, mit der ein Test misst, wird als Reliabilität bezeichnet. Sie klärt die Frage, inwieweit unterschiedliche Messungen zu den gleichen Ergebnissen kommen, und ob ähnliche Informationen erfasst werden, unabhängig davon, ob ein Test ein bestimmtes Kriterium auch zu messen beansprucht (Bortz & Döring, 2006; Lienert & Raatz, 1998). Die im Rahmen der vorliegenden Untersuchung angewandten Verfahren klären

1. die *Intrarater-Reliabilität* und
2. die *interne Konsistenz (Konsistenzanalyse)*.

Bei der *Intrarater-Reliabilität* überprüft ein und derselbe Beurteiler zu zwei Zeitpunkten die drei gemessenen Tests an einem Patienten. Entsprechende Empfehlungen der Literatur legen einen 3-14 Tage Abstand bei Konditions- (Bös, 2001, S. 550) und 2-4 Wochen bei Koordinationstests nahe. Bei letzteren ist nach Bös (2001, S. 550) besonders mit Lern- und Erinnerungseffekten zu rechnen, die durch den entsprechend längeren Zeitraum zwischen den Tests minimiert werden sollten. Biefang et al. (1999) schlagen bei psychologischen Testverfahren einen Abstand von sieben Tage vor.

Die Berechnung der *Konsistenzanalyse* erfolgt über den Alphakoeffizienten nach Cronbach (1951). Ein Test wird demnach in so viele Teile zerlegt wie Items vorhanden sind. In einer itemspezifischen Korrelationsanalyse werden die Autokorrelationen der Items mit sich selbst außer Acht gelassen. Die Durchschnittswerte der Korrelationen gehen in den Schätzwert des Alphakoeffizienten von Cronbach mit ein (Bortz & Döring, 2006). Biefang et al. (1999, S. 25) empfehlen als Richtwert zur internen Konsistenz einen Cronbach Alpha von 0,7. Nach Weise (1975, S. 219) sollte ein Test eine Reliabilität von 0,8 aufweisen, Werte zwischen 0,8 und 0,9 gelten als mittelmäßig und über 0,9 als hoch.

### *Konstruktvalidität*

Die Konstruktvalidität dient dem Ziel, Beziehungen zwischen den in einem Messinstrument gemessenen Verhaltensweisen und einem zu erfassenden Merkmal (= Konstrukt) aufzuklären. Ihre Bedeutung liegt weniger in einer praktisch-diagnostischen Verwertbarkeit, als vielmehr in der theoretischen Klärung dessen, was der betreffende Test misst (Lienert & Raatz, 1998). Die Konstruktvalidität wird deutlich, wenn bedacht wird, dass ein Testverhalten häufig nach „angenommenen bzw. erschlossenen“ (Lienert & Raatz, 1998, S. 226) Eigenschaften erfasst wird. So zielen z.B. Intelligenztests und Fähigkeitstests auf abgeleitete Konstrukte und nicht auf unmittelbar operational erfassbare Einheiten ab (Herrmann, 1969).

Bei der Validierung können unterschiedliche Verfahren zum Einsatz kommen:

„Die Ableitung zugrunde liegender hypothetischer Größen, Faktoren und Testdaten erfolgt zu-  
meist über Faktorenanalysen. Ziel der faktorenanalytischen Untersuchungen ist die Beantwor-  
tung der Frage, durch wie viele Faktoren (Dimensionen) ein Test erklärt wird“ (Bös. 2001, S.  
552).

Nach Cronbach und Meehl (1995) umfassen die wichtigsten Verfahren zur Kon-  
struktvalidierung neben einer Faktorenanalyse des zu validierenden Tests mit ge-  
meinsamen Außenkriterien die Korrelationen mit mehreren Außenkriterien oder mit  
Tests die ähnliche Validitätsansprüche aufweisen.

Die Konstruktvalidität in dieser Untersuchung wird über die Höhe von Zusammen-  
hängen nach dem Korrelationskoeffizienten Spearman-Rho berechnet. Die Beurtei-  
lung der Höhe von Zusammenhängen ergibt für Korrelationen (Bühl & Zöfel, S.  
322):

Bis	0,2	sehr gering,
bis	0,5	gering,
bis	0,7	mittel,
bis	0,9	hoch,
über	0,9	sehr hoch.

### *Kriteriumsvalidität*

Bei der Bestimmung von Gütekriterien für einen Test nehmen der Instrumentenver-  
gleich als Kriteriumsvalidität und die externe Verankerung eine zentrale Bedeutung  
ein. Die Kriteriumsvalidität beschreibt den Grad der Übereinstimmung von Tester-  
ergebnissen einer Stichprobe von Probanden mit einem Außenkriterium (Lienert &  
Raatz, 1998). Nach Bös (2001, S. 553) stellt die kriterienbezogene Validität „das  
wichtigste Maß für die Beurteilung der anwendungspraktischen Relevanz eines  
Tests“ dar. Von Interesse im Rahmen dieser Studie ist beispielsweise, ob die  
Registrierung eines verbesserten Funktionsstatus über den Polla gleichzeitig mit  
einer Verbesserung der subjektiv beurteilten gesundheitsbezogenen Lebensqualität  
des Patienten einhergeht. Derartige Fragestellungen sind häufig Gegenstand von  
Hypothesen und zielen auf die Betrachtung korrelativer Zusammenhänge zwischen  
den Instrumenten (zur Beurteilung der Höhe von Korrelationen siehe *Konstruktvali-  
dität*).

### 2.6.2.2 *Nebengütekriterien*

Nach Lienert und Raatz (1998) schließen sich vier Nebengütekriterien der *Normie-  
rung*, *Vergleichbarkeit*, *Ökonomie* und *Nützlichkeit* als bedingte Forderungen an die  
Hauptgütekriterien eines Tests an.

Das Nebengütekriterium der *Normierung* hat zum Ziel, individuelle Ergebnisse in ein Bezugssystem einzuordnen (Lienert & Raatz, 1998). Dieses Kriterium kann in dieser Untersuchung aufgrund des geringen Stichprobenumfangs nicht erfüllt werden. Bezüglich der *Vergleichbarkeit* wird ein Testinstrumentarium mit einem ähnlichen Verfahren verglichen. Die *Ökonomie* eines Tests wird über eine kurze Durchführungszeit, einfache Handhabung, einen geringen Materialaufwand sowie eine schnelle Auswertung erreicht. Ein Test sollte zudem als Gruppentest durchführbar sein (Lienert & Raatz, 1998). Für die *Nützlichkeit* eines Tests spricht, wenn dieser ein Persönlichkeitsmerkmal oder eine Verhaltensweise misst oder vorhersagt, für deren Untersuchung ein praktisches Bedürfnis besteht (Lienert & Raatz, 1998).

Die Frage danach, ob Tests im Zeitverlauf Veränderungen darzustellen in der Lage sind, wird über das statistische Maß der *Änderungssensitivität* beschrieben. Effekte zu Stichprobenunterschieden zwischen zwei Messzeitpunkten werden über den Standardized Response Mean (SRM) ermittelt (Leonhardt, 2004). Eine Klassifikation der Effekte orientiert sich an den Empfehlungen von Cohen (1988, S. 26 f.). Danach werden Werte von 0,2 als kleine, von 0,5 als mittlere und von 0,8 als große Effektstärken beschrieben.

Die *Unterscheidungsfähigkeit* beschreibt die Fähigkeit, zwischen unabhängigen Stichproben zu unterscheiden. So stellt sich die Frage, ob über die Testergebnisse der angewandten Verfahren zwischen Gesunden und Patienten unterschieden werden kann. Folglich werden die Signifikanz zu Unterschieden bzw. die Effekte zur Signifikanz von Unterschieden ermittelt. Zur Bestimmung von Effektgrößen signifikanter Unterschiede wird Eta angegeben. Bei einer Varianzanalyse ergeben sich kleine Effekte bei Eta-Werten von 0,10, mittlere bei 0,25 und große Effekte bei 0,4 (Bortz & Döring, 2006).

### **2.6.3 Vorzüge von Tests und Einflüsse auf Ergebnisse**

Bei aller Bedeutung bezüglich des Einsatzes von Testverfahren muss die Problematik des Zustandekommens von Testergebnissen und deren u. U. eingeschränkter Tragweite zumindest angedeutet werden. Testergebnisse bleiben vielfältigen Einflüssen ausgesetzt und können auch zu falschen Schlüssen führen. In Tabelle 4 unterscheidet Simmonds (1997) zwischen *methodischen* (z.B. Aufwärmung, Testreihenfolge), *psychologischen* (z.B. Erwartung, Schmerzen) und Einflüssen, die sich aus *Patientendaten* ergeben; dazu zählen z.B. die Tagesform, unerwartete Schmerzzustände, Ermüdungseffekte, Motivationsschwächen oder das Alter.

Tab. 4. Darstellung von Faktoren, die Muskelkraftmessungen beeinflussen (Simmonds 1997, aus Cabri, 2001, S. 211).

Methodik	Psychologie	Patientendaten
Ausstattung/Geräte	Motivation	Gesundheit
Ausgangsstellung	Lernen	Krankheit
Stabilisation	Können	Schmerzen
Aufwärmen	Leidensdruck	Geschlecht
Vordehnen	Depression	Alter
vorherige Übung	Wahrnehmung	Größe
Pausen	Erwartung	Gewicht
„Anfeuern“	Selbstwirksamkeit	Aktivitätsniveau
Reihenfolge der Tests		
Isometrisch/dynamisch		
Konzentrisch/exzentrisch		
Bewegungsgeschwindigkeit		
eingesetzte Parameter		
Tester		
<b>Messung</b>	<b>Muskulatur</b>	<b>Neuronale Faktoren</b>
operationale Definition	Fasertyp	Rekrutierung
Reliabilität	Struktur	Kodierung der Geschwindigkeit
Validität	Größe	Inhibition
Objektivität	Fiederungswinkel	Synchronisation
Sensibilität	Insertionswinkel	
	Hebelwinkel	

Wichtig im Hinblick auf die Beurteilung von isolierten Teilergebnissen (z.B. Beweglichkeitsmessung) ist die Einbindung und Betrachtung dieser Ergebnisse im Kontext aller Lebensbedingungen eines Individuums. So könnte die Beweglichkeitsmessung an einer Extremität die Verbesserung im Therapieverlauf zwar objektiv positiv belegen, der betroffene Patient das Therapieergebnis aus seiner subjektiven Sicht aufgrund fehlender sozialer Unterstützung oder berufsbedingter Schwierigkeiten, allerdings negativ beurteilen.

Bei allen Nachteilen, die beim Einsatz und der Interpretation von Testergebnissen entstehen können, überwiegen jedoch die Vorteile. Die Vorteile der Anwendung von Testverfahren beschreibt Cabri (2001, S. 196). Sie helfen

1. „auf objektiven und subjektiven Daten basierend ein effektives Clinical-Reasoning durchzuführen,
2. bei klinischen Problemen Ziele zu formulieren,
3. bei der Überwachung von Behandlungsfortschritten,
4. einen Überblick über die Veränderungen infolge einer Therapie zu geben,

5. bei der Bewertung der Wirksamkeit physiotherapeutischer Maßnahmen und bei der Entwicklung neuer Behandlungsstrategien.“

### **2.6.4 Gliederung von Testverfahren**

Zur Einordnung von Tests und zur Begründung der Auswahl verschiedener Verfahren für die vorliegende Untersuchung werden an dieser Stelle zwei Unterscheidungen getroffen. Zum einen wird aus der Beurteilerperspektive folgend nach

1. Instrumenten der Selbsteinschätzung,
2. Instrumenten der Fremdbeurteilung und
3. kombinierten Testverfahren unterschieden.

Eine zweite angewandte Gliederung lehnt sich an Guyatt, Jaeschke, Feeny und Patrick (1996) sowie Patrick und Erickson (1993) an. Sie wurde im deutschsprachigen Raum von Biefang, Potthoff und Schliehe (1999) aufgegriffen und hat sich in der medizinischen Rehabilitation weitgehend etabliert. Biefang et al. (1999) unterscheiden zwischen

1. generischen (krankheitsübergreifenden) Instrumenten und
2. spezifischen Instrumenten.

*Generische Instrumente* messen global die gesundheitsbezogene Lebensqualität. Dazu gehören der Funktionsstatus, das Wohlbefinden, die Behandlungszufriedenheit, das Gesundheitsverhalten, die Rehabilitationsmotivation, die soziale Unterstützung und die Krankheitsbewältigung, die unabhängig von der jeweiligen Krankheit oder Störung des Patienten eingesetzt werden können sowie der Schmerz (Biefang et al., 1999). Sie sollten zusammen mit Angaben zur Person, zu soziodemographischen Merkmalen, Gesundheitsrisiken und zur Komorbidität über bewährte Selbsteinschätzungsverfahren gemessen werden (Biefang & Potthoff, 1996).

*Spezifische Instrumente* richten sich nach der im Vordergrund stehenden Störung – also dem Krankheitsbild – und umfassen z.B. klinische Funktionstests, Performance-Tests sowie Scores und Schweregradklassifikationen auf der Basis medizinischer Befunde (Biefang et al., 1999).

#### **2.6.4.1 Instrumente der Selbsteinschätzung**

Während vor allem biomedizinische Fremdbeurteilungen lange zum Standardrepertoire gehörten, werden Selbsteinschätzungsfragebögen bei Patienten erst seit

Beginn der achtziger Jahre zunächst im angloamerikanischen Bereich systematisch in die medizinische Forschung und Praxis einbezogen (Najman & Levine, 1981; Walker & Rosser, 1988). Mittlerweile ist unumstritten, dass Patienten in der Lage sind, ihre Symptome und Funktionseinschränkungen zuverlässig zu beurteilen, insbesondere wenn standardisierte und relevante Fragen gestellt würden (Stucki, Stucki & Sangha, 1997).

### *Generische Instrumente*

Für den Bereich der generischen Instrumente in Form von Selbsteinschätzungen kommen als zentrale und krankheitsunabhängige Verfahren insbesondere die Überprüfung des Funktionszustands und des Wohlbefindens, der Behandlungszufriedenheit, des Gesundheitsverhaltens und der Motivation zu Rehabilitation, die soziale Unterstützung und Krankheitsbewältigung sowie der Schmerz in Frage (Biefang & Schuntermann, 2000). Tabelle 5 gibt einen Überblick über ausgewählte Verfahren.

Tab. 5. *Generische Instrumente für Gesundheitszustand und intervenierende Merkmale (Biefang & Schuntermann, 2000, S. 110, Auszüge).*

<b>Funktionszustand und Wohlbefinden</b>	
<b>ALLTAG</b>	Fragebogen Alltagsleben
<b>EuroQol</b>	European Quality of Life Questionnaire (deutsche Version)
<b>FLZ<sup>M</sup></b>	Fragen zur Lebenszufriedenheit-Module
<b>IHRES-2</b>	Indikatoren des Reha-Status-Version 2
<b>MLDL</b>	Münchener Lebensqualitätsdimensionen Liste
<b>SF-36</b>	Fragebogen zum Gesundheitszustand (deutsche Version SF-36 Health Survey)
<b>SIP</b>	Sickness Impact Profile (deutsche Version)
<b>WHOQOL-100</b>	The World Health Organization Quality of Live Assessment (deutsche Version)

Muthny, Bullinger und Kohlmann (1999) empfehlen zur Überprüfung der subjektiven Gesundheit und Lebensqualität den Short Form Health Survey (SF-36, Ware & Sherbourne, 1992; Bullinger & Kirchberger, 1998), den IRES-Fragebogen (Indikatoren des Reha-Status, Gerdes & Jäckel, 1995) sowie den WHOQOL Fragebogen.

Der WHOQOL ist ein neueres, in noch internationaler Testung befindliches Verfahren mit vier Dimensionen und 24 Subdimensionen, für das aber bisher eine Empfehlung nicht ausreichend geprüft ist (Muthny et al., 1999). Der IRES-2 mit 161 Items, der mittlerweile in einer dritten Version IRES-3 mit 144 Items vorliegt, umfasst zur Unterstützung der Diagnostik und zur Ergebnismessung in der Rehabilitation die Selbsteinschätzung der somatischen Gesundheit, der Schmerzen, des

Gesundheitsverhaltens, der körperlichen und beruflichen Funktionsfähigkeit, des psychischen Befindens sowie die Selbsteinschätzung der Krankheitsbewältigung und der sozialen Integration.

### *Der SF-36 Fragebogen*

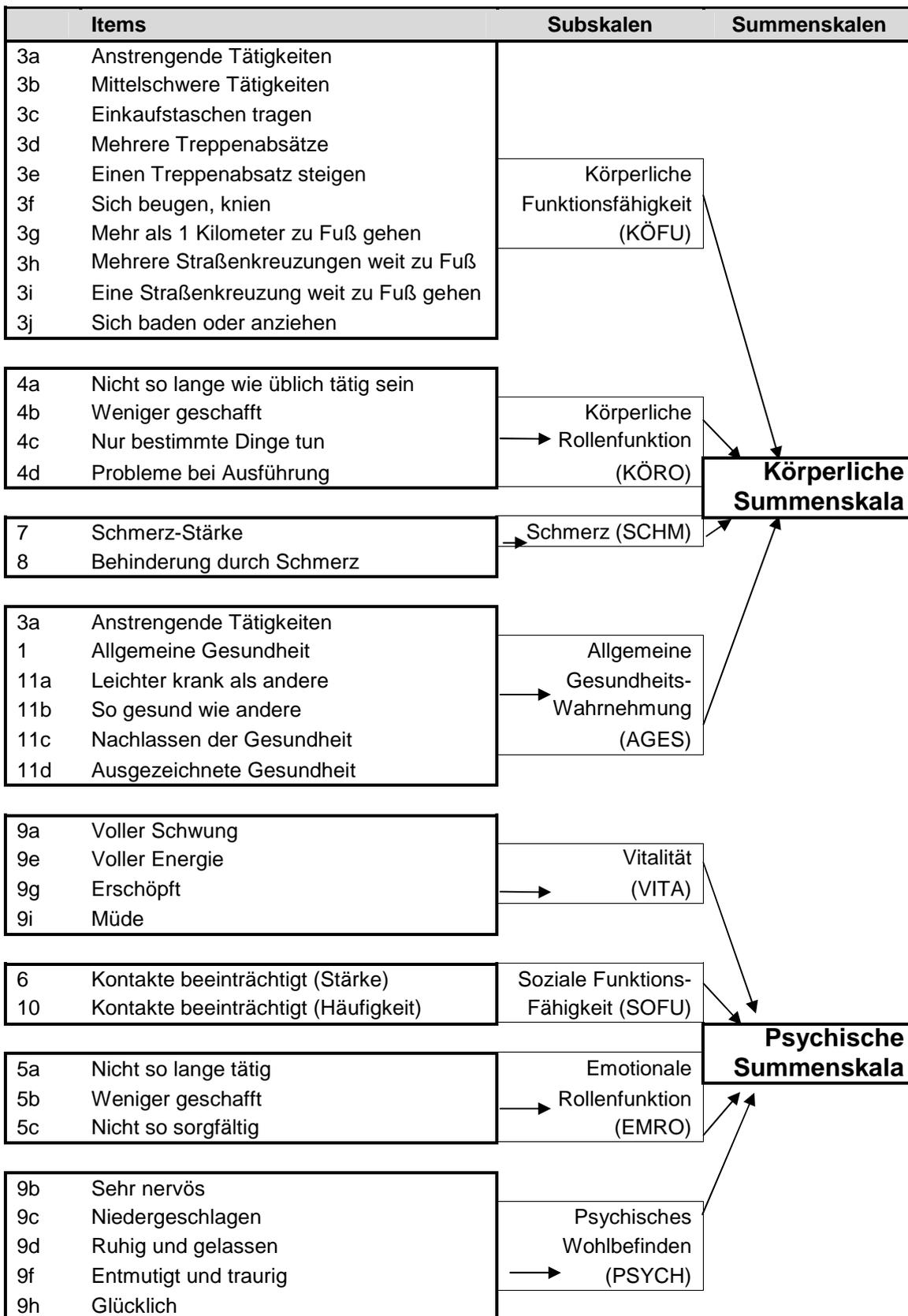
Der Short Form (SF)-36 Health Survey ist ein psychometrisch in verschiedenen Populationen geprüfter generischer Fragebogen zur Erfassung der gesundheitsbezogenen Lebensqualität. Er ist weltweit der am häufigsten eingesetzte Fragebogen zu diesem Thema (Biefang et al., 1997). Patienten füllen den Fragebogen aus, indem sie bei jeder Frage eine Antwort ankreuzen, die ihrem Erleben am ehesten entspricht. Die Zeit für das Ausfüllen wird auf etwa 10 Minuten angesetzt. Der Fragebogen kann in unterschiedlichen Zeitfenstern von z.B. einer oder vier Wochen eingesetzt werden.

Hintergrund seiner Entwicklung in den sechziger Jahren war der Versuch, die Leistungen von Versicherungssystemen in Amerika zu prüfen. Aus einer zunächst umfangreichen Sammlung wurden die Fragen ausgewählt, welche ein Konstrukt der subjektiven Gesundheit abzubilden in der Lage waren (Bullinger & Kirchberger, 1998). Einer Definition zur *subjektiven Gesundheit* waren verschiedene Arbeiten und Expertensitzungen vorausgegangen.

Die 36 Items lassen sich verschiedenen Themenbereichen zuordnen. So ist jedes Item entweder selbst eine übergeordnete Skala oder Teil einer Skala. Der SF-36 erfasst acht Dimensionen zur gesundheitsbezogenen Lebensqualität (ANHANG 2). Die Dimensionen sind als *Körperliche Funktionsfähigkeit* (KÖFU), *Körperliche Rollenfunktion* (KÖRO), *Schmerz* (SCHM), *Allgemeine Gesundheitswahrnehmung* (AGES), *Vitalität* (VITA), *Soziale Funktionsfähigkeit* (SOFU), *Emotionale Rollenfunktion* (EMRO) und *Psychisches Wohlbefinden* (PSYC) dichotom oder mehrstufig Likert-skaliert (Bullinger & Kirchberger, 1998). Die Itemzuordnungen zu den einzelnen Summenskalen sind der Tabelle 6 zu entnehmen.

Im Norddeutschen Verbund für Rehabilitationsforschung (NVRF) findet der SF-36 zurzeit in mehreren Studien Anwendung. Die Vorteile seines Einsatzes ergeben sich aus einer höheren Untersuchungs- und Testökonomie (36 Items) gegenüber dem IRES-2 (161 Items) bei gleichzeitig größerer internationaler Vergleichbarkeit, wobei allerdings im deutschsprachigen Raum begrenzt bevölkerungsbezogene Normwerte und Daten aus klinischen Gruppen vorliegen (Muthny et al., 1999).

Tab. 6. Itemzuordnung zu den Summenskalen des SF-36 (Bullinger &amp; Kirchberger, 1998, S. 66).



Die testtheoretischen Gütekriterien wie die Reliabilität zeigen für die Subskalen, außer der Skala *Allgemeine Gesundheitswahrnehmung*, zufrieden stellende Koeffizienten für die Interne Konsistenz von Cronbach Alpha  $>0.70$  (Bullinger et al., 2003). Die Validität, gemessen über korrelative Vergleiche mit unterschiedlichen Messinstrumenten, zeigt insgesamt gute Ergebnisse (Bullinger et al., 2003). Hinsichtlich der Änderungssensitivität zeigt der SF-36 bei einer Stichprobe von 2295 Personen im Allgemeinen niedrige Werte über verschiedene Indikationen hinweg (Igl, Zwingmann, Faller, Beutel, Beyer, Bischoff, Fritschka, Gustson, Hillert, Knickenberg, Kühn, Kulick, Lamprecht, Nübling, Rief, Schmidt, Schmitz, Vauth, Vogel, Wagner, Wallesch & Wittmann, 2006).

Kritisch anzumerken ist, dass der SF-36 in der Literatur nicht unwidersprochen bleibt. Der SF-36 wurde ursprünglich für epidemiologische Untersuchungen entwickelt, deren Normwerte durch Erhebungen im häuslichen Umfeld gewonnen wurden. Zwingmann, Metzger und Jäckel (1998) diskutieren Fragen, inwiefern sein Einsatz in Kliniken zu gleichermaßen sinnvollen und gleichwertigen Aussagen führt wie sein Einsatz zuhause. Müller, Franke, Schuck und Resch (2001) bemängeln zu hohe Prozentsätze an fehlenden Werten bei einigen Skalen. Zudem scheinen die Ergebnisse auch bei den vier psychosozialen Skalen zu Überschätzungen der Lebensqualität zu führen. Dem widersprechen Bullinger et al. (2003). Sie konnten, mit Ausnahme der Rollenfunktionsskalen, eine sehr geringe Fehlerquote von unter 2% bei 6728 Patienten nachweisen. Kritisch anzumerken seitens des Autors ist, dass die Formulierungen bei Patienten häufiger zu Verwirrungen und Nachfragen führen.

### *Spezifische Instrumente*

Die Messebenen spezifischer Verfahren zur Überprüfung von Aktivitäten des Bewegungsapparates an der unteren Extremität weisen in der Literatur Instrumente zur Selbst- und Fremdeinschätzung sowie kombinierte Verfahren auf. Biefang und Schuntermann (2000, S.110; Tab. 7) listen sieben Instrumente für Skelett, Muskeln und Bindegewebe auf.

Tab. 7. *Spezifische Instrumente für Gesundheitszustand und intervenierende Merkmale (Biefang & Schuntermann, 2000, S. 110, Auszüge).*

<b>Skelett, Muskeln und Bindegewebe</b>	
<b>ACA-RA</b>	American College of Rheumatology 1991 revise criteria of global functional status in Rheumatoid Arthritis
<b>AIMS 2</b>	Arthritis Impact Measurement Health Status Questionnaire – Version 2 (deutsche Version)
<b>FFbH-PI-/R-/OA</b>	Funktionsfragebogen Hannover: polyartikuläre Erkrankungen /Rückenschmerzen/ Arthrosen
<b>HAQ-G</b>	Health Assessment Questionnaire (deutsche Version CH/D)
<b>WOMAC</b>	Western Ontario and McMaster Universities Osteoarthritis Index (deutsche Version)

Der *Health Assessment Questionnaire* (HAQ) erfasst mit seinen 24 Items Dimensionen physischer und psychischer Behinderung, Schmerzen, Schweregrad der Erkrankung, Berufstätigkeit, Einkommen sowie die Kosten der Behandlung und Nebenwirkungen von Medikamenten (Sangha & Stucki, 1997).

Der *Western Ontario and McMasters Universities Osteoarthritis Index* (WOMAC) oder die vergleichbare und kostenlose Version *des Knee and Osteoarthritis Outcome Score* (KOOS) beinhalten dagegen neben Fragen zu Alltagsaktivitäten, Schmerzen und Steifigkeit, die Abfrage von Aspekten zu Sport, Freizeit sowie zur Lebensqualität (WOMAC, Bellamy & Buchanan, 1988; deutsche Version: Stucki, Meier, Stucki, Michel, Tyndall, Dick, Theiler, 1996; KOOS; Roos, 1999).

Auch der aus dem englischen Sprachraum stammende und von Ludwig, Melzer, Grimmig und Daalman (2002) ins deutsche adaptierte und evaluierte Lequesne-Index mit 11 Items erfasst Aktivitäten und Schmerzen bei Menschen mit Hüft- und Kniegelenkserkrankungen bzw. nach operativen Eingriffen am Hüft- und Kniegelenk. Der XSMFA-D (Wollmerstedt et al., 2006) als Selbsteinschätzungsinstrument für die Evaluation der Aktivitätsfunktion der oberen und unteren Extremität erscheint mit mehr als 46 Items für den physiotherapeutischen Routineeinsatz recht umfangreich. Beide Verfahren sind nicht in Tabelle 7 abgebildet, da es sich um neuere Verfahren handelt.

Lediglich der von Kohlmann, Knahr, Krysylin-Exner, Heinrichs und Peschel (1996) evaluierte Funktionsfragebogen Hannover-Osteoarthritis (FFbH-OA, Tab. 7) erfasst mit 18 Items als einer der ersten deutschsprachigen spezifische Instrumente ausschließlich Einschränkungen von Aktivitäten an der unteren Extremität. Allerdings fehlen dem Verfahren Abfragen zu Aktivitäten des Hockens und Kniens, die sich vor allem bei Personen in handwerklichen Berufen wie Fliesenlegern oder Heizungsinstallateuren stellen und im klinischen Setting häufig als Einschränkungen beschrieben werden. Außerdem sind nach Operationen am oberen Sprunggelenk Einschränkungen in der Beweglichkeit des Fußes mit Herabsetzung des Aktivitätsniveaus beim Zehen- oder Hackenstand zu beobachten. Schließlich fehlen Aktivitäten wie die des Radfahrens, das durchaus eine Einschränkung der Mobilität außer Haus darstellen kann. Die Überprüfung dieser Aktivitäten erscheint aber für den diagnostischen Prozess vor dem Hintergrund eines effektiven Clinical-Reasoning sinnvoll.

#### 2.6.4.2 Instrumente der Fremdbeurteilung

Unter der Begrifflichkeit der *Fremdbeurteilung* werden an dieser Stelle Verfahren nach der Einteilung von Biefang, Birkner, Thien, Härtel und Bullinger (1997)

gefasst, die das Outcome von Therapie aus der Sicht von Ärzten und Therapeuten in Form von biomedizinischen Indikatoren verstehen. Darunter fallen zum Beispiel Kraft- und Bewegungsmessungen, aber auch die Beurteilung von Performance-Tests oder komplexere Verfahren, bei denen eindeutig ist, dass diese von Therapeuten oder Ärzten überprüft werden, ohne dass eine subjektive Aussage des Patienten notwendig ist.

#### 2.6.4.3 Kombinierte Testverfahren

Der Vollständigkeit halber werden Verfahren erwähnt, die als kombinierte Instrumente einzuordnen sind. Sie beinhalten sowohl eine subjektive Selbsteinschätzung des Patienten als auch eine objektive Fremdbeurteilung. Die Kombination ergibt sich zwangsläufig durch Abfragen, wie sie in einem Patienten-Setting nicht immer objektiv überprüfbar sind, sondern auf Informationen des Patienten oder Hilfs- oder Pflegepersonen angewiesen sind.

Unter diesen Verfahren sind zum Beispiel so genannte Therapiescores, wie sie Krämer und Maichl (1993) zusammengefasst haben. Die Scores kombinieren beispielsweise Selbsteinschätzungen wie z.B. das Giving-Way Phänomen, die Gelenksteifigkeit, die Schmerzintensität oder die Gelenkstabilität mit Abfragen von Aktivitäten oder klinischen Parametern aus Ärzte- oder Therapeutesicht. Hier sind sowohl Messungen der antero-posterioren Kniebandstabilität, die Bewegungsmessung nach der Neutral-Null-Methode (Debrunner, 1971) als auch der Strukturbefund in Form einer Umfangsmessung oder das Coping zu erwähnen. Zu den Therapiescores zählen u.a.

- die ursprünglich für Hüfte und Knie validierte Version des *Lequesne-Index* (Lequesne, Mery, Samson & Gerard, 1987),
- der *Lysholm-Score* (Lysholm & Gillquist, 1982),
- der Score der *Knie Society* (Insall, Dorr, Scott & Scott, 1989) oder
- der *Harris-Hipscore* (Harris, 1969).

Verfahren wie der *Barthel-Index* (Mahony & Barthel, 1965) lassen zum Beispiel in ambulanten Therapiesituationen nicht uneingeschränkte Fremdbeurteilung zu (z. B. Darmkontrolle, Blasenkontrolle, Essensaufnahme, Baden, Toilettenbenutzung). Hier sind ambulant tätige Therapeuten auf die Beurteilungen aus stationären Einrichtungen oder auf die von Pflegepersonen und Angehörigen angewiesen. Vor einer Erläuterung der Systematisierungs-Problematik und einer Analyse zum Stand der Literatur von Performance-Tests wird folgend die Einbindung von Testverfahren in den Rahmen von Qualitätssicherung und Leitlinien beschrieben.

### 2.6.5 Testverfahren im Kontext von Qualitätssicherung

In der DIN ISO 9004 (1992) wird Qualität als „die Gesamtheit von Eigenschaften und Merkmalen eines Produktes oder einer Dienstleistung, die sich auf deren Eignung zur Erfüllung festgelegter oder vorausgesetzter Erfordernisse beziehen“ bezeichnet (Dorenburg & Tiefensee, 2000, S. 198). Demnach ist „Physiotherapie“ als eine Dienstleistung aufzufassen, in deren Mittelpunkt die Orientierung am Patienten steht.

Die Dienstleistung „Physiotherapie“ geht über die Behandlung im Sinne einer hands-on-Zeit, also der Zeit die ein Therapeut am Patienten verbringt, ihm Übungen zeigt oder verschiedene Techniken anwendet, hinaus. Die Leistung umfasst neben der Bereitstellung einer Struktur, in der sich Patienten wohl fühlen können, beispielsweise eine patientenorientierte Terminplanung in Abstimmung mit den Möglichkeiten einer Einrichtung. Als Dienstleistung wird von Therapeuten erwartet, Patienten in einem Behandlungsprozess dorthin zu geleiten, wo Selbstmanagement möglich ist.

Ausgehend von einem formalen Behandlungsvertrag durch das „Einlösen“ einer Verordnung zu Therapiebeginn wird ein Prozess initiiert, der über eine Befunderhebung hinausgehend die Fähigkeit eines Therapeuten zu einem Clinical-Reasoning beinhaltet. Dieses umfasst neben der Berücksichtigung der Wechselbeziehungen zwischen seiner Persönlichkeit, dem therapeutischen Erfahrungshintergrund und der individuellen Patientenbedürfnisse eine kritische Reflexion evidenzbasierter Erkenntnisse und deren Implementierung in die praktische Behandlungsstrategie, die stetig hinterfragt und damit professionell gesteuert wird. Dabei müssen auftretende Diskrepanzen in der Patienten-Therapeutenbeziehung, beispielsweise über unterschiedliche Vorstellungen der Behandlungsinhalte oder Schwierigkeiten hinsichtlich der Möglichkeiten der Krankheitsbewältigung, bedacht oder offen abgestimmt werden. Sie können eine begründete klinische Behandlungsstrategie durch verstärkte Einbindung in eine interdisziplinäre Zusammenarbeit mit Therapeuten anderer Fachgebiete oder Ärzten erforderlich machen. Demzufolge kann Qualitätssicherung als

„Oberbegriff für eine Reihe unterschiedlicher Maßnahmen verstanden werden, deren globales Ziel darin besteht, die konkrete Versorgungspraxis systematisch und kontinuierlich zu hinterfragen, zu bewerten und zu fördern – und im Falle von Mängeln oder Schwachstellen – zu verbessern“ (Nübling & Schmidt, 1998, S. 5).

Die Voraussetzungen für qualitätssichernde Maßnahmen können an der Einsicht auf Führungsebene sowie an der Einbindung und Motivation von Mitarbeitern und Therapeuten zu eigenverantwortlichem Handeln gemessen werden. Die Etablierung einer Unternehmensphilosophie und deren Verinnerlichung von Seiten aller Beteiligten bildet eine wichtige Grundlage für den Erfolg einer Einrichtung. Die mög-

lichen Ziele von Qualitätssicherung beschreiben Dorenburg und Schliehe (1998) durch die Gewährleistung und Verbesserung

1. der Wirksamkeit von Leistungen,
2. einer bedarfs- und patientengerechten Versorgung,
3. einer Transparenz der Leistungserstellung,
4. der Erschließung von Wirtschaftlichkeitsreserven.

Fehlinterpretationen der Ziele sind in der Praxis keine Seltenheit. So wird Transparenz mit Kontrolle verwechselt oder die Verbesserung der Wirksamkeit durch Einführung von Testverfahren als „Überfrachtung“ oder Verlust an Therapiezeit beschrieben, in der man sich nicht dem Patienten widmen könne. Dabei fehlt es oft lediglich an geeigneten Argumenten, eine standardisierte Befunderhebung zu rechtfertigen, deren Ergebnisse von allen anderen Teilnehmern der Behandlungskette verstanden würden.

Dorenburg & Tiefensee (2000) unterscheiden in der systematischen Beschreibung von Qualitätssicherung Dimensionen, die über die Behandlungsebene hinausgehen und nennen

1. die Strukturqualität,
2. die Prozessqualität und
3. die Ergebnisqualität.

Unter *strukturellen Merkmalen* sind z.B. personelle Bedingungen, räumliche Gegebenheiten oder die technische Ausstattung zu verstehen. *Prozesse* erfassen die „Angemessenheit diagnostischer und therapeutischer Leistungen sowie die Beurteilung der sachgerechten Durchführung“ (Dorenburg & Tiefensee, 2000, S. 203). Die *Ergebnisqualität* spiegelt den Grad der Therapiezielerreichung wider, der über Assessmentverfahren aus Patientensicht und aus Sicht der Leistungserbringer erfasst werden kann. In dieser Betrachtungsweise „liefern Assessmentverfahren einen wichtigen Beitrag zur Qualitätssicherung von Praxis und Forschung in der Medizin und Gesundheitsversorgung“ (Biefang et al., 1999, S. 15), denn alle Maßnahmen sollten darauf ausgerichtet sein, ein möglichst gutes Therapieergebnis zu erreichen, was nur auf der Basis effizienter Strukturen und Prozesse möglich erscheint.

Allerdings gehen *einschlägige Verfahren der Qualitätssicherung* über den Einsatz von Assessmentverfahren hinaus und umfassen Verfahren, die auf allen drei Qualitätsebenen (Struktur-, Prozess-, Ergebnisebene) zum Einsatz kommen. Tabelle 8 macht einige dieser Verfahren deutlich. Dazu gehören neben Ist-Analysen zu Struk-

turen und Prozessen therapeutischer Einrichtungen auch Ergebnisvergleiche mit äußeren Standards oder das Benchmarking (Klinikvergleich). Probleme und entdeckte Qualitätsmängel können über Verfahren wie dem Qualitätszirkel aufgegriffen und im Sinne einer Problemlösung bearbeitet werden. Die Problemlösung wird in die Struktur, in den Prozess oder als Ergebnisüberprüfung in der Einrichtung implementiert und hinsichtlich ihrer Effektivität evaluiert (Routinemonitoring, Evaluation, Tab. 8).

Tab. 8. Verfahren der Qualitätssicherung in den einzelnen Stadien des Qualitätssicherungsprozesses (Dorenburg & Tiefensee, 2000, S. 205).

Stadien der Qualitätssicherung	Verfahren der Qualitätssicherung
Ist-Analyse	Erhebungsbögen zu Struktur-, Prozess- und Ergebnisqualität
Aufbereitung und Rückmeldung der Daten	Vergleich mit Standards Benchmarking
Problemanalyse und Erarbeitung von Problemlösungen	Qualitätszirkel
Evaluation der Problemlösung	Erneutes Routinemonitoring Katamnesen Evaluationsforschung

Als Bestandteil der Ergebnisqualität sind Testverfahren damit in einen übergeordneten Rahmen von Qualitätssicherung eingebettet, dessen zuletzt gesetzliche Legitimation sich aus der Verankerung im §135a SGB V (DGQ & ZVK, 2003) ergibt. Für alle Einrichtungen des Gesundheitswesens empfahl die Gesundheitsministerkonferenz 1999 entsprechend die Einführung eines an dem Stand der Wissenschaft und Technik orientierten Qualitätsmanagements bis zum 1. Januar 2005. Damit sind Leistungserbringer zur Sicherung und Weiterentwicklung der Qualität und der von ihnen erbrachten Leistungen verpflichtet, was in der Physiotherapie mit einer Dokumentationspflicht in Form einer regelmäßigen Verlaufskontrolle und den Empfehlungen zum Einsatz von Bewertungs- und Dokumentationsverfahren einherging (DGQ & ZVK, 2003).

Zusätzlich wurde mit Inkrafttreten der Heilmittelrichtlinien im Jahre 2001 eine interdisziplinäre Zusammenarbeit zur verbesserten Kommunikation der an der physiotherapeutischen Behandlung beteiligten Fachleute geregelt (Beyer, 2001/2002). Diese sah regelmäßige Rückmeldungen der Therapeuten an den behandelnden Arzt in Form von Therapieberichten vor. Dass Berichtsformen über den Behandlungsstand von Patienten nicht ohne Testverfahren denkbar sind, liegt auf der Hand. Eine generell verpflichtende Berichterstattung wurde allerdings 2003 zugunsten eines nur noch nach Aufforderung des behandelnden Arztes geäußerten Wunsches verändert. Neben Kostengründen zeigten Ärzte nur wenig Bereitschaft, im dichten Arbeitsalltag zusätzlich Berichte zu lesen. Den Ansprüchen bei der

Umsetzung von Qualitätssicherung tun sich damit zumindest kurzfristig Grenzen auf. Mittel- oder langfristig wird einer standardisierten Therapieergebnisbeurteilung nicht ausgewichen werden können, denn ihre Vorteile liegen auf der Hand: Sie führen zu einer verbesserten Datenqualität mit der Möglichkeit zu einer vergleichenden Berichterstattung oder sogar zu einem Vergleich unterschiedlicher Therapieformen miteinander (Broda & Beckmann, 2000).

Eine standardisierte Therapieergebnisbeurteilung sichert unter der Benutzung von Assessmentverfahren die externe Qualität und steigert damit die Wettbewerbsfähigkeit einer therapeutischen im Vergleich zu anderen Einrichtungen. Künftig könnten über Kollektivverträge hinaus Krankenkassen Einzelverträge mit Heilmittelerbringern schließen, um über zusätzliche Wirtschaftlichkeitsanreize einen Wettbewerb zu ermöglichen, der bessere Qualität fordert und fördert (Esser, 2006).

Versicherte der gesetzlichen Krankenversicherungen haben zwar einen Anspruch auf ärztliche Behandlung durch die Versorgung mit Heilmitteln (§32 SGB V), aber die Selbstverständlichkeiten von Verordnungen zu physiotherapeutischer Behandlungen gehören der Vergangenheit an, und so müssen Physiotherapeuten dokumentieren, wenn sie ihre gegenwärtige Stellung als Heilmittelerbringer künftig behalten wollen.

Assessmentverfahren finden neben der Einbindung in die Qualitätssicherung übergeordnet auch Eingang dort, wo Leitlinien ihre Wirkungen entfalten. Auf die Bedeutung von Leitlinien im Zusammenhang mit Testverfahren soll im Folgenden kurz eingegangen werden.

### **2.6.6 Testverfahren im Kontext von Leitlinien**

Unter *Leitlinien* werden systematisch entwickelte Darstellungen und Empfehlungen verstanden, deren *Hauptziel* es ist, unter Berücksichtigung der vorhandenen Ressourcen gute klinische Praxis zu fördern und die Öffentlichkeit darüber zu informieren. So sollen Ärzte und Patienten bei der Entscheidung über angemessene Maßnahmen der Krankenversorgung unter spezifischen medizinischen Umständen unterstützt werden (AWMF, 2005). Diese Maßnahmen betreffen sowohl die Prävention, die Diagnostik und Therapie als auch die Nachsorge. Die Leitlinien geben dabei den Stand des Wissens über effektive und angemessene Krankenversorgung zum Zeitpunkt der "Drucklegung" wieder.

Darüber hinaus sind Leitlinien wesentlicher *Bestandteil des Qualitätsmanagements* im Gesundheitswesen und haben in diesem Zusammenhang die Aufgabe, das umfangreiche Wissen (wissenschaftliche Evidenz und klinische Erfahrung) zu speziellen Versorgungsproblemen zu bewerten, gegensätzliche Standpunkte zu

klären und unter Abwägung von Nutzen und Schaden das derzeitige Vorgehen der Wahl zu definieren. In Anbetracht der unausbleiblichen Fortschritte wissenschaftlicher Erkenntnisse müssen *periodische Überarbeitungen*, Erneuerungen und Korrekturen unternommen werden (AWMF, 2005).

In den Kriterien für die Qualität von Leitlinien werden zur Überprüfung der Anwendung von Leitlinien Verfahren aufgezeigt, mit denen die Akzeptanz von Empfehlungen in der Praxis ermittelt werden kann. So werden in den Leitlinien für Diagnostik und Therapie in der Orthopädie und der orthopädischen Chirurgie (AWMF & ÄZQ, 2005) zu den krankheitsspezifischen Leitlinien für die unteren Extremität systematisch

1. die klinische Diagnostik mit
  - a. *Inspektion* (z.B. Muskelatrophie, Beinlängendifferenz, Gangbild),
  - b. *Palpation* (z.B. Überwärmung, Erguss, Schwellung) und die
  - c. *spezifischen Funktions- und Schmerztests*,
2. die *apparative Diagnostik* (z.B. Röntgen, Sonographie, MRT, CT)
3. und die *klinischen Scores* aufgeführt.

Die Funktionstests sind nicht näher spezifiziert, entsprechend auch nicht standardisiert. Ein expliziter Einsatz von Testverfahren in der Therapie wird auch nicht ausgesprochen, es werden lediglich verschiedene klinische Scores für den wissenschaftlichen Vergleich empfohlen.

Dennoch sind Assessmentverfahren implizit Voraussetzung für die Effektivität von Leitlinien. Nach AWMF und ÄZQ (2005) unterliegt die Wirksamkeit von Leitlinien den Einflüssen

- ihrer Akzeptanz bei den Benutzern,
- der Zuverlässigkeit ihrer Empfehlungen,
- der systematischen Auswahl und Bewertung der Evidenz,
- der strukturierten Konsensusfindung,
- der Orientierung am Ergebnis für den Patienten (Outcome-Bewertung)
- der Abwägung von Nutzen und Risiken (Entscheidungsanalysen) und
- der Nachvollziehbarkeit des Versorgungsablaufs.

Dass die Wirksamkeit von Leitlinien auch am Therapieergebnis von Patienten orientiert gemessen wird, setzt bei der Outcome-Bewertung den Einsatz von Assessmentverfahren voraus.

### **2.6.7 Zur Bedeutung von Performance-Tests**

In Sport und Rehabilitation wurden vor allem ab Mitte der achtziger Jahre verschiedene motorische Testverfahren zur Beurteilung der Funktion der unteren Extremität entwickelt (Bös, 2001). Guralnik et al. (1989) bewerten dabei Performance-Tests mit den Vorzügen einer hohen Augenscheinvalidität, besserer Reproduzierbarkeit, größerer Veränderungssensitivität, insbesondere wenn quantitative Messungen vorgenommen würden. Lin et al. (2001) erheben Performance-Tests nicht zum Goldstandard, empfehlen sie aber ergänzend im Zusammenhang mit krankheitsspezifischen Fragebögen. Auch Guralnik, Simonsick, Ferrucci, Glynn, Berkman, Blazer, Scherr und Wallace (1994) sind der Ansicht, dass Performance-Tests über Selbsteinschätzungsfragebögen hinaus eine weitere wertvolle Quelle darstellen, um aus der objektiven Überprüfung von Aktivitäten Informationen zu gewinnen.

Bührlen et al. (2002) äußern die Problematik von Fremdbeurteilungsinstrumenten und beobachteten sehr häufig systematische Überschätzungen von Rehabilitationserfolgen durch Ergebniseinschätzungen von Ärzten. Ursächlich führen sie an, dass Patienten gegen Ende der Rehabilitation möglicherweise eine Verbesserung angeben, um Dankbarkeit für die Behandlung auszudrücken.

Guralnik et al. (1989) proklamieren für die Zukunft einen umfassenden Nutzen aus Performance-Tests. Dies vor allem unter der Perspektive der Identifikation von Personen mit einem hohen Funktionsstatus

- im interkulturellen Vergleich,
- als Outcome-Indikator bei Interventionsstudien,
- bei der Darstellung von Veränderungen des Funktionszustandes im Lebensverlauf und schließlich
- als Quelle bedeutsamer Informationen für Ärzte.

Glatz, Anneken, Heipertz, Schüle, Mozdzanowski und Schian (2006) liefern Hinweise, dass durch Leistungstests im Hinblick auf eine Beurteilung der arbeitsbezogenen Leistungsfähigkeit von Patienten eine größere Treffsicherheit zu erwarten ist, als durch eine ausschließliche Beurteilung von Patienten aufgrund der Aktenlage, der Anamnese oder einer körperlichen Untersuchung.

Nachteile von Performance-Tests sind dem gegenüber ein höherer Zeitbedarf, eine notwendige Einweisung der Untersucher in die Handhabung des jeweiligen Instruments sowie eine erhöhte Verletzungsgefahr. Außerdem spiegeln nach Guralnik et al. (1989) die Ergebnisse einfacher Tests nicht unbedingt die Adaptation an die Umgebung im täglichen Leben wider. Einschränkungen deuten auch Wind, Goutteborge, Kuijer, Sluiter und Frings-Dresen (2006) bei medizinisch nicht geklärten Problematiken, negativer Selbstwahrnehmung und bei Problemen im sozialen Umfeld der Patienten an. Ein komplementärer Einsatz von Instrumenten der Fremdbeurteilung und Selbstbeurteilung ist demnach empfehlenswert (Kohlmann, 2004b).

### **2.6.8 Zur Systematik von Performance-Tests**

Die Absicht einer inhaltlichen Zuordnung von motorischen Funktions- oder Performance-Tests zeigt nach Bührlen et al. (2002) eine Grundproblematik bestehender Systematiken auf. Sie beinhalten bisher keine ICF dimensionskonformen Operationalisierungen in Anlehnung an alltagsnahe Tätigkeiten. Sowohl nach einer Literaturrecherche wie auch nach den Ergebnissen von Expertendiskussionen im Rahmen der Operationalisierung von Therapiezielen bleibt die Auswahl von Funktionsbereichen zu Funktionstests in gewisser Weise willkürlich (Bührlen et al., 2002).

Eine Systematik zu Tests orientiert sich an der in der Forschung, Lehre und Praxis des Sports geprägten Terminologie der motorischen Hauptbeanspruchungsformen nach Hollmann & Hettinger (1980). Sie umfassen die

1. Koordination,
2. Flexibilität,
3. Kraft,
4. Schnelligkeit und die
5. Ausdauer.

In einer dieser nicht völlig entsprechenden Systematik wird in dem „Handbuch motorische Tests“ von Bös (2001) – dem wohl umfangreichsten deutschsprachigen Werk zu motorischen Tests – unterschieden zwischen:

1. Motorischen Verhaltenstests (z.B. Kondition, Fitness, Koordination),
2. motorischen Funktionstests (z.B. Haltung, Muskelfunktion, Ausdauer),

### 3. Fragebögen zur körperlich-sportlichen Aktivität und sportpsychologischen Diagnose.

Die Testformen dienen der Erfassung ein oder mehrerer Körperstrukturen, bestimmter Funktionen und Erkrankungen oder definierter Zielgruppen. Sie reichen von einer einfachen Beweglichkeitsmessung oder dem Einbeinstand, über Testbatterien mit zahlreichen Items bis hin zu komplexen arbeitsbezogenen Leistungsüberprüfungen an der Schnittstelle zwischen medizinischer und beruflicher Rehabilitation. Hier sind vor allem das System der *Evaluation der funktionellen Leistungsfähigkeit* (EFL) nach Susan Isernhagen (1988) und das *Arbeitssimulationsgerät ERGOS* (Kaiser, Kersting & Schian, 2000) zu nennen. Sie besitzen eine höhere Validität, sind aber in einem physiotherapeutischen Setting nur mit großem Zeitaufwand, hohen Kosten sowie umfangreicher technischer oder personeller Ausstattung umzusetzen. In kleineren Praxen oder auch vielen ambulanten Rehabilitationseinrichtungen sind sie nicht zu leisten.

Eine Literatursichtung zum Thema Performance-Tests ergab eine Reihe von Testverfahren mit einer Vielzahl unterschiedlicher Messintentionen und methodischer Ansätze, eine vollständige Darstellung erwies sich als unmöglich (Bührlen, et al. 2002). Im folgenden Kapitel sollen daher ausgewählte Verfahren beschrieben werden.

#### **2.6.9 Forschungsstand und Forschungsdefizit**

So umfangreich sich der gegenwärtige Literaturstand zum Thema Performance-Tests darstellt, so eingeschränkt ist das Auffinden geeigneter Aktivitätsüberprüfungen für die untere Extremität, die für ein postoperatives Patientenkontinuum in einer Altersspanne von 19-75 Jahren anwendbar sind.

Der *Timed-Up & Go-Test*, in diesem Zusammenhang einer der am häufigsten eingesetzten Performance-Tests, stellt sich als Weiterentwicklung des Up & Go-Test von Mathias, Nayak und Isaacs (1986) dar. Sie zielten mit der Entwicklung des Up & Go-Tests auf die Erfassung grundlegender Aktivitäten der Beine wie von einem Stuhl oder Bett aufstehen, auf die Toilette gehen und zu Fuß gehen. Eine weitere Darstellung von Verfahren beschränkt sich auf die Überprüfung von Aktivitäten, wobei Aspekte des Gleichgewichtes nicht ausgeschlossen werden können.

*Testbatterien*, die eine Überprüfung des Bewegungsapparates der unteren und oberen Extremität erfassen, sind in Tabelle 9 dargestellt. Die *Physical Capacity Evaluation* (PCE, Daltroy, Phillips, Eaton, Larson, Partridge, Logigian & Liang, 1995) testet bei einer Zielgruppe von 17-90-jährigen neben Funktionen der oberen Extremität zwei Funktionen der unteren Extremität (Tandemstand, 30s und auf den

Füßen tippeln) sowie die zwei Aktivitäten des Strümpfe Anziehens und des Timed-Up & Go-Test.

Reuben & Siu (1990) zielen mit ihrem *Physical Performance Test* (PPT) auf 65 bis 99-jährige Menschen. In neun Items werden fein- und großmotorische Funktionen der oberen und unteren Extremitäten getestet. Für die untere Extremität sind enthalten: Das Aufheben eines Geldstückes vom Boden, eine 360 Drehung im Stand, ein 15,2 m (50 ft) Gehtest, das Steigen auf eine Treppenstufe sowie das Treppensteigen, wobei die Anzahl der Stufen gezählt wird (Tab. 9).

Tab. 9. *Performance-Tests für die obere/untere Extremität (Ziel: Gleichgewicht, Funktionszustand).*

Quelle	Testaufgaben	Zielgruppe/ Itemzahl
PCE, Daltroy et al. (1995)	Greifkraft, Schreiben, Umdrehen von Karteikarten, mit einem Schlüssel ein Schloss verschließen, Purdue Pegboard Test (Tiffkin & Asher, 1948)	Ab 65 J. / 13
Harding et al. (1994)	10-Min.-Gehtest, Ganggeschwindigkeit, Treppauf- und -absteigen, Einbeinstand, Aufstehen und Hinsetzen von einem Stuhl, Sit-ups, Kraftausdauer für die obere Extremität, Greifkraft, Lungenfunktionstest	Erwachsene / 9
BFT, Keitel et al. (1993)	z.B. aktives Spreizen der Beine im Liegen, Außenrotation im Hüftgelenk, Aufrichten aus dem Liegen, Aufstehen vom Hocker, 15s Zehen- und Fersenstand, Hocken, Einbeinstand, 30 m Gehen (Normzeit 20s), Treppensteigen (10 Stufen nach oben und unten, Normzeit je 7s)	Erwachsene / 24
PPT, Reuben & Siu (1990)	Schreiben, simuliertes Essen, ein Buch in ein Regal stellen, an-/ ausziehen, aufheben, 360 Drehung im Stand, 50 ft Gehtest, steigen auf eine Treppenstufe, Treppensteigen (Anzahl der Stufen wird gezählt)	über 65 J. / 7
ADL, Katz & Akpom (1974)	Baden, Ankleiden, auf die Toilette gehen, Transfer vom Bett auf einen Stuhl, Kontinenz und Essen	Erwachsene / 6
Barthel-Index, Mahony & Barthel (1965)	Essen (Nahrungsaufnahme), Wechsel von Bett in Rollstuhl, Gesicht waschen, Zähne putzen, kämmen, rasieren, Toilettenbenutzung, Baden, Gehen zu ebener Erde (bzw. Rollstuhlnutzung), Treppensteigen (auf und ab), Anziehen (einschl. Schnürsenkel zu binden oder Klettverschlüsse schließen), Darmkontrolle, Blasenkontrolle	Erwachsene / 10

Der *Bewegungsfunktionstest* nach Keitel, Hoffmann, Weber und Krieger (BFT, 1993, Tab. 9) beinhaltet 24 Einzeltests mit dem Ziel der Erfassung von Bewegungsfunktionen an der unteren und oberen Extremität. Neben funktionellen Bewegungen (z.B. aktives Spreizen der Beine im Liegen, Außenrotation im Hüftgelenk) sind auch Aktivitäten wie das Aufrichten aus dem Liegen, das Aufstehen vom Hocker, ein 15s Zehen- und Fersenstand, das Hocken sowie ein Einbeinstand, 30m Gehen (Normzeit 20s) oder ein Treppensteigen (10 Stufen nach oben und unten, Normzeit je 7s) enthalten.

Den *Barthel-Index* (Mahony & Barthel, 1965) und den ADL nach Katz und Akpom (1974, Tab. 9) als kombinierte Instrumente zwischen Selbsteinschätzung und Fremdbeurteilung, eignen sich für den Einsatz bei Patienten mit neurologischen Behinderungen, die insbesondere (noch) nicht in der Lage sind ihre Funktionsbeeinträchtigung selbst zu beurteilen.

Harding, Williams, Richardson, Nicholas, Jackson, Richardson und Pithers (1994, Tab. 9) zielen mit ihrer *Battery of measures for assessing physical functioning* auf die Erfassung von Funktionseinschränkungen durch chronische Schmerzen bei Erwachsenen. Der Test enthält acht Überprüfungen von vier Aktivitäten der Beine. Dazu gehören ein 10 Minuten-Gehtest, die Ganggeschwindigkeit, das Treppenauf- und -absteigen, der Einbeinstand sowie das Aufstehen und Hinsetzen von einem Stuhl ohne Armlehne (Tab. 9).

Der *Physical Performance Test* nach Simmonds, Olson, Jones, Hussein, Lee, Novy und Radwan (PPT, 1998) zielt mit seiner Überprüfung der funktionellen Kapazität auf Erwachsene mit lumbalen Rückenschmerzen. In acht Aufgaben sind vier Tests für die untere Extremität vorgesehen: Der Timed Up & Go-Test, das Aufstehen von einem Hocker (10x so schnell wie möglich), 50 Fuß so schnell wie möglich gehen und ein Fünf-Minuten-Gehen (die Strecke wird gemessen).

Die *Short Physical Performance Battery* (SPPB, Guralnik, Simonsick, Ferruci, Glynn, Berkman, Blazer, Scherr & Wallace, 1994, Tab. 10) prüft bei einer Zielgruppe von 60 bis 99-jährigen das Gleichgewicht, die Ganggeschwindigkeit sowie die Kraft und Ausdauer der unteren Extremität. Die einzelnen Tests umfassen das Stehen mit geschlossenen Augen (10s) im Tandem- und Halbtandemstand, die Ganggeschwindigkeit über 8 Fuß sowie die Kraft- und Ausdauerleistung über das fünfmalige Aufstehen von einem Stuhl ohne Zuhilfenahme der Arme (Tab. 10). Die Tests erscheinen, wie die weiteren in Tabelle 11 dargestellten Verfahren, für Jüngere als ungeeignet.

Tab. 10. *Performance-Tests für die untere Extremität (Ziel: Gleichgewicht, Funktionszustand).*

Quelle	Testaufgaben	Zielgruppe/ Itemzahl
PPME, Winograd et al. (1994)	Aufsetzen im Bett, vom Bettrand zu einem Stuhl gehen und sich setzen und aufstehen, auf einen Stuhl setzen und aufstehen (5x), Sitzpositionen halten, Stufe steigen, 5m Gehtest	Ab 60 J. /6
SPPB, Guralnik et al. (1994)	Tandem- und Semi-Tandemstand, Side-by-side-Stand, 8 ft-walk, 5x-rise-from-Chair	über 71 J. /5
POMA, Tinetti (1986)	Balanceübungen im Sitzen, Stehen mit geschlossenen Augen, Einbeinstand, Kopfdrehung, Rückenstreckung im Stehen, Hochreichen an ein Regal, Bücken und Hinsetzen	Ältere /13

Tab. 11. Performance-Tests zur Beurteilung der Kniefunktion nach Knieverletzungen (Bös, 2001, S. 354 f).

Quelle / Zielsetzung	Testaufgaben	Bewertung	Bemerkungen/Testgüte
Marshall, Fetto & Botero (1977)	Kniebeuge, Sprünge, Entengang	subjektiv, dichotom, „kann“/„kann nicht“	Einer der ersten dokumentierten Tests zur Beurteilung der Kniefunktion, keine teststatistischen Angaben
Jensen, Conn, Hazelrigg & Hewett, (1984)	Laufen im Stand, Hüpfen auf der Stelle, Entengang	subjektiv, dreistufige Skala: normal/asymmetrisch/nicht möglich	keine teststatistische Absicherung
Tegner, Lysholm, Lysholm & Gillquist (1986)	Achterlauf, einbeiniger Weitsprung, Lauf auf Wendeltreppe, Lauf auf Steigung/Gefälle	metrisch, Erfassung von Weiten (einbeiniger Weitsprung) und Zeiten, Vergleich mit dem kontralateralen Bein	Normwerte wurden an einer Vergleichsgruppe gesunder Sportler erhoben, Normwerte wurden je nach Test aber auch von 35%-62% der Sportler erreicht, Teststatistische Absicherung der Reliabilität durch Ageberg et al. (1998)
Barber, Noyes, Mangine, McCloskey & Hartmann (1990)	Einbeiniger Weitsprung, einbeiniger Jump & Reach Sprung, einbeinige Sprünge über eine Distanz von 6 m, Läufe über 6 m	metrisch, Erfassung von Weiten und Höhen; Vergleich der Messdaten mit einem Symmetrieindex	Keine Angaben zur Teststatistik; der Symmetrieindex wurde an gesunden Probanden ermittelt, als normal gelten Indizes > 85%, bei einzelnen Tests erreichten 50% der Patienten normale Indizes, die Autoren empfehlen daher die Durchführung von mindestens zwei verschiedenen Tests
Lephart, Pincivero, Giraldo & Fu (1994)	Kokontraktionstest, Carioaca-Test (seitl. Überkreuzen), Linienläufe	metrisch, Messung der benötigten Zeiten, Vergleich mit dem kontralateralen Bein	Für die Tests werden Reliabilitätskoeffizienten zwischen $r = .92$ und $r = .96$ angegeben
Wilk, Romaniello, Soscia, Arrigo & Andrews, (1994)	Einbeiniger Weitsprung, einbeiniges Hüpfen über eine Distanz von 6m, einbeiniges Überkreuzhüpfen über 6m	metrisch, Erfassung von Weite (einbeiniger Weitsprung) und Zeit; Vergleich mit dem kontralateralen Bein, vgl. mit Symmetrieindex (normal bei mind. 85%)	Keine teststatistischen Angaben, 64% der untersuchten Patienten hatten Symmetrieindizes > 85%
Risberg & Ekeland, (1994)	Einbeiniger Jump & Reach-Test. einbeiniger Dreisprung, einbeiniges „Treppauf- und Treppabhüpfen“, Side-Jump-Test	Metrisch, Erfassung von Zeiten und Weiten, Vergleich mit dem kontralateralen Bein	Text stellt methodische Arbeit zu funktionellen Tests dar, Autoren kritisieren die fehlende teststatistische Absicherung
PAR-Physical Activity Restriction, Rejeski, Ettinger, Schumaker, James, Burns & Elam, (1995)	6 Minuten-Gang, Treppauf- und Treppabgehen, Anheben und Tragen einer Last (22 kg), Ein-/ Aussteigen in PKW	Metrisch (Strecke, bzw. Zeiten)	Test-Reliabilität: 6 min-Gang: $r = .87$ ; Treppensteigen $r = .93$ , Last tragen: $r = .92$ , Ein-/Aussteigen PKW: $r = .88$ , Cronbachs $\alpha = .92$

Forschungsdefizite hinsichtlich *Performance-Tests* zeigen sich nach Bührlen et al. (2002, S. 77), die davon überzeugt sind, „dass ein praktikabler Performance-Test ein wichtiges Instrument für die Erfassung eines der wichtigsten Therapieziele in der Rehabilitation, nämlich die Funktionsfähigkeit im Alltag, darstellt“, gegenwärtig nicht vorhanden ist. Bührlen et al. (2002) hatten sich im Rahmen der Operationalisierung von Therapiezielen für die Überprüfung ihrer Erreichung, die Entwicklung eines „Test für Alltagsfunktion“ (TAF) zur Aufgabe gemacht. Mit ihm sollten krankheitsübergreifend Alltagsfunktionen und deren Einschränkungen erfasst werden können. Es zeigten sich allerdings sowohl bei der Literaturrecherche als auch in Expertentreffen divergierende Ansichten über die dimensionale Struktur von Alltagsfunktionen, so dass die Erprobung einer Testbatterie nicht in die Praxis umgesetzt werden konnte (Bührlen et al., 2002). Die Empfehlungen enthalten für die untere Extremität die folgenden Tests:

- Chair-rising-Test (Anzahl in einer Minute)
- Up & Go-Test
- Tandemgang auf einem Balancierbalken (Meter oder Fehler auf 2 m)
- aus dem Stand Hinlegen und wieder aufstehen (Zeit für 3x)
- ein Stockwerk hoch und wieder herunter Treppensteigen (Zeit in Sek.)
- Sprunghöhe (Hochhüpfen mit Maßband am Fuß)
- 3-Minuten-Stufentest (Anzahl der Stufen)
- 6-Minuten-Gehtest (Anzahl der Meter in sechs Minuten).

Zu den Defiziten bezüglich der Übersichten zu Tests von McGill (1994) und Pfeifer (1996) äußert Bös (2001, S. 354): „Den meisten Tests gemeinsam ist eine ungenügende teststatistische Absicherung und auch im Hinblick auf die Problematik der Sensitivität und Spezifität besteht noch Klärungsbedarf“. In der Auflistung von Bös (2001, S. 354 f., Tab. 11) – die keinen Anspruch auf Vollständigkeit erhebt – sind zur Beurteilung der unteren Extremität nach Kniebandverletzungen Testverfahren wie Treppabhüpfen, einbeinige Weitsprünge oder Entengang enthalten, die sich ausschließlich an Jüngere oder Leistungssportler wenden. Patienten, insbesondere nach Operationen, wollen und müssen zunächst einmal wieder grundlegende Aktivitäten beherrschen, die für die alltägliche häusliche oder berufliche Belastung notwendig sind: Aufstehen von einem Hocker oder aus einem Bett, Gehen, Hocken, Knien, Treppensteigen oder Laufen.

Ingesamt zeigen die meisten Tests nach Bös (2001, S. 354) „eine ungenügende teststatistische Absicherung“ sowie Klärungsbedarf bezüglich Sensitivität/Spezifität und Bührlen et al. (2002, S. 64) bescheinigen zahlreichen Verfahren: „Die Testgüte der Verfahren ist bei den einfachen Maßen teilweise sehr gut, teilweise aber auch unklar bzw. nicht mitgeteilt, teilweise zwar geprüft, aber mit mangelhafter Methodik.“

Auch zeigen sich Defizite in dem vom Zentralverband für Physiotherapeuten der Öffentlichkeit vorgestellten Handbuch zu Testverfahren. In diesem ist für die Überprüfung von Patienten mit muskuloskelettalen Beschwerden kein einziger Performance-Test gelistet (ZVK, 2006a). Lediglich im Zusammenhang mit neurologischen Beschwerden und Problemen innerer Organe werden mit dem Timed-Up & Go-Test von Podsiadlo und Richardson (1991), einem 6-Minuten-Gehtest, einem Gehstreckentest und dem Barthel-Index (Mahony & Barthel, 1965) einige Verfahren beschrieben.

Bezüglich einer möglichen Testauswahl für den Polla erschien bei der geplanten Zielgruppe von 19-75 Jahren die Erkenntnis von Sayers, Guralnik, Newman, Brach und Fielding (2006) wichtig, dass sich bei einer substantiellen Anzahl von noch gut konditionierten älteren Menschen Einschränkungen eher bei einem 400m Gehtest andeuteten, als bei der *Short Physical Performance Battery* (SPPB), also bei insgesamt höherer Leistungsanforderung (zum Vergleich s. Tab. 10: SPPB). Die Aufnahme einer Tätigkeit in den Polla, der über die Dauer von wenigen Sekunden hinaus geht und strukturell höhere Belastungen abfordert, war daher dringend notwendig.

Eine Beschreibung weiterer Instrumente könnte fortgesetzt werden. Bei aller Unterschiedlichkeit einzelner Testbatterien ergeben sich im Hinblick auf mögliche Aktivitäten für die untere Extremität immer wieder Gemeinsamkeiten. Die Ergebnisse der Literatursichtung und die ICF-Domänen *Mobilität* und *Selbstversorgung* werden im Folgenden als Entscheidungsgrundlage für die Konstruktion einer vorläufigen Performance-Test-Batterie dargestellt.

## 2.7 Ableitung der Items für den Polla

Die Auswahl von Aktivitäten erfolgt zunächst für eine Pilotversion, bei der die Praktikabilitätsprüfung während des Behandlungsprozesses im Vordergrund steht. Die Kriterien für die Ableitung zentraler Aktivitäten für den Polla ergeben sich 1.) aus den in der Literatursichtung am häufigsten angeführten Aktivitäten, 2.) den in den Domänen *Mobilität* und *Selbstversorgung* beschriebenen Aktivitäten sowie 3.) aus den Empfehlungen zu einem Performance-Test nach Bührlen et al. (2002). Weiterhin wird die Auswahl 4.) von den Gegebenheiten der therapeutischen Praxis bestimmt. Der physiotherapeutische Alltag muss bei der Überprüfung alltagsnaher Aktivitäten mit einem geringen zeitlichen und organisatorischen Aufwand auskommen. Ausgeschlossen werden daher Aktivitäten aus ICF-Kategorien, die sich einer objektiven Überprüfung während der Therapie entziehen (z.B. „klettern, steigen“, „sich mit Transportmitteln fortbewegen“) und Aktivitäten, die nicht im Zusammen-

Tab. 12. Zuordnung von ICF-Kategorien zu Tests der Literatur und Ableitung von Items für den Polla.

Code	ICF-Kategorie <sup>1</sup>	Literatur	Polla
<b>d410 Eine elementare Körperposition wechseln</b>			
d4100	<b>Sich hinlegen:</b> In oder aus einer liegenden Position zu gelangen, wie aufstehen oder sich hinsetzen...	Bührlen et al. (2002), Podsiadlo & Richardson (1991)	1. Hinlegen auf den Boden und wieder aufstehen
d4101	<b>Hocken:</b> In eine oder aus einer mit angezogenen Knien auf dem Hinterteil oder den Fersen sitzenden oder kauernenden Stellung zu gelangen...	Winograd et al. (1994), Keitel et al. (1993)	2. Hocken
d4102	<b>Knien:</b> In oder aus einer Position zu gelangen, bei der der Körper durch die Knie bei gebeugten Beinen unterstützt wird...	-	3. Hinknien
d4103	<b>Sitzen:</b> In oder aus einer sitzenden Position zu gelangen oder die Körperposition von einer sitzenden in jede andere Position zu wechseln...	Bührlen et al. (2002), Harding et al. (1994)	4. Aufstehen vom Hocker 5. 1-Minute-Up & Go-Test
d4104	<b>Stehen:</b> In oder aus einer stehenden Position zu gelangen oder die Körperposition von einer stehenden in jede andere Position zu wechseln...	Guralnik et al. (1994), Tinetti (1986)	6. Einbeinstand
<b>d415 In einer Körperposition verbleiben</b>			
d430	<b>Gegenstände anheben und tragen:</b> Einen Gegenstand anzuheben...	Reuben & Siu (1990)	7. Aufheben eines Gegenstandes
<b>d450 Gehen</b>			
d4500	<b>Kurze Entfernungen gehen:</b> Weniger als einen Kilometer zu gehen...	Bührlen et al. (2002)	8. 2-Minuten-Gehtest
<b>d455 Sich auf andere Weise fortbewegen</b>			
d4552	<b>Rennen:</b> Sich mit schnellen Schritten in der Weise zu bewegen, dass beide Füße gleichzeitig vom Boden abgehoben sind...	Jensen et al. (1984), Barber et al. (1990)	9. 2-Minuten-Lauftest
d4553	<b>Springen:</b> Durch Beugen und Strecken der Beine den Boden verlassen, wie auf einem Bein springen, hopsen, hüpfen und ins Wasser springen...	Bührlen et al. (2002), Marshall et al. (1977)	10. Einbeinhüpfen
<b>d460 Sich in verschiedenen Umgebungen fortbewegen</b>			
d4600	<b>Sich in seiner Wohnung umherbewegen:</b> ...Inkl. Sich von Stockwerk zu Stockwerk, auf einem Balkon, auf dem Hof, auf der Veranda oder im Garten bewegen...	Bührlen et al. (2002), Winograd et al. (1994), Risberg & Ekeland, (1994)	11. Treppenlaufen-Test
<b>d540 Sich kleiden</b>			
d5402 d5403	<b>Schuhwerk anziehen / ausziehen:</b> Die koordinierten Handlungen und Aufgaben durchzuführen, welche das Anziehen von Socken, Strümpfen und Schuhwerk betreffen...	Mahony & Barthel (1965)	12. Strümpfe aus- und wieder anziehen

<sup>1</sup>(Dimdi, 2005)

hang mit der unteren Extremität stehen (z.B. „*feinmotorischer Handgebrauch*“).

Aktivitäten mit geringer praktischer Relevanz für das geplante Stichprobenalter sind ausgeschlossen (z.B. „*sich verlagern*“). Leichte Items würden bei dem insgesamt mobilen Patientenklientel im Behandlungsverlauf schnell zu Bodeneffekten führen. Zu den alltagsnahen Aktivitäten, die in der Literatur häufiger anzutreffen sind, zählen der Einbeinstand, das An- und Ausziehen von Schuhen oder Strümpfen, das Aufstehen von einem Hocker sowie das Treppensteigen. Sie bieten in der frühen postoperativen Phase auch Jüngeren ausreichend Anspruch (vgl. 2.6.9). Weitere Verfahren wie der Timed Up & Go-Test, die Ganggeschwindigkeit und ein 10 Minuten Gehetest sind ebenfalls häufige Bestandteile. Die Tabelle 12 zeigt die relevanten ICF-Kategorien, die in der Literatur erwähnten Aktivitäten und die in den Polla aufgenommenen Items.

Nicht operationalisiert ist das Heben schwerer Gegenstände. Hier wäre eine zu starke Teststandardisierung notwendig, die sich ergibt, wenn das Heben über die Beinbeugung geprüft werden sollte. Der natürliche Bewegungsablauf des Hebens erfolgt nach Operationen erfahrungsgemäß über die gestreckten Beine und würde daher keine zusätzlichen Informationen zum Hocken oder Hinknien liefern.

Den Empfehlungen von Bührlen et al. (2002) wird im Hinblick auf den Tandemgang auf einem Balancierbalken, die Testung der Sprunghöhe (mangelnder Alltagsorientierung) sowie den 3-Minuten-Stufentest (Überlastungsgefahr) nicht entsprochen werden. Außerdem fehlen aus Autorensicht die wichtigen Aktivitäten des Knien oder Hockens vollständig.

Insgesamt enthält die Testbatterie 12 Aktivitäten (Tab.12). Das Knien findet sich in keiner Testbatterie der Literatur, dessen Überprüfung scheint aber aus therapeutischer Sicht wichtig für Patienten in handwerklichen Berufen (z.B. Fliesenleger). Von daher wurde für eine Pilotphase vom Hocken abgesehen und vorerst das Knien aufgenommen. Ebenso wurde das Strümpfe an- und ausziehen zunächst herausgelassen und das Aufstehen vom Hocker aufgrund der wichtigen Beugebelastung in zwei Versionen aufgenommen (Anzahl der Male Aufstehen in 10 Sekunden und ein 1-Minute-Up & Go-Test). Zur Vermeidung von Deckeneffekten im spätpostoperativen Zeitraum wurde ein Dreibeinsprung (Triple-Jump-Test) hinzugefügt. Dieser ist keine Aktivität im alltäglichen Sinne, bietet aber noch bei jüngeren Rehabilitanden im Übergang zur sportlichen Belastung oder bei konservativ behandelten Patienten ausreichend Anspruch. Grundsätzlich sollen alle Tests in der Pilotphase, wie von Bührlen et al. (2002) empfohlen, quantitativ überprüft werden (Zeit, Meter, Anzahl in einer bestimmten Zeit). Vor der Darstellung der Ergebnisse der Pilotphase werden im Folgenden die Hypothesen hergeleitet.

## 2.8 Herleitung der Hypothesen

Die Überprüfung des Entwicklungsstands von Patienten während einer physiotherapeutischen Behandlungsreihe ist aus unterschiedlicher Sicht möglich. Die Messung kann aus Patientenperspektive durch schriftlich-subjektive Selbsteinschätzungen oder aus therapeutischer Sicht beispielsweise durch klinische Parameter erfolgen.

Unter der Perspektive einer bio-psychozialen Sichtweise des Patienten vor seinem gesamten Lebenshintergrund, geht die Beurteilung von Therapie über eine Anwendung klinischer Parameter und damit zumeist struktureller oder funktioneller Messinstrumente hinaus und umfasst die Betrachtung von Behinderungen auf Aktivitäts- und Teilhabeebene in Anlehnung an die Dimensionen der ICF.

Von besonderer Bedeutung ist ein möglicher Informationsgewinn darüber, inwieweit die Bewältigung von Aktivitäten des Alltags – gemessen über Performance-Tests – und der allgemeine Gesundheitszustand von Patienten miteinander *zusammenhängen*. Die sich ergebenden Fragen zielen darauf, ob Patienten mit einem hohen Testergebnis im Performance-Test auch hohe Werte im Gesundheitszustand erzielen.

Darüber hinaus kommt der Darstellung von Veränderungen zwischen verschiedenen Messzeitpunkten eine wichtige Bedeutung zu. Kann ein Messinstrument sowohl nach kürzeren Behandlungsintervallen von sechs Wochen, als auch nach längeren Zeiträumen von über drei Monaten, Veränderungen statistisch dort belegen, wo nur noch kleine Fortschritte im Genesungsprozess zu erwarten sind? Zeigt ein Patient also beispielsweise, dass er sich zum Therapieende ohne Einschränkungen hinlegen oder schneller zwei Treppenabsätze bewältigen kann, dann verdeutlicht diese Feststellung Unterschiede zu verschiedenen Zeitpunkten und kann als *Änderungssensitivität* aufgefasst werden. Bei der Konstruktion der Testbatterie wurde zur Vermeidung von Boden- und Deckeneffekten auf ein breites Aktivitätsspektrums mit unterschiedlichen Schwierigkeitsgraden Wert gelegt. Im Zusammenhang mit der Erhebung von Einschränkungen wie Schmerzen, Hilfen, Hilfsmittel oder Ausweichbewegungen könnte der Polla, da er krankheitsspezifisch ausgerichtet ist, Veränderungen leichter als die *Körperliche Funktionsfähigkeit* (SF-36) abbilden und damit eine größere Änderungssensitivität aufweisen.

Leistet ein Performance-Test die Aufgabe der Änderungssensitivität nicht, können Veränderungen im Rahmen von Behandlung nicht aufgezeigt werden und stellen zumindest nach extern, gegenüber anderen Teilnehmern der Behandlungskette, die therapeutische Arbeit in Frage. Aus den Generalhypothesen abgeleitet sind die folgenden Hypothesen Grundlage der statistischen Auswertung:

*Korrelationshypothese*

Es bestehen Zusammenhänge zwischen dem objektiven Performance-Test, der subjektiven Selbsteinschätzung zu Einschränkungen von Aktivitäten und der gesundheitsbezogenen Lebensqualität (SF-36) zu den Messzeitpunkten Therapiebeginn, Therapieende und Follow-up.

*Unterschiedshypothese*

Bei Mittelwertvergleichen zwischen dem Therapiebeginn, dem Therapieabschluss und dem Follow-up zeigt der Performance-Test gegenüber dem SF-36 und der Selbsteinschätzung von Aktivitäten größere Unterschiede auf, er ist damit änderungssensitiver.

Vor der Hypothesenüberprüfung und der Betrachtung von Testgütekriterien zum Performance test of lower limb (Polla) werden im folgenden Kapitel zur Methode zunächst die Vorarbeiten in einer Pilotphase dargestellt, diskutiert und die Einflussnahme der Ergebnisse auf die Konstruktion des Polla dargelegt. Im Weiteren werden das Studiendesign, die unabhängigen und abhängigen Variablen sowie die Erhebungsinstrumente betrachtet.

## 3 METHODE

### 3.1 Pilotstudie

Die vorbereitenden Maßnahmen zur Validierung umfassen eine explorative Erprobung des Polla im Hinblick auf seine Praktikabilität und augenscheinliche Validität. Aus einer anschließenden Auswertung und Diskussion der Pilotphase wird die endgültige Form des Performance-Tests abgeleitet.

#### 3.1.1 Erprobung und Ergebnisse

Die aus zwölf Einzelaktivitäten bestehende Pilotversion wurde an 22 Patienten in der postoperativen orthopädisch-traumatologischen Nachbehandlung zu zwei Zeitpunkten, als Test und als Retest, durchgeführt. Die Retests erfolgten in einem Abstand von 3-4 Tagen zum Test. Der Testzeitpunkt betrug bis zu drei Monate nach Operation. Es nahmen 12 weibliche und 10 männliche Personen im Alter von 39 Jahren ( $\pm 17,2$ ) teil, wobei die Testzeit ohne Erläuterung etwa 30 Minuten betrug. Die Testbatterie umfasst Aktivitätsüberprüfungen, die entsprechend den Empfehlungen nach Bührlen et al. (2002) quantitativ gemessen wurden:

1. Einbeinstand (Abbruch nach 60 Sekunden)
2. Aufstehen vom Hocker (Anzahl in 10s)
3. Aufheben (Zeit für 5x)
4. Hinlegen/Aufstehen (Zeit für 3x)
5. 1-Minute-Up & Go-Test (Anzahl in einer Minute)
6. 2-Minuten-Gehtest (Meter in 2 Minuten)
7. Treppaufsteigen (einbeinig auf Stepp: Anzahl in 30s.)
8. Lauftest (Meter in bis zu 6 Minuten)
9. Einbeinsprung (Achterspringen: Anzahl in 30s)
10. Triple-Jump-Test (Einbeiniger Dreisprung, vgl. ANHANG 8)
11. Treppenlaufen-Test (Ein Stockwerk hoch/runter, in Sekunden)
12. Hinknien (möglich/nicht möglich).

Von 528 möglichen Tests (Test: 264 und Retest: 264, Tab. 13) konnte etwa die Hälfte (Test: 126 und Retest: 133) durchgeführt werden. 176 Tests waren aus Gründen des frühen postoperativen Zeitpunkts oder aufgrund von Schmerzen nicht möglich (Tab. 13). Der reguläre Arbeitsprozess, die Patientenbedürfnisse zur Behandlung, aber auch organisatorische Probleme (z.B. abgesagte Termine oder Krankheit von Therapeuten) ermöglichten keine durchgängige Durchführung der Tests, woraus sich keine Angaben zu 81 Tests (Tab. 13) ergaben.

Tab. 13. *Ergebnisse der Pilotphase: Tests in der Reihenfolge deren Durchführung nicht möglich (n.m.) oder möglich (m.) war. Daneben die Anzahl überprüfter Tests (Test), Keine Angaben (K.A.) sowie das Gesamttestaufkommen (Ges).*

N=22 Aktivität	Zeitpunkt Test					Zeitpunkt Retest				
	n.m.	m.	Test	K.A.	Ges	n.m.	m.	Test	K.A.	Ges
1. Triple-Jump-Test	21	1	22	-	22	17	1	18	4	22
2. Lauftest	20	2	22	-	22	13	2	15	7	22
3. Einbeinsprung	11	4	15	7	22	11	4	15	7	22
4. Treppenlaufen-Test	14	6	20	2	22	7	9	16	6	22
5. Treppaufsteigen	9	12	21	1	22	7	11	18	4	22
6. Hinknien	7	9	16	6	22	7	9	16	6	22
7. Up & Go-Test	5	12	17	-	22	4	15	19	3	22
8. Gehtest	5	14	19	3	22	4	12	16	6	22
9. Hinlegen	3	15	18	1	22	2	17	19	3	22
10. Aufstehen	3	17	20	2	22	1	17	18	4	22
11. Einbeinstand	2	14	16	2	22	1	18	19	3	22
12. Aufheben	1	20	21	1	22	1	18	19	3	22
<b>Gesamt</b>	<b>101</b>	<b>126</b>	<b>227</b>	<b>25</b>	<b>264</b>	<b>75</b>	<b>133</b>	<b>208</b>	<b>56</b>	<b>264</b>

Tab. 14. *Deskriptive Ergebnisse der Pilotphase.*

N=22 Aktivität	Zeitpunkt Test			Zeitpunkt Retest		
	N <sup>1</sup>	M	s	N	M	s
1. Aufheben (für 5x in Sek.)	20	10,6	4,1	18	9,8	4,5
2. Hinlegen (für 3x in Sek.)	18	27	17	17	20,7	10,8
3. Einbeinstand in Sekunden	16	19,4	18,2	18	31,8	22,6
4. Up & Go Test (Anzahl in 1 Min.)	17	10,1	3,3	15	10,5	3
5. Aufstehen (Anzahl in 10 Sek.)	15	24	10,6	17	20,7	7,2
6. Gehtest (Meter in 2 Min.)	14	141	42,4	12	153,9	46,1
7. Treppaufsteigen (Anzahl 30 Sek.)	12	43,6	16,4	11	46,3	20,1
8. Treppenlaufen-Test (in Sek.)	6	23	4,8	9	22,3	10,6
9. Einbeinsprung (Anzahl in 30 Sek.)	4	24,3	12	4	15,5	9,0
10. Lauftest (Meter in 6 Min.)	3	456	353,5	2	950,0	188,1
11. Triple-Jump-Test (in Metern)	1			1	Differenz: 108cm	
12. Hinknien	Nur qualitativ an N=9 überprüft					

Anm.: <sup>1</sup>Anzahl der durchgeführten Überprüfungen zu dem jeweiligen Test.

Die Ergebnisse der Testphase zeigen, dass insbesondere die strukturell anspruchsvollen Formen wie der Triple-Jump-Test, der Lauf- und der Einbeinsprung-Test in dem frühen Nachbehandlungszeitraum von nur bis zu vier Probanden durchgeführt werden konnten. Bei einer Aufnahme aller Tests würden sich entsprechend schnell Bodeneffekte zeigen (Tab. 13). Tabelle 14 zeigt die deskriptiven Ergebnisse.

### **3.1.2 Kritische Bewertung und Konstruktion des Polla**

Die wichtigste Erkenntnis besteht in einer mangelnden Selbsteinschätzungsfähigkeit und übertriebenen Motivation zur Ergebnisverbesserung, die mit einer raschen Überlastungsgefahr durch die Ausführung der Tests unter Zeitdruck einhergeht. Die Grenze der ohnehin gestörten und zur Reizung neigenden Körperstruktur ist mit einer Testdauer von 30 Minuten schnell erreicht. Unter der Betrachtung der Empfehlungen von Bührlen et al. (2002; Kap. 2.6.9) stellt sich daher die Frage nach dem Sinn ausschließlich quantitativer Testung. Welche Informationen sollen beispielsweise aus dem fünfmaligen *Aufheben eines Gegenstandes* oder dem zweimaligen *sich Hinlegen auf den Boden* auf Zeit über die tatsächliche alltägliche Belastungsfähigkeit gewonnen werden? Wichtig erscheint es auch Antworten auf Fragen zu erhalten, wie: Mit welcher Art von Einschränkung und in welcher Qualität wird eine Aktivität des Alltags bewältigt? Hieraus sind insbesondere für den diagnostischen Prozess und die Behandlung wichtige Informationen zu gewinnen.

Aus genannten Gründen wird den Empfehlungen nach Bührlen et al. (2002) eine Kombination aus qualitativer Beurteilung und quantitativer Messung vorgezogen. Therapeuten sollen sich zunächst in zehn Tests qualitativ einen Eindruck über das Aktivitätsniveau der Patienten verschaffen und anschließend weitere vier Tests zuerst qualitativ beurteilen, um sie dann quantitativ zu messen (ANHANG 8). Die Testzeit sollte zur Verringerung der Verletzungsgefahr insgesamt 20 Minuten nicht überschreiten.

Alle Tests der Pilotphase werden in den endgültigen Polla aufgenommen. Lediglich der Triple-Jump-Test wird für die Validierung entfernt, in den Empfehlungen zum Praxiseinsatz und zur Vermeidung von Deckeneffekten bei Jüngeren Sportlern aber wieder aufgenommen (ANHANG 8).

Die Tests im Polla werden allerdings zur Reduktion des Verletzungsrisikos und zu einem methodisch angelegten Einschleichen in größere Belastungen leicht modifiziert. Der *Chair-rising-Test* (10x auf Zeit) und der *Up & Go-Test* werden in einem *1-Minute-Up & Go-Test* kombiniert; ein qualitativ zu beurteilendes *dreimaliges Aufstehen von einem Hocker* wird vorgeschaltet. Die Testanweisung zum Timed-Up & Go-Test nach Podsiadlo und Richardson (1991) wird durch eine leistungs-

orientierte Durchführungsanweisung über eine Minute modifiziert, um einem altersbedingt breiten Patientenspektrum in der Untersuchung entgegen zu kommen. Ursprünglich war der „Get-up and Go“ Test nach Mathias, Nayak und Isaacs (1986) für Personen mit einem Altersdurchschnitt von 79,5 Jahren konzipiert.

Obwohl der Lauftest, der Einbeinsprung-Test und der Treppenlaufen-Test in der Pilotphase nur von wenigen Patienten durchgeführt werden konnten, erscheint eine Aufnahme dennoch wichtig, um in einer späteren Therapiephase Deckeneffekte zu vermeiden. Zur Einhaltung des Zeitrahmens von höchstens 20 Minuten wird der „6-Minuten-Gehtest“ nach Bührlen et al. (2002) auf zwei Minuten reduziert. Dem 2-Minuten-Lauftest auf dem Laufband – er konnte in der Pilotphase von nur wenigen durchgeführt werden und ein Laufband ist nicht in jeder Praxis vorhanden – wird eine qualitative Beurteilung des *Laufens auf der Stelle* als methodische Einleitung vorgeschaltet. Dem quantitativ gemessenen Treppenlaufen-Test wird das qualitativ zu beurteilende *Aufsteigen auf eine Stufe* zugefügt, denn ein Treppenstockwerk ist ebenfalls nicht in jeder Praxis vorhanden. Die Endform des Polla enthält demnach 14 Tests.

Vor der Darstellung der Erhebungsinstrumente im Einzelnen wird zunächst das Studiendesign erläutert.

## **3.2 Hauptstudie**

### **3.2.1 Studiendesign**

In einer Längsschnittstudie zu den Zeitpunkten Therapiebeginn und Therapieende wurden die Ergebnisse verschiedener Messungen zu physiotherapeutischen Behandlungen an einem Kollektiv von 81 Patienten erfasst (Abb. 4). Die abhängigen Variablen des Polla, der Selbsteinschätzung zu Aktivitäten sowie der gesundheitsbezogenen Lebensqualität (SF-36) wurden zu der zweifach gestuften unabhängigen Variablen „Messzeitpunkt“ erhoben. Eine Gruppe von 30 Patienten wurde zusätzlich mit einer Vergleichsgruppe von 30 Gesunden parallelisiert und untersucht.

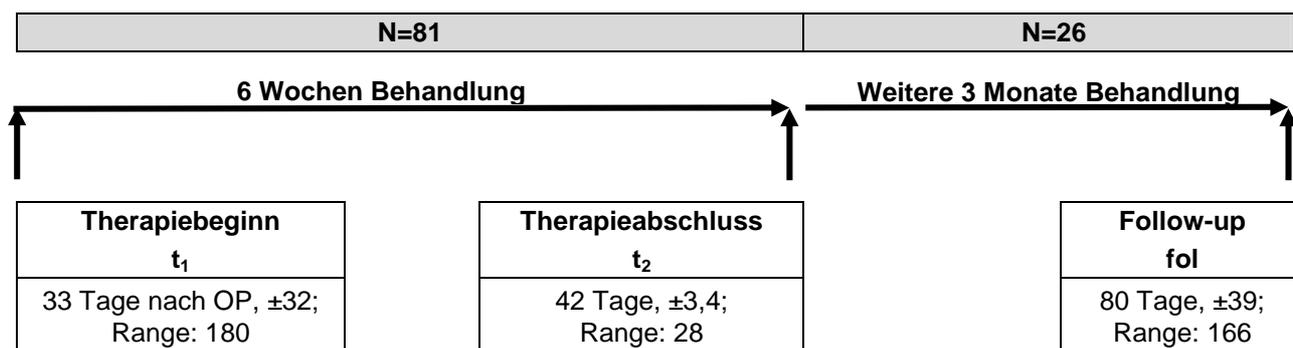


Abb. 4. Messzeitpunkte in der Längsschnittstudie.

26 Patienten konnten darüber hinaus zum einem Zeitpunkt *Follow-up* noch einmal vergleichend getestet werden. (Abb. 4).

### 3.2.2 Patientenauswahl und Parallelisierung

Zur Eingrenzung der Altersstreuung wurde die Experimentalgruppe zunächst auf ein Alter von 19 bis 65 Jahren festgelegt. Eine Erweiterung des Einsschlusskriteriums auf 75 Jahre erfolgte gegen Mitte der Untersuchung aufgrund unerwartet niedriger Fallzahlen unter 65 Jahren. Die Patientenaufnahme beschränkte sich auf Krankheitsbilder nach operativem Eingriff an der unteren Extremität und betraf Hüft-, Knie-, Unterschenkel- und Fußpatienten.

Ausschlüsse betrafen Patienten mit einem Belastungsverbot oder einer Teilbelastung von über zwei Wochen nach Therapiebeginn und Personen, die sprachliche Verständnisschwierigkeiten bzw. Einschränkungen hinsichtlich der Compliance aufwiesen. Zu letzterem sind Personen zu nennen, die kein Interesse an einer Untersuchungsteilnahme äußern würden, was allerdings in keinem Fall auftrat.

Der Abgleich der Ergebnisse von Patienten mit den von Gesunden erfolgte gegen Ende der Untersuchungsdurchführung im Rahmen der Parallelisierung von 30 Patienten und 30 Gesunden. Dazu wurden in Frage kommende gesunde Personen, vornehmlich Angestellte des *Ambulanten Rehasentzentrums Kiel* oder des *Lubinus Clinicums* sowie Personen, die das private Gesundheitsangebot der Einrichtung nutzten, hinsichtlich ihrer Bereitschaft zur Teilnahme an der Untersuchung befragt. Die Auswahl der Vergleichspersonen orientierte sich an den jeweiligen Vorgaben zu Alter, Geschlecht und zum Body-Mass-Index (BMI) der Patienten aus der Untersuchungsgruppe. Für die Vergleichsgruppe kamen lediglich Personen in Betracht, die keine Schädigungen an der unteren Extremität aufwiesen.

### 3.3. Erhebungsinstrumente

#### 3.3.1 Der Performance-Test (Polla)

Der Polla stellt das Kernstück der Untersuchung und der objektiven Testung in Form standardisierter Fremdbeobachtung dar (ANHANG 1). Für die weitere Abhandlung wird an dieser Stelle aus Gründen der unterschiedlichen Messniveaus eine Abgrenzung in zwei Testkonstrukte vorgenommen:

1. Performance-Test of lower limb activities (Polla) und
2. die gemessenen Tests im Polla.

Der *Polla* umfasst 14 Einzeltests (Tab. 16), deren Beurteilung zur Absicherung und Vermeidung von Überlastungssituationen zunächst qualitativ erfolgt. Beispielsweise soll der Proband dreimal ohne Zuhilfenahme der Hände von einem Hocker aufstehen und sich wieder hinsetzen. Es erfolgt eine dreistufig rangskalierte Messwertaufnahme (Tab. 15), wobei einem hohen Wert eine hohe Funktionskapazität entspricht.

Tab. 15. Die Beurteilungskategorien zu den Polla Items.

Antwortmöglichkeiten	Item Werte
Nein, die Aktivität ist nicht möglich	0
Die Aktivität ist eingeschränkt möglich	1
Die Aktivität ist uneingeschränkt möglich	2

Die erreichbare Punktzahl für jeden einzelnen Test beträgt zwei und summiert sich bei 14 Tests auf maximal 28 Punkte. Einen Überblick über die Tests gibt Tabelle 16. Die Qualität der Aktivitätsausführung wurde mit „eingeschränkt“ belegt, sofern nach einer standardisierten Beurteilung (ANHANG 8)

1. Schmerz, Druckgefühl oder Ziehen,
2. leichte oder starke Ausweichbewegungen oder
3. Hilfen (durch eine Person) oder Hilfsmitteln registriert wurden.

Die folgenden vier Tests im Polla wurden zusätzlich quantitativ erhoben (gemessene Tests im Polla), der 1-Minute-Up & Go-Test nach Podsiadlo und Richardson (1991), ein 2-Minuten-Gehtest, den Treppenlaufen-Test (Risberg & Ekeland, 1994) und einen 2-Minuten-Lauftest. Für die Testdurchführung waren etwa 15-20 Minuten veranschlagt.

Tab. 16. Die Aktivitäten im Polla (1-14) inklusive der gemessenen Tests (11-14).

	Testname	Erläuterungen zu den Tests	Anzahl
1.	<b>Strümpfe aus- und wieder anziehen</b>	Sitz auf einem Hocker von 45 cm Höhe.	1x
2.	<b>Aufstehen von einem Hocker</b>	Vom Sitz auf einem Hocker von 45 cm Höhe soll das Aufstehen aus dem Parallelstand der Füße und ohne Unterstützung der Hände erfolgen.	3x
3.	<b>Einbeinstand</b>	Der Einbeinstand muss ohne Bodenkontakt des Spielbeines und ohne Berührung des Standbeines in leichter Kniebeugstellung durchgeführt werden.	10 Sek.
4.	<b>Aufheben eines leichten Gegenstands vom Fußboden</b>	Der Gegenstand ist ein Verkehrshut von 10 cm Höhe. Die Art der Ausführung hinsichtlich rücken-gerechten Verhaltens obliegt dem Patienten.	3x
5.	<b>Hinlegen auf den Fußboden und wieder Aufstehen</b>	Die Unterlage bildet eine weiche Therapiematte. Ausgangs- und Endposition ist der Parallelstand.	1x
6.	<b>Treppaufsteigen einbeinig auf einen Stepp</b>	Der Stepp hat eine Höhe von 20 cm. Nach dem Aufsteigen erfolgt das langsame Absteigen über dasselbe Bein nach vorne. Jeweils mit dem betroffenen und dem gesunden Bein.	3x
7.	<b>In die Hocke gehen</b>	Die Hocke soll soweit als möglich beidbeinig parallel erfolgen.	2x
8.	<b>Hinknien auf den Fußboden</b>	Auf einer Therapiematte kniend soll 20x auf der Stelle „getippelt“ werden.	1x
9.	<b>Laufen</b>	Auf der Stelle.	10 Sek.
10.	<b>Einbeinhüpfen auf der Stelle</b>	Der Fuß muss sich von der Unterlage lösen	10x
<b>Zusätzlich zur qualitativen Beurteilung gemessene Tests zum Polla</b>			
11.	<b>1-Minute-Up &amp; Go-Test</b>	Gezählt wird die Anzahl der Wegstrecken innerhalb einer Minute, die ein Patient aus dem Sitz von einem Hocker bis zu einer drei Meter entfernten Wand und wieder zurück bis zum Hocker zurücklegt. Jedes Hinsetzen wird als ganzer oder, wenn die Zeit an der Wand zu Ende ist, als halber Punkt gezählt.	<b>Anzahl</b>
12.	<b>2-Minuten-Gehtest</b>	Gemessen wird die Gehstrecke in Metern innerhalb von zwei Minuten, die zwischen zwei 14 Meter voneinander entfernten Markierungen zurückgelegt wird.	<b>Meter</b>
13.	<b>Treppenlaufen-Test</b>	Gemessen wird die Zeit in Sekunden, die für die Bewältigung zweier Treppenabsätze hinauf und wieder herunter benötigt wird (40 Stufen à 18cm).	<b>Sek.</b>
14.	<b>2-Minuten-Lauftest</b>	Gemessen wird die Anzahl der gelaufenen Meter auf dem Laufband innerhalb von zwei Minuten.	<b>Meter</b>

### 3.3.2 Die Gesundheitsbezogene Lebensqualität (SF-36)

Zur Überprüfung des allgemeinen Gesundheitszustands wurde entsprechend den Empfehlungen von Muthny et al. (1999) der SF-36 mit seinen 36 Items in acht Dimensionen zur gesundheitsbezogenen Lebensqualität ausgewählt (ANHANG 2).

Der Fragebogen wurde zu zwei Zeitpunkten im Zeitfenster von einer Woche eingesetzt und nach der Behandlung in der Einrichtung vom Patienten ausgefüllt. Die Zeit für das Ausfüllen war auf etwa 10 Minuten angesetzt.

### 3.3.3 Die Selbsteinschätzung von Aktivitäten

Der in der Untersuchung eingesetzte Fragebogen zur Selbsteinschätzung geht auf die Erfahrungen eines in der klinischen Praxis entwickelten spezifischen Instrumentes zurück und operationalisiert die Abfrage von Einschränkungen zu Aktivitäten an der unteren Extremität (ANHANG 4). Die 27 Aktivitäten sind fünfstufig intervallskaliert. Der niedrigste Wert entspricht „Ich kann diese Aktivität nicht“ und der höchste „Ich habe keine Probleme“ (Tab. 17).

Tab. 17. Die Beurteilungskategorien zu den Items der Selbsteinschätzung zu Aktivitäten.

Antwortmöglichkeiten	Item Werte
Ich kann diese Aktivität nicht	0
Ich habe große Probleme	1
Ich habe mäßige Probleme	2
Ich habe leichte Probleme	3
Ich habe keine Probleme	4

Die schriftliche Einzelbefragung erfolgte im Zeitfenster von einer Woche, schloss sich an eine Behandlung an und erforderte einen Zeitaufwand von etwa fünf Minuten.

### 3.3.4 Befundaufnahme und spezifische Zielkriterien

Die *Befundaufnahme* beinhaltete neben Angaben zum Alter, Geschlecht, Body-Mass-Index (BMI) auch Angaben zur Diagnose und Anamnese sowie zu den Verordnungsarten und der Behandlungsfrequenz. Die *Abschlussdokumentation* enthielt zusätzlich Angaben über die Anzahl der Verordnungen und die Behandlungsanzahl sowie die Behandlungsfrequenz (ANHANG 5).

Tab. 18. Die Erfassung der spezifischen Zielkriterien in ihren Merkmalsausprägungen.

Zielkriterium	Skalenniveau	Stufen	Merkmalsausprägung
<b>Medikation</b>	rangskaliert	1-3	1 = regelmäßig 2 = bei Bedarf 3 = nie
<b>Hauptproblem</b>	rangskaliert	1-8	1 = ständig 2 = 2x im Jahr 3 = 2x im Monat 4 = 1x die Woche 5 = mehrmals die Woche 6 = täglich 7 = mehrmals täglich 8 = nie
<b>NRS</b>	intervallskaliert	0-10	0 = kein Schmerz 10 = maximal vorstellbarer Schmerz
<b>Schmerzdauer</b>	rangskaliert	1-6	1 = ständig 2 = mehrere Stunden täglich 3 = bis zu einer Stunde 4 = einige Minuten 5 = kurze Schmerzen 6 = keine Schmerzen
<b>Giving-Way</b>	rangskaliert	5	1 = mehrmals täglich 2 = täglich 3 = mehrmals die Woche 4 = weniger als 1x die Woche 5 = kein
<b>Aktive Flexion</b>	intervallskaliert	Grad	Messung in Grad
<b>Kniebeuge</b>	intervallskaliert	Grad	Messung in Grad

Tabelle 18 gibt einen Überblick über die *spezifischen Zielkriterien*, von denen im Untersuchungsverlauf Veränderungen zu erwarten waren. Eine mündliche standardisierte Befragung umfasste die Häufigkeit der Medikamenteneinnahme, des Hauptproblems und des Giving-Way Phänomens – letzteres beschreibt das spontane Wegsacken im Kniegelenk (Buckup, 2000). Intervallskaliert erhoben wurde die Schmerzintensität nach der Numeric-Rating-Scale (NRS, Kool & de Bie, 2001), die aktive Flexion (Beugung) am betroffenen Gelenk nach der Neutral-Null-Methode in Rückenlage (Debrunner, 1971) und die schmerzfreie Kniebeuge in Grad am Kniegelenk gemessen (Tab. 18, ANHANG 5).

## 4 DURCHFÜHRUNG DER STUDIE

### 4.1 Patientenrekrutierung

Alle Patienten, die entsprechende Einschlusskriterien aufwiesen, wurden zunächst von den Versuchsleitern um die Teilnahme an der Untersuchung gebeten. Vor Untersuchungsbeginn fand eine Aufklärung sowohl in mündlicher als auch schriftlicher Form über das Ziel und den Ablauf der Untersuchung statt. Die Bereitschaft zur Teilnahme wurde durch eine schriftliche Einverständniserklärung bestätigt (ANHANG 6).

### 4.2 Erhebungsplan

Zur Erfassung von Veränderungen in der postoperativen Phase ergab sich aus klinischer Erfahrung eine Untersuchungsdauer von sechs Behandlungswochen und 10 Behandlungseinheiten.

Tab. 19. Der Erhebungsplan im zeitlichen Verlauf der Untersuchung.

Zielkriterien	t <sub>1</sub>	t <sub>2</sub>	Intrarater	Follow-up	Gesunde
	N=81 1. Tag	N=81 42 Tage, ±3,4	N=29 2,3 Tage, ±0,8	N=26 80 Tage, ±39	N=30
<b>Befundaufnahme</b>	X				
<b>Anamnese</b>	X				
<b>Medikation</b>	X	X			
<b>Hauptproblem</b>	X	X			
<b>NRS</b>	X	X			
<b>Schmerzdauer</b>	X	X			
<b>Giving-Way</b>	X	X			
<b>Beweglichkeit</b>	X	X			
<b>Kniebeuge</b>	X	X			
<b>Polla</b>	X	X	X	X	X
<b>SF-36</b>	X	X		X	X
<b>Selbsteinschätzung</b>	X	X		X	X
<b>Abschlussbefund</b>		X			
<b>Interrater Polla (N=32)</b>		X			

Die Therapeuten erstellten während der ersten 30-minütigen Behandlungseinheit die Befundaufnahme und nach sechs Wochen den Abschlussbefund. Im Aufnahmebefund wurden neben den demografischen Daten die Diagnose, die Anamnese

sowie die spezifischen Zielkriterien erfasst. Im Anschluss erfolgte die Überprüfung des Polla. Abschließend füllten Patienten die Fragebögen zur Selbsteinschätzung von Aktivitäten und zum SF-36 aus. Die gleichen Parameter (Ausnahme: Anamnese, Demografie) waren Gegenstand der Erhebungen zum Therapieabschluss nach 42 ( $\pm 3,4$ ) Tagen (Tab. 19).

Die *Selbstbeurteilungsinstrumente* wurden zu allen drei Messzeitpunkten im Anschluss an die jeweilige Behandlung ausgehändigt, von den Patienten in der Einrichtung ausgefüllt und dem zuständigen Versuchsleiter oder an der Rezeption übergeben.

### 4.3 Durchführung zur Interrater- und Intrarater Reliabilität

32 der 81 Abschlusstests zum Polla wurden gegen Behandlungsende nach sechs Wochen als *Interrater-Überprüfung* durchgeführt: Zeitgleich beurteilten zwei weitere Therapeuten – unabhängig vom behandelnden Stammtherapeuten – die Ausführungen der Aktivitäten im Polla („ja-eingeschränkt-nein“) und trugen die Ergebnisse auf jeweils einem separaten Dokumentationsbogen ein (ANHANG 1).

Bei der *Intrarater-Reliabilität* überprüft ein und derselbe Beurteiler zu zwei Zeitpunkten die drei gemessenen Tests an einem Patienten. Entsprechende Empfehlungen der Literatur legen einen 3-14 Tage Abstand nach dem ersten Test nahe (Bös, 2001). Aufgrund schneller Veränderungen, die in einem frühen postoperativen Zeitraum zu Therapiebeginn zu erwarten sind, erschien die Erhebung der Retests im Rahmen dieser Untersuchung zu einem Zeitpunkt sinnvoll, wo insgesamt stabilere Ergebnisse zu erwarten waren.

Entsprechend wurden 29 von 81 Patienten im Durchschnitt 2,3 Tage ( $\pm 0,8$ ) nach Untersuchungsende zu einem Retest einbestellt und erneut mit den Tests überprüft.

### 4.4 Kontrolle von Störgrößen und Qualitätssicherung

*Strukturelle Störungen* auf die flexibel im Hinblick auf die Terminplanung reagiert werden musste, ergaben sich durch unvorhergesehene Wechsel, Krankheit oder kurzfristigen Urlaub auf Therapeutenseite und Nichterscheinen, Absage von Terminen oder Therapieabbruch auf Patientenseite.

Eine *Maßnahme der Qualitätssicherung* umfasste die eintägige Einweisung der vier Versuchsleiter unter besonderer Berücksichtigung der standardisierten Testanwendung. Zur Sicherung der korrekten Durchführung der Tests lagen sowohl ein Standardisierungsleitfaden der Durchführungssituation als auch eine schriftliche Anleitung bereit (ANHANG 8).

#### **4.5 Behandlungsinhalte und Verantwortlichkeiten**

Die Durchführung der sechswöchigen physiotherapeutischen Maßnahmen im Untersuchungszeitraum erfolgte unter Routinebehandlung gemäß den ausgestellten Verordnungen entsprechend den in der Krankengymnastik allgemein üblichen Behandlungsinhalten nach operativen Eingriffen an der unteren Extremität. Dazu zählen die Gangschulung, das funktionelle Beinachsentraining und die Schulung funktioneller Aktivitäten im Alltag. Spezifische Behandlungstechniken beinhalteten beispielsweise die Gelenkmobilisation und Muskeldehnungen entsprechend dem Konzept der Manuellen Therapie sowie zur Verbesserung der muskulären Koordination die propriozeptive neuromuskuläre Faszilitation (Reichel, 2005). Ferner kamen Muskeldekontraktionen in Anlehnung an die Brügger-Therapie (Brügger, 2000) und die gerätegestützte medizinische Trainingstherapie zur Anwendung. Ein regelmäßiges Üben der Testsituationen des Polla im Behandlungsalltag war nicht vorgesehen, erfolgte aber in Ausnahmefällen nach individueller Bedürfnislage im Therapieprozess.

Die funktionellen Gerätschaften umfassten die für die untere Extremität bedeutsamen medizinischen Muskeltrainingsgeräte wie Funktionsstemma, Kniebeuger und -strecker sowie Fahrradergometer in aufrechter und halb liegender Position, ein Cross-Trainer und ein Posturomed. Art und Umfang des Einsatzes von Geräten waren durch die Ordnungsweise determiniert. Die für die Untersuchung relevanten Ordnungsweisen beinhalteten

- die Physiotherapie (20-30 Min.),
- die ambulante Rehabilitation (2,5-3 Stunden),
- die Krankengymnastik am Gerät (KGG, eine Stunde)
- sowie eine standardisierte Heilmittelkombination (D1, eine Stunde).

Die Verantwortung für die Durchführung des Projektes lag bei dem Projektleiter und für die Betreuung, Behandlung sowie Testung der Patienten in den Händen der jeweiligen Therapeuten (Projektmitarbeiter).

#### 4.6 Merkmale der Patientenstichprobe und Drop-out

An der Studie nahmen insgesamt 81 Patienten teil. Mit Ausnahme von zwei Personen waren alle Patienten deutscher Staatsangehörigkeit und lebten zu 90 Prozent in einem Ein- oder Zwei-Personen-Haushalt. Über die Hälfte der Teilnehmer war zum Zeitpunkt der Erhebung verheiratet und zwei Drittel lebte mit einem festen Partner zusammen. Von den 81 Patienten war zurzeit der Untersuchung ein Drittel berentet.

Das Alter der Teilnehmer betrug im Durchschnitt 53 Jahre ( $\pm 17,7$ ) in einem Bereich von 19 bis 76 Jahren (Tab. 20). Das Geschlechtsverhältnis ist annähernd ausgeglichen (Frauen 48,1%, Männer 51,9%)

Tab. 20. Verteilung von Alter, Geschlecht und Body-Mass-Index (BMI) im Patientenkollektiv.

N=81	N <sup>3</sup>	Geschlecht	Alter		BMI	
			M <sup>1</sup>	s <sup>2</sup>	M	s
Patienten	20	50% weiblich (10) 50% männlich (10)	56,05	17,24	26,99	3,51
Patienten (Intrarater)	29	48,3% weiblich (14) 51,7% männlich (15)	48,97	20,65	27,39	5,19
Patienten (Interrater)	32	46,9% weiblich (15) 53,1% männlich (17)	55,19	14,69	28,71	5,36
Patienten Gesamt	81	48,1% weiblich (42) 51,9% männlich (39)	53,17	17,71	27,82	4,90

Anm.: <sup>1</sup>Mittelwert, <sup>2</sup>Standardabweichung, <sup>3</sup>Anzahl der Versuchspersonen.

Die Altersverteilung zeigt eine zweigipflige Kurve zwischen 19-45 und 60-76 Jahren. Die Verteilung weicht nach dem Kolmogorov-Smirnov-Anpassungstest signifikant von der Normalverteilung ab ( $Z = 1,82$ ;  $p = ,003$ ). Der Body-Mass-Index (BMI) dagegen ist mit  $p = 0,44$  ( $Z = 0,87$ ) normal verteilt.

Drop-out traten in sechs Fällen auf (Tab. 21). Gründe waren unentschuldigtes Fernbleiben von der Therapie, postoperativ bedingte Komplikationen und Therapiefortsetzung in einer wohnortnäheren Einrichtung ab. Postoperative Schwierigkeiten waren nicht auf die aktive Therapie, sondern beispielsweise auf unspezifisch starke Schmerzen oder auf eine Gelenkersatzlockerung zurückzuführen.

Tab. 21. Drop-out nach Geschlecht, Alter, Behandlungsanzahl und Ursachen des Abbruchs.

	Geschlecht	Alter	Behandlungen	Grund des Abbruchs
1.	Weiblich	20	4	unentschuldigtes Fernbleiben
2.	Männlich	21	2	unentschuldigtes Fernbleiben
3.	Männlich	75	2	Therapiefortsetzung an einem anderen Ort
4.	Männlich	67	1	operativ bedingte Komplikationen
5.	Weiblich	46	14	operativ bedingte Komplikationen
6.	Weiblich	63	12	operativ bedingte Komplikationen

## 4.7 Merkmale der parallelisierten Stichproben

In der Vergleichsgruppe der 30 Gesunden sind alle, mit Ausnahme einer Person, deutscher Staatsangehörigkeit. 60 Prozent leben zum Zeitpunkt der Erhebung in einem Ein- oder Zwei-Personen-Haushalt und 70% der Teilnehmer waren verheiratet. Zweidrittel der 30 Gesunden leben mit einem festen Partner zusammen, ein Drittel ist berentet.

Tab. 22. Gruppenstatistiken: Verteilung von Alter, Geschlecht und BMI in den parallelisierten Gruppen „Patient“ und „Gesunde“ (n=30).

Parallelisierte Gruppen	N <sup>3</sup>	Geschlecht		Alter		BMI	
				M <sup>1</sup>	s <sup>2</sup>	M	s
Gesunde	30	40%	weiblich (12)	49,67	16,78	24,82	2,53
		60%	männlich (18)				
Patienten	30	33,3%	weiblich (10)	48,07	18,58	25,32	2,63
		66,7%	männlich (20)				

Anm.: <sup>1</sup>Mittelwert, <sup>2</sup>Standardabweichung, <sup>3</sup>Anzahl der Versuchspersonen.

Die in die *Parallelisierung* aufgenommenen je 30 Patienten und 30 gesunden Personen zeigen Normalverteilung hinsichtlich des BMI und des Alters. Nach dem T-Test für unabhängige Stichproben weisen die Gruppen weder signifikante Unterschiede im BMI ( $T = 5,11$ ;  $p = ,66$ ) noch bezüglich des Alters auf ( $T = -0,35$ ;  $p = ,73$ ).

Dagegen zeigen sich Unterschiede zwischen der Gruppe von Gesunden (n=30) und der Gesamtstichprobe (n=81) im Hinblick auf den BMI mit ( $T = 4,12$ ;  $p < ,001$ ), nicht dagegen zum Alter ( $T = 0,94$ ;  $p \leq ,35$ ; Tab. 23, deskriptive Statistik Tab. 24).

Tab. 23. Test bei unabhängigen Stichproben zu Unterschieden in Alter und BMI zwischen den parallelisierten Gruppen „Patient“ und „Gesunde“ (N=30).

		Levene-Test		T-Test für die Mittelwertgleichheit					
		F	Sig.	T	df	Sig.	Mittlere Diff.	95% KI	
								Untere	Obere
<b>Alter</b>	Gleiche Varianzen	,072	,789	,94	109	,350	3,51	-3,89	10,90
	ungleiche Varianzen			,96	54,51	,340	3,51	-3,79	10,80
<b>BMI</b>	Gleiche Varianzen	8,36	,005	3,13	109	,002	2,95	1,07	4,80
	ungleiche Varianzen			4,12	97,90	,000	2,95	1,52	4,35

Tab. 24. Deskriptive Statistik zu Alter und BMI in den parallelisierten Gruppen „Patient“ und „Gesunde“.

	Gruppen	N	M	s
<b>Alter</b>	Patient	81	53,17	17,71
	Gesunde	30	49,67	16,78
<b>Body-Mass-Index</b>	Patient	81	27,86	4,90
	Gesunde	30	24,88	2,51

Die 30 Patienten in der Parallelisierung weisen neun Hüftendoprothesen, fünf Knieendoprothesen sowie 16 Knie/Unterschenkeldiagnosen auf. Der Anteil der männlichen Teilnehmer überwiegt mit 60 Prozent den der weiblichen Teilnehmer (40%).

#### 4.8 Diagnosen und Behandlungsstatistik

Die in der Untersuchung aufgenommenen Diagnosen sind nach Körperstrukturen in vier Gruppen zusammengefasst worden. Dies geschah zum einen zur erleichterten Darstellung und zum anderen im Hinblick auf eine spätere Analyse von Gruppenunterschieden zu Diagnosen bezüglich der Testergebnisse des Polla.

Mit 29 und 28 Patienten bilden die Gruppen der *Hüftendoprothesen* und *Kniediagnosen* die größten Gruppen, gefolgt von den Gruppe der *Knieendoprothesen* (n=15) und der *kombinierten Knie-/Unterschenkeldiagnosen* (n=9), dessen Entwicklung im Funktionsstatus erfahrungsgemäß schwieriger zu bewerten ist (Tab. 25). Zu letzterer Gruppe gehören z.B. Diagnosen wie die Versetzung der Tuberositas-Tibae oder Tibiakopffrakturen (Tab. 25).

Am häufigsten verordnet wurde mit 52 Prozent die Krankengymnastik. Hinzu kamen kombinierte Verordnungen in Form von Erweiterter Ambulanter Physiotherapie (EAP) oder die standardisierte Heilmittelkombination (D1).

Tab. 25. Die Häufigkeiten der aufgenommenen Diagnosen (N=81).

Gruppe	Diagnosen	Insgesamt
<b>Hüfte</b>	Endoprothesen	29
<b>Knie</b>	Endoprothesen	15
<b>Knie</b>	Kreuzbandplastiken, Arthroskopien, Lateral-Release, Mediale Raffung	28
<b>Knie/Unterschenkel</b>	Unterschenkelfraktur; Calcaneusfraktur; obere Sprunggelenksfraktur; Tibiakopffraktur (2x); Tuberositas-Tibae-Versetzung; Tibiakopf-Umstellungsosteotomie; Kreuzbandersatzplastik mit den Zusatzdiagnosen: Meniskusresektionen, Innenbandanriss und partielle Patellasehnenruptur; knöcherner vorderer Kreuzbandausriss	9

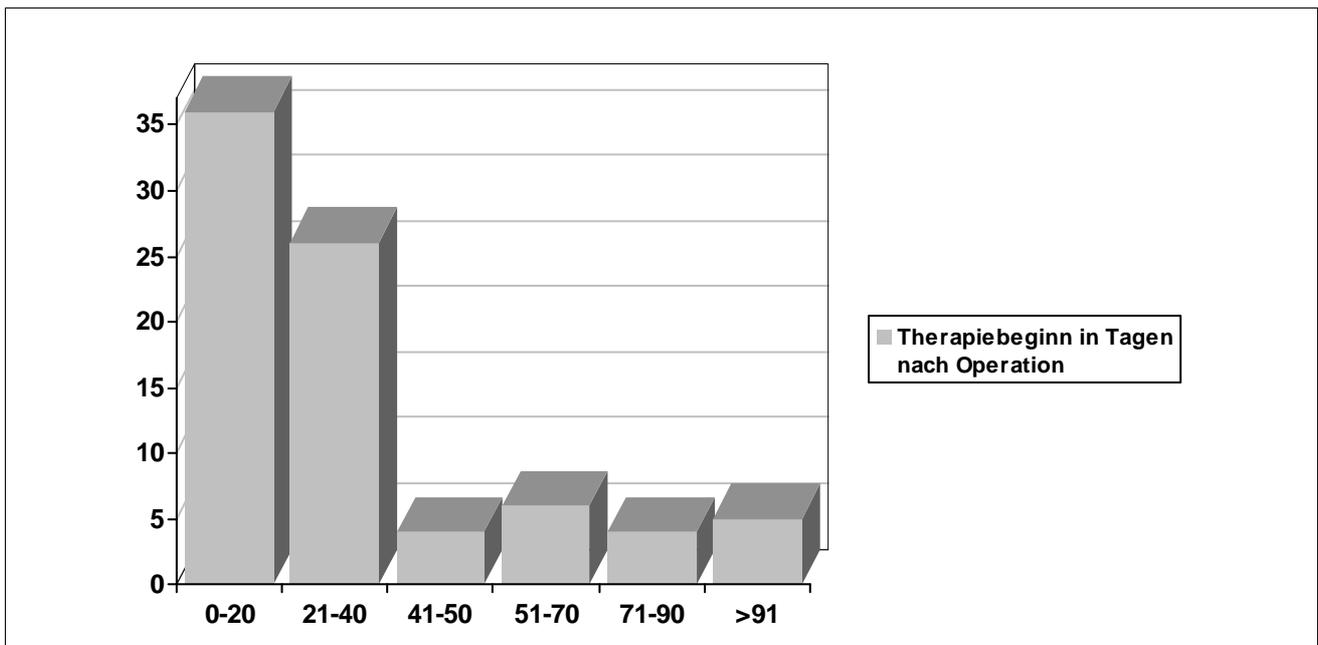


Abb. 5. Postoperativer Therapiebeginn in Tagen. Waagrecht: Fallzahlen; Vertikal: Tage nach Operation.

93 Prozent der Patienten erhielten 6-18 Behandlungseinheiten. Die Behandlungsfrequenz lag zwischen ein und drei Behandlungen pro Woche. Der durchschnittliche Untersuchungszeitraum betrug 42 Tage ( $\pm 3,4$ ) bei einem Maximum 63 und einem Minimum von 35 Tagen.

Die Therapie wurde durchschnittlich 33 Tage nach Operation begonnen ( $\pm 32$ , Range: 180). 36 Patienten (44%) begannen die Therapie innerhalb von 20 Tagen nach dem Operationstag und 26 Patienten (32%) in einem Zeitraum bis zu 40 Tagen danach (Abb. 5).

Die Gründe des späten Therapiebeginns bei einigen Patienten wurden nicht systematisch erfasst. Neben dem Jahreswechsel, der häufig aus privaten Gründen zu einer Verzögerung der Therapieaufnahme führte, wurde die aktive Therapie nicht selten erst mit Beginn der Belastungsaufnahme verordnet, die sich bei langwierigen Krankheitsbildern hinauszögerte.

#### 4.9 Statistische Auswertestrategie

Die Auswertungen zu den Ergebnissen der Untersuchung beschäftigen sich zunächst mit einer *Itemanalyse* zu den Items im Polla und der Selbsteinschätzung und ziehen Konsequenzen im Hinblick auf den Verbleib einzelner Items in der

Testbatterie. Eine anschließende *Faktorenanalyse* klärt vor der Bildung von Summenskalen mögliche inhaltliche Dimensionen beider Instrumente.

Die folgende *Ergebnisdarstellung* gibt Auskunft über die fehlenden Werte, Verteilungseigenschaften und die deskriptive Statistik zu den Messinstrumenten zu den Zeitpunkten Therapiebeginn und Therapieabschluss.

In der sich anschließenden *Inferenzstatistik* werden, neben den über die Messinstrumente erhobenen Veränderungen der Therapieergebnisse gegenüber dem Untersuchungsbeginn, auch Gruppenunterschiede in den Testergebnissen im Hinblick auf Diagnosegruppen und Altersklassen betrachtet.

Die *Validierung von Testgütekriterien* folgt den Empfehlungen von Lienert & Raatz (1998) mit der Bestimmung von Objektivität, Reliabilität und Validität. Die Objektivitätsprüfung zur Interrater Übereinstimmung der ordinalskalierten Polla Items erfolgt über das Kappa-Maß. Eine sich anschließende Intrarater-Reliabilitätsprüfung zu den gemessenen Tests wird über den Intra-Class-Koeffizienten (ICC) geklärt und eine Reliabilitätsprüfung (interne Konsistenz) erfolgt über das statistische Maß Cronbachs Alpha. Die Prüfung der Konstruktvalidität betrachtet über den Korrelationskoeffizienten Spearman-Rho die Zusammenhänge zwischen den qualitativ ordinalskalierten Polla Items und den gemessenen Performance-Tests sowie die Korrelation ihrer Summenskalen untereinander.

Die kriterienbezogene Validität als *eine zentrale statistische Hypothese* klärt die Zusammenhänge zwischen den Summenskalen des Polla, der SF-36-Dimensionen und der Selbsteinschätzung. In einer *zweiten statistischen Hypothese* wird die Änderungssensitivität der Instrumente über den Standardized Response Mean (SRM) verglichen. Schließlich wird die Unterscheidungsfähigkeit der Tests zwischen Patienten und Gesunden über die einfache Varianzanalyse nach dem linearen Modell überprüft.

In dem Exkurs zum *Follow-up* mit 26 Versuchspersonen werden unter dem Blickwinkel des größeren Zeitabstands vom operativen Eingriff die Unterscheidungsfähigkeit der Testergebnisse zwischen Gesunden und Patienten sowie die Änderungssensitivität betrachtet. Die Auswertung erfolgte mit dem Programm SPSS Version 12.

Alle Ergebnisse werden unter dem Gliederungspunkt 7 (*Diskussion*) einer kritischen Betrachtung unterzogen. In Kapitel 8 erfolgt eine abschließende Betrachtung der klinischen Relevanz und es werden Empfehlungen für den Praxiseinsatz eines modifizierten Polla gegeben und ausblickend weitere Untersuchungsfragen aufgeworfen. Kapitel 9 gibt eine Zusammenfassung.

## 5 AUSWERTUNGEN

### 5.1 Entscheidungsgrundlage für die Skalenbildung

#### 5.1.1 Itemstatistik zu Polla und Selbsteinschätzung

Die Ergebnisse der Itemanalyse stellen den Schwierigkeitsgrad der Tests dar (Tab.26).

Tab. 26. Der Schwierigkeitsindex und die Trennschärfen zu den Polla Items zu  $t_1$  und  $t_2$ .

N=81	$t_1$		$t_2$		Korrigierte Trennschärfe	
	Summe	Index	Summe	Index	$t_1$	$t_2$
Strümpfe anziehen	83,00	0,51	125,00	0,77	,34	,32
Aufstehen vom Hocker	89,00	0,55	134,00	<b>0,83</b>	,46	,37
Einbeinstand	48,00	0,30	122,00	0,75	,49	,35
Aufheben	147,00	<b>0,91</b>	160,00	<b>0,99</b>	<b>,25</b>	<b>,20</b>
Hinlegen	87,00	0,54	138,00	<b>0,85</b>	,38	,39
Treppauf eine Stufe	39,00	0,24	109,00	0,67	,64	,63
Hocken	7,00	<b>0,04</b>	40,00	0,25	,59	,47
Hinknien	28,00	<b>0,17</b>	87,00	0,54	,48	,52
Auf der Stelle laufen	16,00	<b>0,10</b>	106,00	0,65	,44	,49
Einbeinhüpfen	2,00	<b>0,01</b>	21,00	<b>0,13</b>	,42	,47
1 Minute-Up & Go-Test	94,00	0,58	115,00	0,71	,43	,58
2-Minuten-Gehtest	96,00	0,59	106,00	0,65	,44	,41
Treppenlaufen-Test	88,00	0,54	116,00	0,72	,56	,59
2-Minuten-Lauftest	0,00	<b>0,00</b>	7,00	<b>0,04</b>	-	,29

Anm.: Werte unter 0,2 = schwere Items; Werte über 0,8 = leichte Items.

Besonders schwierige oder sehr leichte Aufgaben enthalten wenige Informationen, da sie keine Personenunterschiede deutlich machen (Bortz & Döring, 2006) und können damit u. U. aus der Testbatterie entfernt werden. Erstrebenswert sind mittlere Itemschwierigkeiten von 0,5 (Bortz & Döring, 2006). Ein Wert unter 0,2 spricht für ein schweres und ein Wert über 0,8 für ein leichtes Item. Grundlage der Berechnung von Aufgabenschwierigkeiten in der vorliegenden Untersuchung ist der Schwierigkeitsindex für mehrstufige Items nach Dahl (1971). Der Schwierigkeitsindex errechnet sich aus der erreichten Punktsumme aller Patienten auf dem jeweiligen Item dividiert durch die maximal erreichbare Summe auf diesem Item.

Das Aufheben eines Gegenstands stellt sich zu beiden Zeitpunkten mit einem Wert von 0,91 und 0,99 als leichtestes Item dar. Lediglich ein Patient konnte den Test

nicht durchführen. Der Lauftest erwies sich als besonders schwierig und konnte zu Therapiebeginn (0,00) von keinem und gegen Ende (0,04) von nur vier Patienten durchgeführt werden. Die Aktivitäten des Hockens, Hinkniens, Laufens und Einbeinhüpfens wiesen zu Beginn große Schwierigkeiten von unter 0,2 auf, die sich zum Messzeitpunkt Therapieende aber nur noch lediglich für das Hüpfen manifestieren. Der korrigierte Trennschärfekoeffizient, als die Korrelation der Beantwortung eines Items mit dem Gesamtestwert (Bortz & Döring, 2006), weist für Personen, die im Gesamtergebnis einen hohen Wert erreichen, auf einem trennscharfen Item ebenfalls eine hohe Punktzahl auf. Für den Polla gibt die Trennschärfe überwiegend mittlere und damit akzeptable Werte von 0,3 bis 0,5 für beide Zeitpunkte an (Weise, 1975, S. 219). Darunter liegt lediglich das Aufheben.

Tab. 27. Der Schwierigkeitsindex und die Trennschärfen zur Selbsteinschätzung zu  $t_1$  und  $t_2$ .

N=81	$t_1$		$t_2$		Trennschärfe	
	Summe	Index <sup>1</sup>	Summe	Index <sup>1</sup>	$t_1$	$t_2$
Auf dem Boden knien	98,80	0,30	195,60	0,60	,42	,49
Strümpfe anziehen	200,00	0,62	269,00	<b>0,83</b>	,41	,57
Hinlegen und wieder aufstehen	168,80	0,52	242,00	0,75	,52	,62
Eine Stunde auf einem Stuhl sitzen	200,00	0,62	257,00	0,79	,38	,55
Aus einem Auto aussteigen	197,00	0,61	245,80	0,76	,45	,60
Aufstehen nach längerem Sitzen	189,00	0,58	216,00	0,67	,41	,42
Auf dem betroffenen Bein stehen	164,00	0,51	262,00	<b>0,81</b>	,51	,71
20 – 30 Min stehen	132,00	0,41	201,00	0,62	,58	,77
Eine Straßenkreuzung gehen	219,30	0,68	283,00	<b>0,87</b>	,52	,72
Mehrere Straßenkreuzungen gehen	163,00	0,50	239,00	0,74	,65	,74
Einen langen Spaziergang machen	116,00	0,36	205,50	0,63	,69	,76
Eine Einkaufstasche tragen	183,00	0,56	274,00	<b>0,85</b>	,59	,55
Auf den Zehenspitzen gehen	133,40	0,41	235,00	0,73	,77	,74
Auf den Hacken gehen	124,50	0,38	233,10	0,72	,74	,77
Ein Stockwerk hinaufgehen	218,00	0,67	279,00	<b>0,86</b>	,48	,68
Ein Stockwerk hinuntergehen	217,00	0,67	270,00	<b>0,83</b>	,43	,68
In die Hocke gehen	83,00	0,26	159,00	0,49	,56	,56
Auf dem betroffenen Bein hüpfen	37,00	<b>0,11</b>	141,00	0,44	,48	,67
Auf der Stelle laufen	133,30	0,41	249,00	0,77	,54	,68
Ein kleines Stück rennen	43,50	<b>0,13</b>	138,00	0,43	,52	,68
Eine Zeit lang joggen	41,10	<b>0,13</b>	130,00	0,40	,46	,62
Schwimmen gehen	101,00	0,31	189,00	0,58	,35	,33
Einen Schuh vom Boden aufheben	266,50	<b>0,82</b>	301,00	<b>0,93</b>	<b>,28</b>	,57
Radfahren im Freien	60,20	<b>0,19</b>	162,80	0,50	,35	,37
Über Kopfsteinpflaster gehen	160,00	0,49	239,00	0,74	,75	,76
Die Toilette benutzen	269,00	<b>0,83</b>	306,00	<b>0,94</b>	,41	,47
Eine Kiste Wasser heben	123,00	0,38	222,20	0,69	,48	,51

Anm.: <sup>1</sup>Schwierigkeitsindex nach Dahl (1971), fettgedruckte Zahlen: Werte unter 0,2 = schwere Items; Werte über 0,8 = leichte Items.

Der Schwierigkeitsindex (Dahl, 1971) bezüglich der Selbsteinschätzung von Aktivitäten liegt zu  $t_1$  für das Einbeinhüpfen, ein kleines Stück rennen, eine Zeit lang Joggen sowie für das Radfahren im Freien bei Werten unter 0,2 und charakterisiert diese als schwierige Items (Tab. 27).

Das Aufheben eines Schuhs sowie die Toilette benutzen weisen Werte von über 0,8 auf und gelten damit als leicht. Zu  $t_2$  zeigen sich dagegen sechs Items mit geringem Schwierigkeitsgrad von über 0,8. Die Trennschärfe zeigt mittlere bis hohe Werte für beide Zeitpunkte außer beim Aufheben eines Schuhs.

### 5.1.2 Faktorenanalyse zum Polla

Zur Abschätzung, in wie weit in dem Konstrukt des Polla unterschiedliche inhaltliche Dimensionen abgebildet werden, ist eine explorative Faktorenanalyse zu Therapiebeginn durchgeführt worden. Der Lauftest musste aufgrund seiner Nullvarianz aus der Testbatterie entfernt werden. Die Ergebnisse extrahieren bei einem Eigenwertkriterium von über eins vier Faktoren mit einer Aufklärung von 62 Prozent der Gesamtvarianz (Tab. 28). Nach der Extraktionsmethode der Hauptkomponentenanalyse (Tab. 29) und einer Oblimin-Rotation mit Kaiser-Normalisierung lassen die Faktoren keine eindeutigen inhaltlichen Dimensionen erkennen (Tab. 30).

Tab. 28. Erklärte Gesamtvarianz nach der Extraktionsmethode der Hauptkomponentenanalyse ( $t_1$ ).

N=81 Komponente	Anfängliche Eigenwerte		
	Gesamt	% der Varianz	Kumulierte %
1	4,32	33,23	33,23
2	1,40	10,76	44,00
3	1,25	9,62	53,62
4	1,04	8,03	61,65
5	0,90	6,89	68,54

Tab. 29. Extraktionsmethode der Hauptkomponentenanalyse: Komponentenmatrix ( $t_1$ ).

N=81	Komponente					Komponente			
	1	2	3	4		1	2	3	4
Strümpfe anziehen	,37	<b>,61</b>	,01	,21	Hinknien	<b>,57</b>	,16	,42	-,39
Aufstehen vom Hocker	<b>,60</b>	-,05	<b>-,50</b>	,07	Laufen	<b>,55</b>	,40	<b>-,54</b>	-,10
Einbeinstand	<b>,58</b>	,20	,02	-,10	Einbeinhüpfen	<b>,53</b>	,11	-,28	-,25
Aufheben	,29	,35	,11	<b>,73</b>	Up & Go-Test	<b>,61</b>	-,48	-,13	,27
Hinlegen	,42	,39	<b>,54</b>	-,02	Gehtest	<b>,58</b>	-,35	,12	,31
Treppauf eine Stufe	<b>,77</b>	-,06	-,08	-,19	Treppenlaufen	<b>,71</b>	-,32	,25	,07
Hocken	<b>,73</b>	-,25	,18	-,15					

Anm.: Nur fettgedruckte Werte  $\geq \pm ,50$  werden für eine Dimensionierung als relevant erachtet.

Tab. 30. Hauptkomponentenanalyse mit Oblimin-Rotation ( $t_1$ ).

N=81	Oblimin-Rotation			
	Komponente			
	1	2	3	4
Strümpfe anziehen	-,01	,40	,37	<b>,60</b>
Aufstehen vom Hocker	,47	<b>,71</b>	-,00	,10
Einbeinstand	,31	<b>,50</b>	,46	,17
Aufheben	,22	,11	,10	<b>,84</b>
Hinlegen	,14	,12	<b>,73</b>	,36
Treppauf eine Stufe	<b>,58</b>	<b>,65</b>	,46	-,02
Hocken	<b>,68</b>	,41	<b>,51</b>	-,09
Hinknien	,29	,31	<b>,80</b>	-,04
Auf der Stelle laufen	,13	<b>,85</b>	,16	,23
Einbeinhüpfen	,26	<b>,64</b>	,27	-,04
1 Minute-Up & Go-Test	<b>,80</b>	,34	,01	,04
2-Minuten-Gehtest	<b>,74</b>	,21	,19	,17
Treppenlaufen-Test	<b>,77</b>	,29	,44	,04

Anm.: Nur fettgedruckte Werte  $\geq \pm ,50$  werden für eine Dimensionierung als relevant erachtet.

### 5.1.3 Faktorenanalyse zur Selbsteinschätzung

Eine Faktorenanalyse zur Selbsteinschätzung zu Therapiebeginn zeigt nach der Extraktionsmethode der Hauptkomponentenanalyse (Eigenwert >1) sieben Faktoren mit einer Varianzaufklärung von 71 Prozent (Tab. 31). Inhaltliche Dimensionen lassen sich auch hier nicht eindeutig zuordnen.

Tab. 31. Die erklärte Gesamtvarianz nach der Extraktionsmethode der Hauptkomponentenanalyse.

N=81 Komponente	Anfängliche Eigenwerte		
	Gesamt	% der Varianz	Kumulierte %
1	8,67	32,09	32,09
2	2,99	11,06	43,14
3	2,20	8,18	51,39
4	1,66	6,14	57,46
5	1,43	5,31	62,76
6	1,20	4,43	67,20
7	1,04	3,85	<b>71,04</b>

Tab. 32. Extraktionsmethode nach der Hauptkomponentenanalyse mit Oblimin-Rotation zu  $t_1$ .

N=81	Komponente						
	1	2	3	4	5	6	7
Auf dem Boden knien	,47	-,15	-,00	<b>,56</b>	,19	-,15	,15
Strümpfe anziehen	,44	,28	,44	-,20	,21	-,24	,19
Hinlegen und wieder aufstehen	<b>,55</b>	,15	,26	-,08	,22	-,38	,20
Eine Stunde auf einem Stuhl sitzen	,42	,48	,36	,04	,22	-,06	-,43
Aus einem Auto aussteigen	,49	<b>,50</b>	,30	,08	,13	-,08	-,19
Aufstehen nach längerem Sitzen	,45	<b>,55</b>	,24	,10	,18	-,04	-,06
Auf dem betroffenen Bein stehen	<b>,58</b>	,04	-,08	-,09	<b>-,55</b>	-,31	-,16
20-30 Min stehen	<b>,64</b>	,18	-,32	-,17	,17	-,01	-,03
Eine Straßenkreuzung gehen	<b>,60</b>	,26	-,40	-,35	,18	,27	,01
Mehrere Straßenkreuzungen gehen	<b>,71</b>	,16	-,29	-,35	,23	,11	-,03
Einen langen Spaziergang machen	<b>,75</b>	-,15	-,30	-,28	,02	-,13	-,17
Eine Einkaufstasche tragen	<b>,63</b>	-,27	,08	-,44	-,28	-,03	,00
Auf den Zehenspitzen gehen	<b>,78</b>	-,16	-,06	,26	-,30	-,16	-,06
Auf den Hacken gehen	<b>,79</b>	-,08	-,01	,22	-,29	-,15	-,06
Ein Stockwerk hinaufgehen	<b>,55</b>	,31	<b>-,53</b>	,11	-,03	,21	,28
Ein Stockwerk hinuntergehen	<b>,52</b>	,40	-,49	,26	-,01	-,01	,30
In die Hocke gehen	<b>,61</b>	-,18	,02	,30	,14	-,23	,10
Auf dem betroffenen Bein hüpfen	<b>,54</b>	<b>-,53</b>	-,06	,03	,20	-,03	-,03
Auf der Stelle laufen	<b>,59</b>	-,31	,02	,30	-,21	,28	-,08
Ein kleines Stück rennen	<b>,56</b>	-,36	,02	,36	,24	,27	-,18
Eine Zeit lang Joggen	,48	<b>-,55</b>	,15	,10	,34	,19	-,18
Schwimmen gehen	,38	-,45	,21	-,39	,28	,02	,19
Einen Schuh vom Boden aufheben	,31	,39	,36	,05	-,13	<b>,60</b>	,00
Radfahren im Freien	,37	-,38	,37	-,13	-,01	,07	<b>,51</b>
Über Kopfsteinpflaster gehen	<b>,79</b>	,01	-,08	-,16	-,15	,03	-,16
Die Toilette benutzen	,43	,45	,41	,14	-,26	,15	,29
Eine Kiste Wasser heben	<b>,50</b>	-,24	,47	-,23	-,28	,20	-,00

Anm.: Nur fettgedruckte Werte  $\geq \pm ,50$  werden für eine Dimensionierung als relevant erachtet.

Tabelle 32 zeigt die einzelnen Faktorladungen nach der Extraktionsmethode nach der Hauptkomponentenanalyse mit Oblimin-Rotation, die überwiegend auf Faktor 1 laden.

#### 5.1.4 Konsequenzen für die Summenskalenbildung

Wird eine Betrachtung der Werte aus den explorativen Faktorenanalysen von unter 0,5 vernachlässigt, so lassen sich im Hinblick auf die Zuordnung inhaltlicher Dimensionen beim Polla und bei der Selbsteinschätzung keine eindeutigen Möglichkeiten erkennen, sie laden überwiegend auf einem Faktor. Von einer Faktoren-

bildung wird daher abgesehen und mit den Summenskalen zu beiden Instrumenten weitergearbeitet. Aufgrund der Itemanalyse findet der Laufstest als schwieriges und das Aufheben eines Gegenstandes als besonders leichtes Item bei der Summenskalenbildung keine Berücksichtigung. Die beiden Tests werden allerdings in den Ergebnissen zu den einzelnen Items im Polla dargestellt.

## 5.2 Datenaufbereitung

Die Aufbereitung des Datenmaterials (Polla, Selbsteinschätzung, SF-36) zielt zur besseren Vergleichbarkeit der Ergebnisse miteinander auf die Bildung von Summenskalen und eine Transformation auf Prozentniveau.

Die Berechnung der Skalenrohwerter des SF-36 lehnt sich an das Test-Manual von Bullinger und Kirchberger (1998) an. Nach der Umkodierung und Gewichtung (ANHANG 3) von zehn Items sowie der Berechnung von Skalenwerten durch Addition der Items zu Subskalen (Skalenrohwerter) wurden die Werte in eine Prozentskala gebracht (Transformierte Skalenwerte):

$$\text{Prozentskala} = \frac{(\text{Tatsächlicher Rohwert} - \text{niedrigst möglicher Rohwert}) \times 100}{\text{Mögliche Spannweite der Rohwerte}}$$

Die Werte der Subskalen reichen von 0 (schlechteste Lebensqualität) bis 100 (beste Lebensqualität). Neben der Auswertung auf Ebene der Subskalen ist die Bildung zweier übergeordneter Summenskalen (Körperliche Funktion und Psychisches Befinden) möglich, die in diesem Zusammenhang nicht vollzogen wurde, da zur Bearbeitung der zentralen Fragestellung von Zusammenhängen die Körperliche Funktionsfähigkeit als eigenständige Subskala notwendig war. Fehlende Werte sind nach dem standardisierten Scoring-Algorithmus durch den Mittelwert der vorhandenen Items ersetzt, sofern mindestens 50% einer Dimension beantwortet wurde (Bullinger & Kirchberger, 1998).

Die Berechnung der *Summenskala zum Polla* erfolgt durch die Addition der zwölf Items jeweils zu den Zeitpunkten  $t_1$  und  $t_2$  sowie zum Follow-up und zur Gruppe von Gesunden (ausgeschlossen: Aufheben und Laufstest). Entsprechend waren bei 12 uneingeschränkt durchführbaren Tests maximal 24 Punkte zu erreichen.

Die *Datenverarbeitung zur Selbsteinschätzung* von Aktivitäten umfasste die Bildung von Mittelwerten auf Itemebene, ein Ersetzen fehlender Werte durch die Zuweisung von Mittelwerten und die Bildung von Summenskalen zu den Zeitpunkten Therapiebeginn und -abschluss, zur Vergleichsgruppe und zum Follow-up. Bei

einer erreichbaren Punktzahl von vier auf jedem Item, ergibt sich eine Gesamtpunktzahl von 108 für die Selbsteinschätzung.

In Anlehnung an die oben genannte Formel erfuhren der Polla, die Selbsteinschätzung und die gemessenen Tests ebenfalls eine *Transformation* auf Prozentniveau.

Für die Transformation der gemessenen Tests dienten der jeweils kleinste aufgenommene Rohwert zum jeweiligen Test aus allen Gruppen und zu allen Zeitpunkten sowie die Spannweite der Rohwerte als Berechnungsgrundlage. Sie setzen sich wie folgt zusammen:

<i>1-Minute-Up &amp; Go-Test:</i>	Kleinster Wert: 3	Spannweite:	12,5
<i>2-Minuten-Gehtest:</i>	Kleinster Wert: 56	Spannweite:	218
<i>Treppenlaufen-Test:</i>	Kleinster Wert: 132,97	Spannweite:	121,97

Abschließend wurde aus den drei transformierten Skalen der gemessenen Tests eine Summenskala gebildet, deren Anwendung lediglich im Hinblick auf die Betrachtung der Konstruktvalidität findet.

## 6 ERGEBNISSE

### 6.1 Deskriptive Statistik zu Therapiebeginn

#### 6.1.1 Deskriptive Statistik zum Polla

Die *Datenerhebung* konnte zu Untersuchungsbeginn ohne fehlende Werte abgeschlossen werden. Die *Verteilungen zum Polla* zeigen, dass über 90 Prozent der Teilnehmer nicht in der Lage waren, die Aktivitäten des Hockens, Einbeinhüpfens oder des Laufens auszuführen. 70-90 Prozent konnten weder auf der Stelle laufen noch auf dem Boden knien (Tab. 33).

Tab. 33. Die Häufigkeiten zu den Polla Items zu  $t_1$ .

N=81	Gültige Prozente		
	Nein	Eingeschränkt	Ja
Strümpfe anziehen	12,3	72,8	14,8
Aufstehen vom Hocker	2,5	85,2	12,3
Einbeinstand	50,6	39,5	9,9
Aufheben	1,2	16,0	82,7
Hinlegen/Aufstehen	32,1	28,4	39,5
Treppauf eine Stufe	58,0	35,8	6,2
Hocken	92,6	6,2	1,2
Hinknien	77,8	9,9	12,3
Laufen auf der Stelle	84,0	12,3	3,7
Einbeinhüpfen	98,8	0	1,2
Up & Go-Test	0	<b>84,0</b>	16,0
Gehtest	0	<b>81,5</b>	18,5
Treppenlaufen	2,5	<b>86,4</b>	11,1
Lauftest	100	0	0

30-60 Prozent der Teilnehmer konnten keinen Einbeinstand ausführen, sich auf den Boden hinlegen oder eine Stufe treppauf steigen. Von besonderer Bedeutung im Hinblick auf die Konstruktvalidität ist, dass es allen 81 Patienten möglich war, den Up & Go-Test, den Gehtest sowie den Treppenlaufen-Test durchzuführen, bei diesen Tests traten aber Einschränkungen in über 80% aller Fälle auf (Tab. 33).

Die *Verteilungen zu den gemessenen Tests im Polla* veranschaulichen für den 1-Minute-Up & Go-Test, dass die Strecke vom Hocker bis zur Wand und wieder zurück zum Hocker, vier bis achtmal von 81 Prozent (n=66; Abb. 6) und neun bis 15 Mal von 17 Prozent (n=14) der Patienten bewältigt wurde. Der Mittelwert der zu-

rückgelegten Wegstrecken beträgt 6,6 ( $\pm 2$ ), bei einem Maximum von 13 und einem Minimum von drei Wegstrecken.

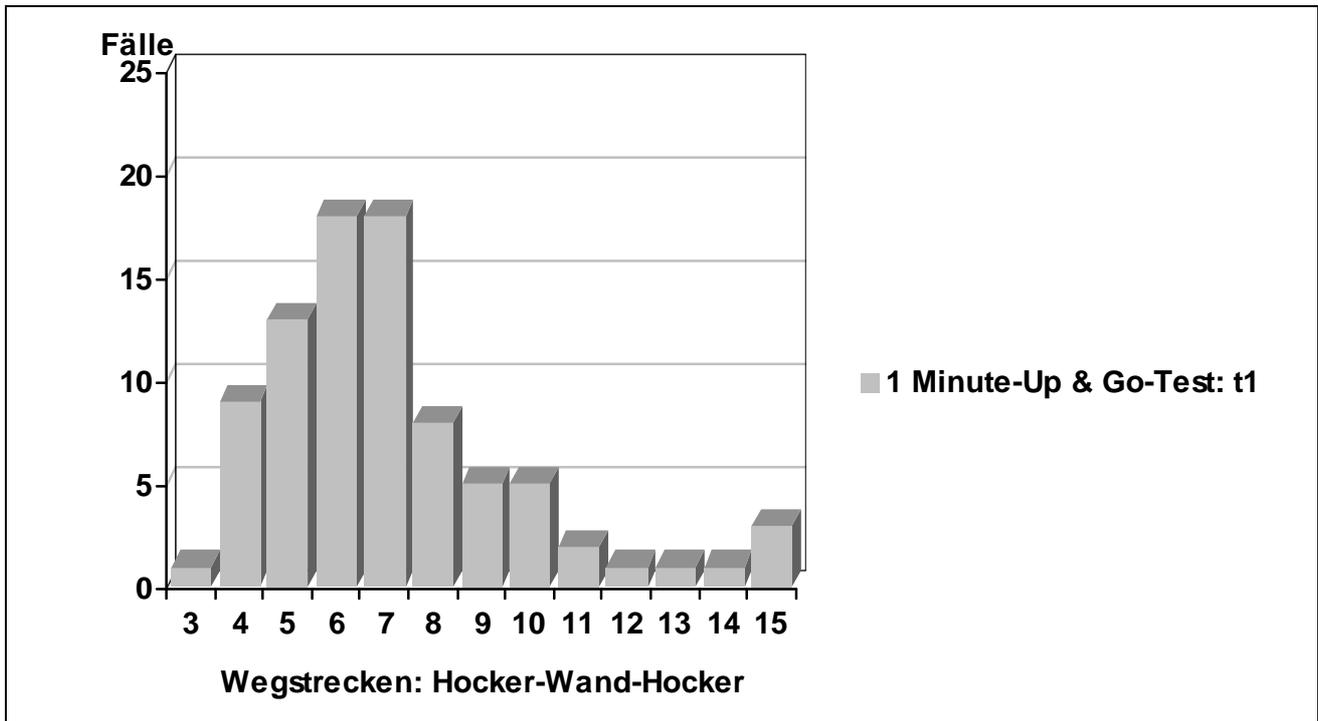


Abb. 6. Die Verteilungen zum 1-Minute-Up & Go-Test zu t<sub>1</sub> (N=81).

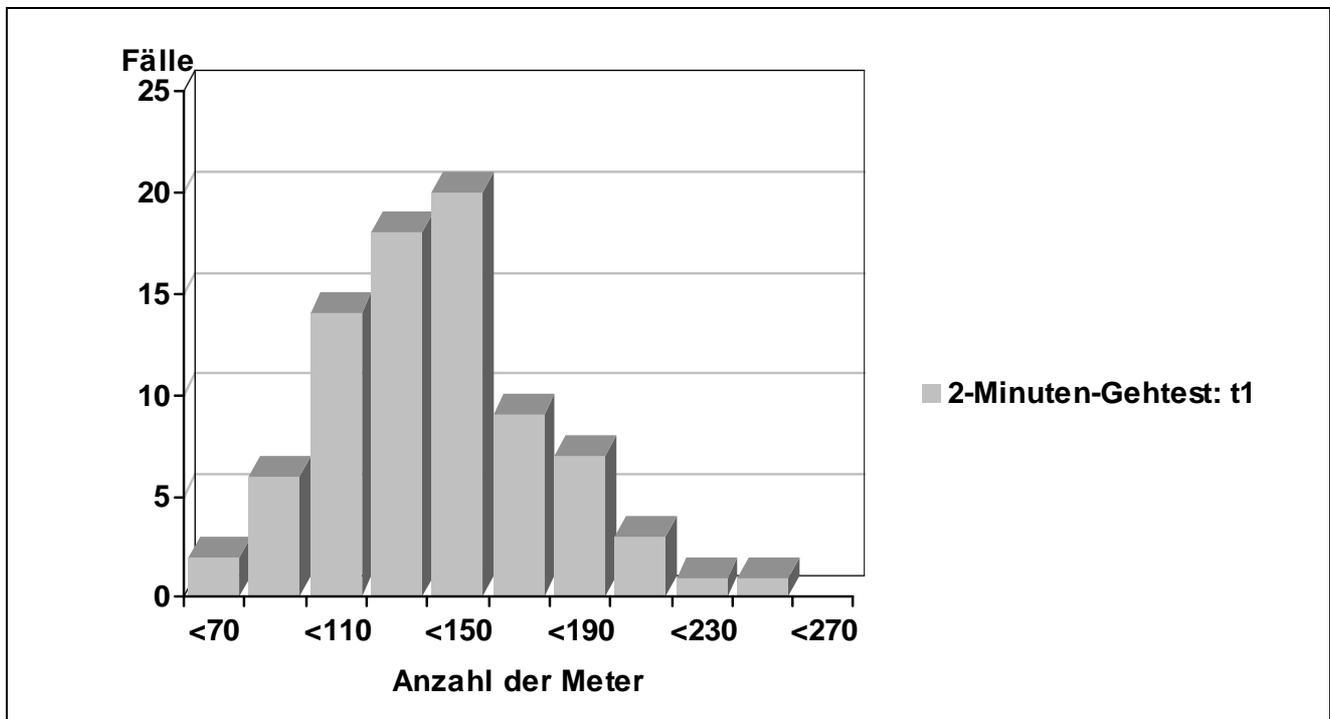


Abb. 7. Die Verteilungen zum 2-Minuten-Gehtest zu t<sub>1</sub> (N=81).

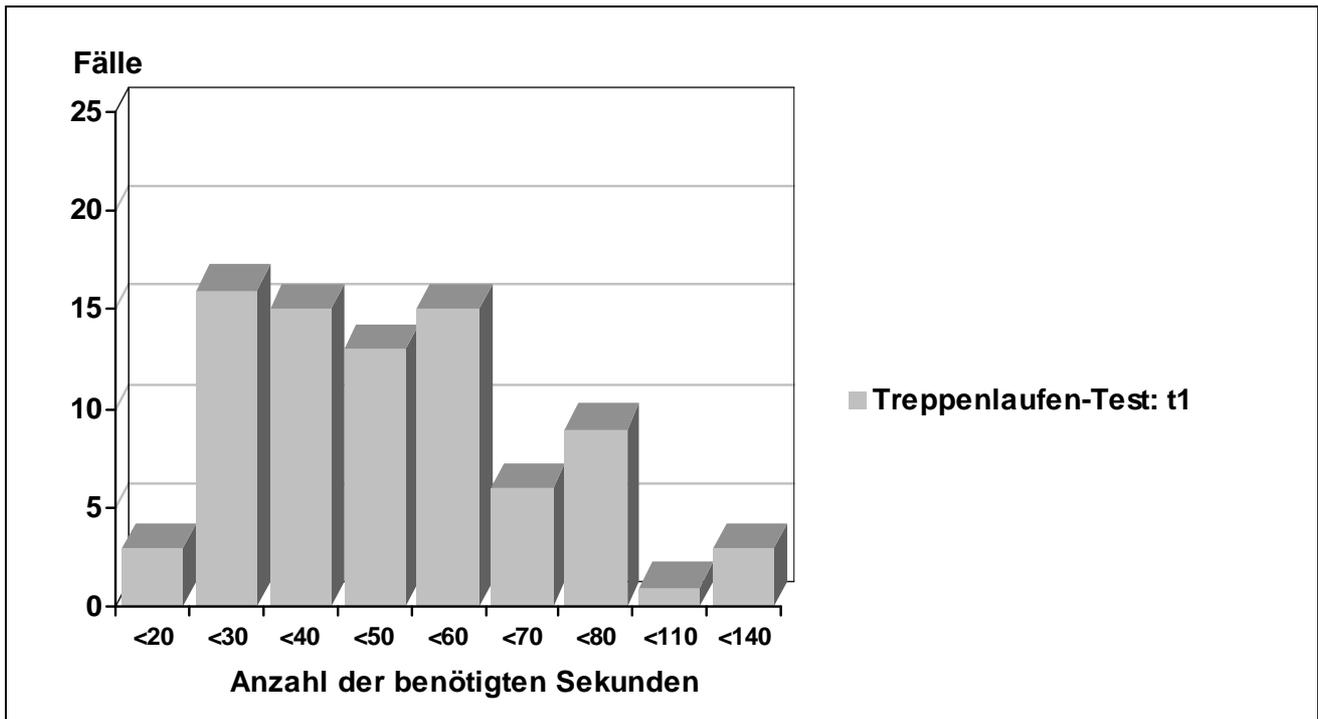


Abb. 8. Die Verteilungen zum Treppenlaufen-Test zu  $t_1$  (N=81).

Im *2-Minuten-Gehtest* legten 27 Prozent (n=22) eine Strecke zwischen 56 und 110 Metern sowie 70 Prozent (n=57) eine Distanz zwischen 110 und 210 Metern zurück (Abb. 7), zwei Personen gingen weiter als 210 Meter. Bei einem Maximum von 240 Metern und einem Minimum von 56 Metern betrug die Gehstrecke im Mittel 132 Meter ( $\pm 35,5$ ).

Für die Bewältigung der beiden Treppenabsätze benötigen 80 Prozent (n=65) zwischen 20 und 69 Sekunden und 16 Prozent (n=13) länger als 70 Sekunden (Abb. 8). Zwei Patienten, die den *Treppenlaufen-Test* nicht absolvieren konnten, wurde der Wert von 133 Sekunden zugewiesen, um diese damit gewichtet mit in die Auswertung aufnehmen zu können. Der Wert von 133 Sekunden entspricht der längsten, für die Bewältigung der beiden Treppenabsätze benötigten Zeit. Im Durchschnitt waren, bei einem Maximum von 133 und einem Minimum von 14 Sekunden, 48 Sekunden ( $\pm 25$ ) für das Treppenlaufen nötig. Einen abschließenden Überblick über die deskriptive Statistik der gemessenen Tests gibt Tabelle 34.

Tab. 34. Die deskriptive Statistik zu den gemessenen Tests zum Messzeitpunkt  $t_1$ .

N=81	Min	Max	M	s
1-Minute-Up & Go-Test <sup>1</sup>	3,00	13,00	6,56	1,98
2-Minuten-Gehtest <sup>2</sup>	56,00	240,00	131,86	35,54
Treppenlaufen-Test <sup>3</sup>	14,00	132,97	48,36	24,78
2-Minuten-Lauftest <sup>2</sup>	,00	,00	,00	,00

Anm.: Angaben in <sup>1</sup>Wegstrecken; <sup>2</sup>Metern; <sup>3</sup>Sekunden

### 6.1.2 Deskriptive Statistik zu den Summenskalen

Die Verteilungen der transformierten Summenskalen zum Polla, zur Selbsteinschätzung und zur Körperlichen Funktionsfähigkeit (SF-36) zu Therapiebeginn finden an dieser Stelle Betrachtung. Die weiteren Skalen zum SF-36 werden unter Kapitel 6.3.3 beschrieben. Fehlende Werte zur Körperlichen Funktionsfähigkeit traten in einem Prozent und in 2,7 Prozent zur Selbsteinschätzung auf.

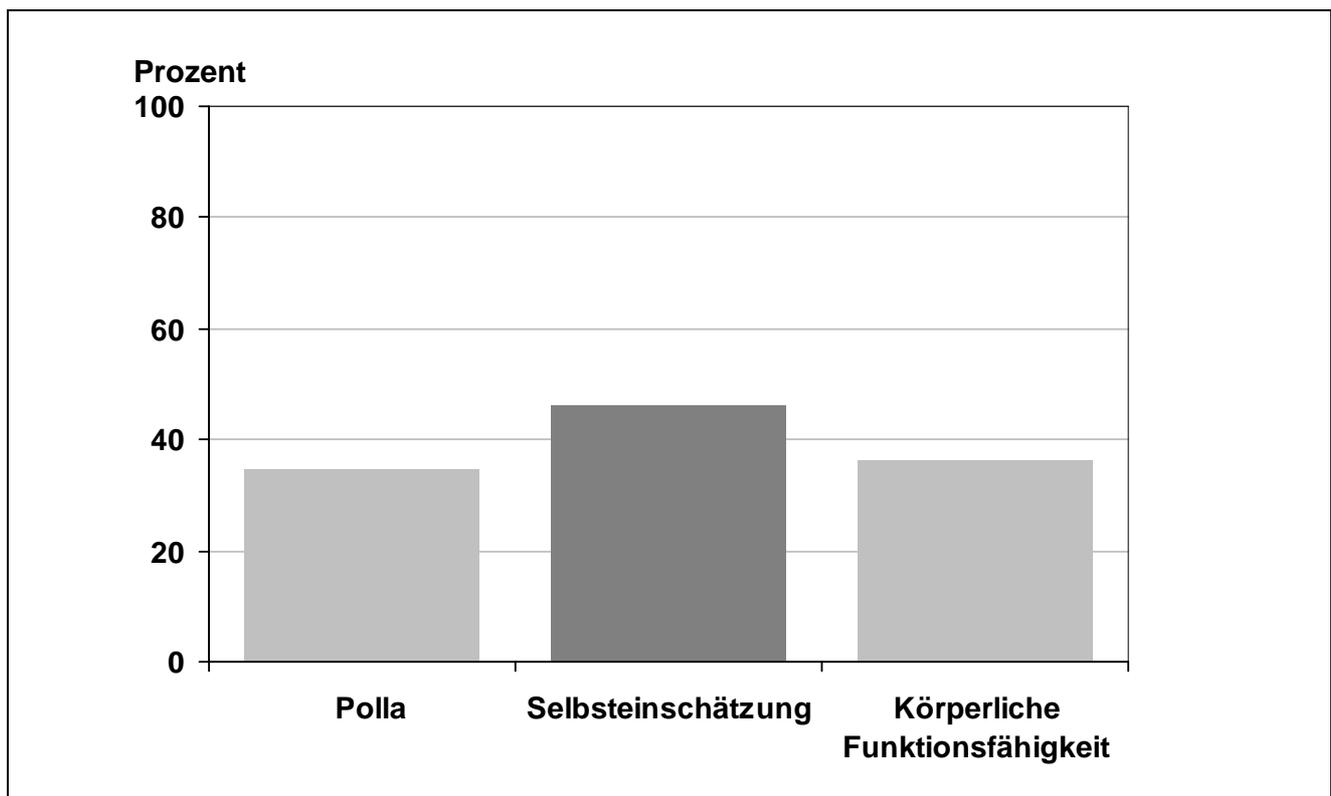


Abb. 9. Die Mittelwerte der Summenskalen zu  $t_1$  (N=81).

60 Patienten (74%) zum Polla, 52 Patienten (64%) zur Selbsteinschätzung und 44 Patienten (65%) zur Körperlichen Funktionsfähigkeit erreichen 21-50 Prozent des maximal erreichbaren Punktwertes auf der jeweiligen Prozentskala.

Tabelle 35 und Abbildung 9 veranschaulichen die deskriptive Statistik zu den Summenskalen der Messinstrumente.

Tab. 35. Mittelwerte und Standardabweichungen zu den Summenskalen (0-100) zu  $t_1$ .

N=81	M	S
Polla	34,83	14,38
Selbsteinschätzung	46,19	16,56
Körperliche Funktionsfähigkeit	36,38	19,19

### 6.1.3 Ergebnisse der spezifischen Zielkriterien

Die *Datenerhebung* der spezifischen Zielkriterien zum Untersuchungsbeginn konnte aus Gründen unvollständig ausgefüllter Befunde nicht ohne fehlende Werte abgeschlossen werden. Erfasst wurden das Hauptproblem (n=76), die Medikamenteneinnahme (n=71), das Giving-Way Phänomen (n=73), die Schmerzdauer (n=72), die Schmerzintensität (NRS, n=76), die aktive Flexion (n=78) und die schmerzfreie Kniebeuge (n=48).

Die Ergebnisse zu dem *Hauptproblem* zeigen, dass für 59,2 Prozent der Fälle (n=45 von 76) das von dem einzelnen Patienten angegebene Problem permanent gegenwärtig war. 38,1 Prozent (n=29) äußerten dies ein oder mehrmals täglich und 2,6 Prozent (n=2) lediglich wöchentlich.

Eine regelmäßige *Medikamenteneinnahme* zu Therapiebeginn gaben 43,7 Prozent (n=31 von 71) an, 25,4 Prozent (n=18) äußerten eine nach Bedarf und 31 Prozent (n=22) gaben keine Medikamenteneinnahme an.

Das *Giving-Way Phänomen* trat in 12,4 Prozent (n=9 von 73) der Fälle täglich, in 31,5 Prozent (n=23) wöchentlich, dagegen in 56,2 Prozent (n=41) der Fälle nicht auf.

Die *Schmerzdauer* wurde von 18,1 Prozent (n=13 von 72) als ständig, von 20,8 Prozent (n=15) als täglich zwischen einer und mehreren Stunden und von 54,1 Prozent (n=39) als kurz oder bis zu einigen Minuten anhaltend beschrieben. Keine Schmerzen gaben 6,9 Prozent (n=5) an.

Die Mittelwerte der intervallskalierten *spezifischen Zielkriterien* zeigen für die Schmerzintensität (NRS) einen Wert von 3,8, für die schmerzfreie Kniebeuge 63 Grad Beugung und die Beweglichkeit 94 Grad an (Tab. 36).

Tab. 36. Die deskriptive Statistik zur Schmerzintensität, Flexion und Kniebeuge zu  $t_1$ .

N=81	N	Min	Max	M	s
Schmerzintensität (NRS)	76	0	8,50	3,82	2,06
Aktive Flexion in Grad	78	10,0	142,00	94,17	20,29
Schmerzfremde Kniebeuge in Grad	48	0	145,00	63,04	31,62

## 6.2 Deskriptive Statistik zu Therapieende

### 6.2.1 Deskriptive Statistik zum Polla

Die Datenerhebung zum Polla konnte gegen Therapieabschluss ohne fehlende Werte abgeschlossen werden. Über 80 Prozent der Patienten waren nicht in der Lage, auf einem Bein zu hüpfen oder den Laufstest durchzuführen und für 58 Prozent war die Aktivität des Hockens nicht möglich (Tab. 37).

Dagegen waren das Strümpfe anziehen, Aufstehen vom Hocker, Aufheben, Hinlegen, der Einbeinstand, der Up & Go-Test, der Gehstest sowie das Treppenlaufen von nahezu allen Patienten durchführbar. Die gemessenen Tests im Polla weisen mit über 50 Prozent auch zu  $t_2$  mehr Einschränkungen auf, als die anderen Tests im Polla. Lediglich das Aufheben konnte von 80 Patienten ohne Einschränkungen ausgeführt werden (Tab. 37).

Tab. 37. Die Häufigkeiten zu den Polla Items zu  $t_2$ .

N=81	Gültige Prozente		
	Nein	Eingeschränkt	Ja
Strümpfe anziehen	1,2	43,2	55,6
Aufstehen vom Hocker	0	34,6	65,4
Einbeinstand	2,5	44,4	53,1
Aufheben	1,2	0	98,8
Hinlegen/Aufstehen	1,2	27,2	71,6
Treppauf eine Stufe	4,9	55,6	39,5
Hocken	58,0	34,6	7,4
Hinknien	33,3	25,9	40,7
Laufen auf der Stelle	17,3	34,6	48,1
Einbeinhüpfen	80,2	13,6	6,2
Up & Go Test	0	<b>58,0</b>	42,0
Gehstest	0	<b>69,1</b>	30,9
Treppenlaufen-Test	0,00	<b>56,8</b>	43,2
Laufstest	95,1	1,2	3,7

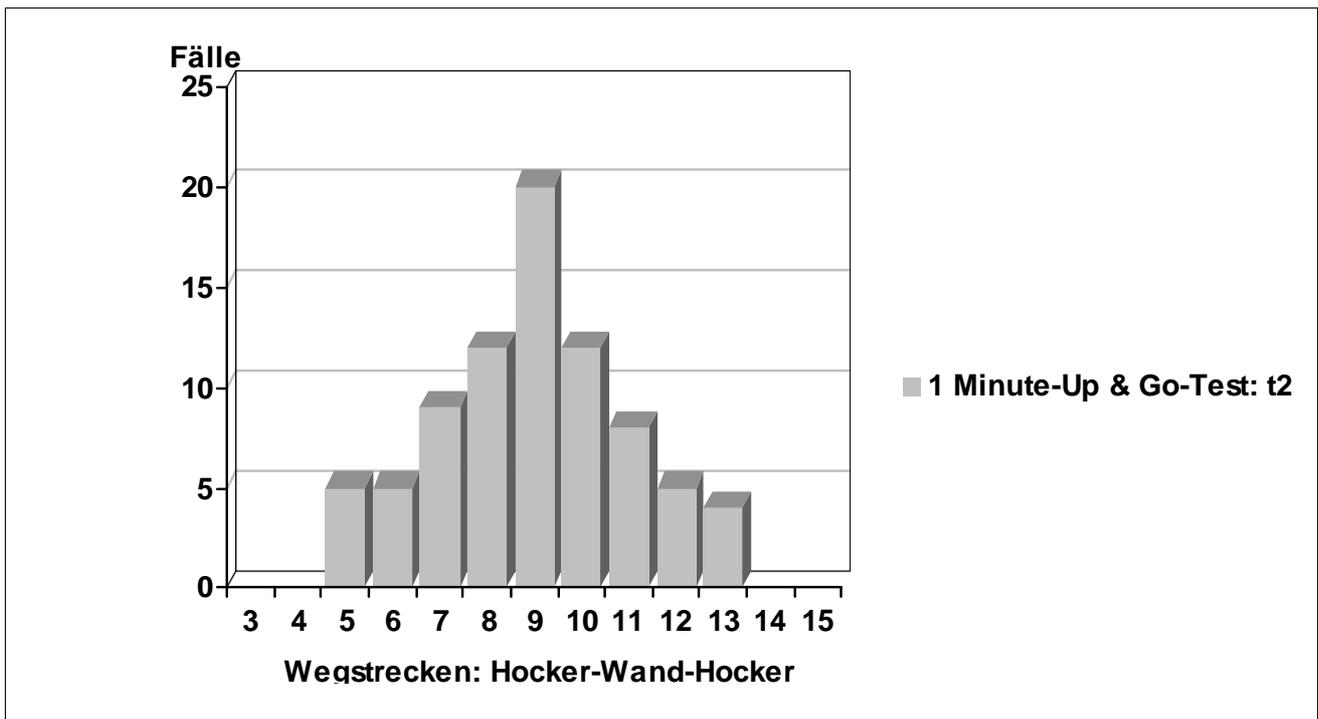


Abb. 10. Die Verteilungen zum 1-Minute-Up & Go-Test zu  $t_2$  (N=81).

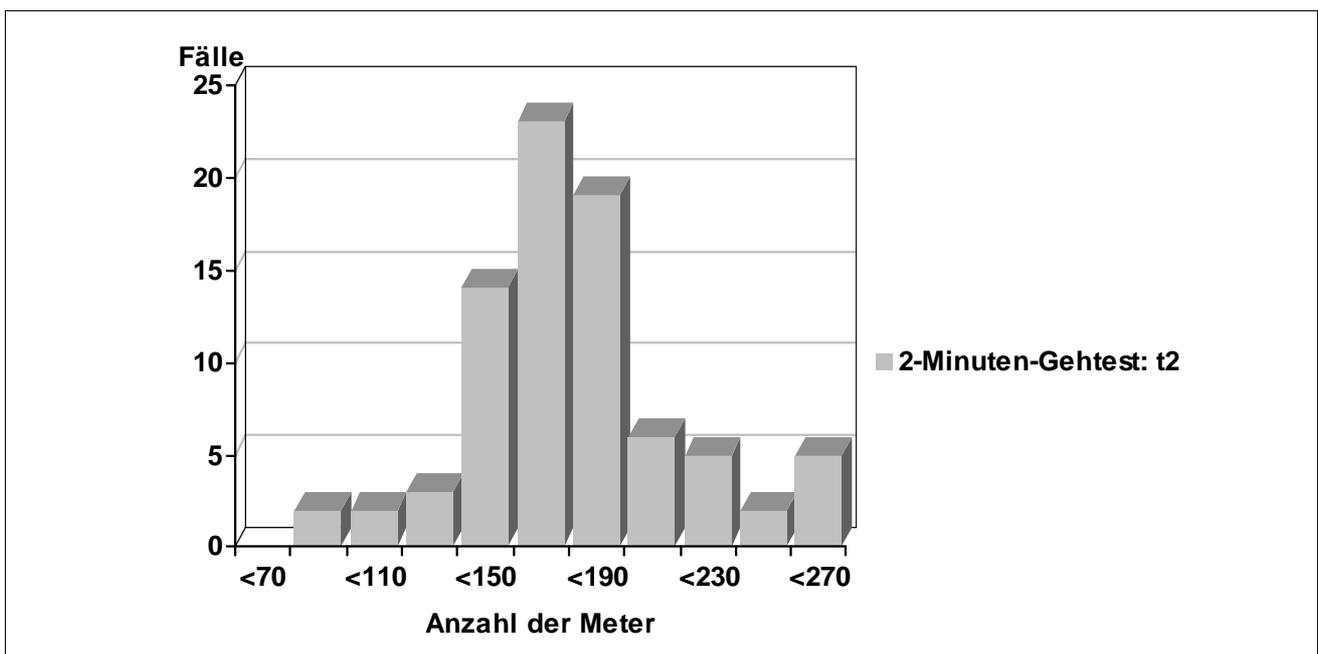


Abb. 11. Die Verteilungen zum 2-Minuten-Gehtest zu  $t_2$  (N=81).

Die *Verteilungen* im Hinblick auf die gemessenen Tests veranschaulichen zum 1-Minute-Up & Go-Test, dass 35 Prozent der Patienten ( $n=28$ ) die Strecke fünf bis achtmal, 64 Prozent ( $n=49$ ) neun bis 13 Mal und vier Personen mehr als 13 Mal zurücklegten (Abb. 10). Die Anzahl der Wegstrecken liegt im Mittel bei  $9,2 (\pm 2,2)$ , bei einem Maximum von 15 und einem Minimum von fünf Wegstrecken.

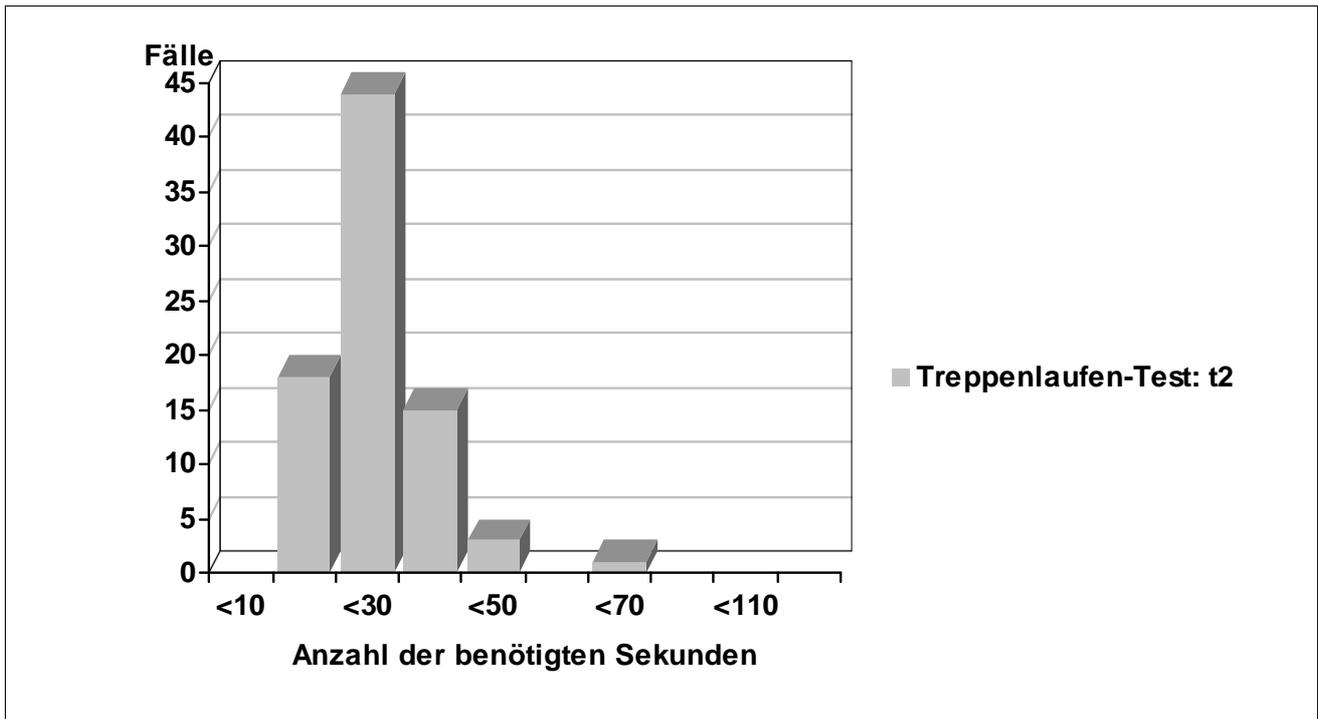


Abb. 12: Die Verteilungen zum Treppenlaufen-Test zu t<sub>2</sub> (N=81).

Im *2-Minuten-Gehtest* legten 5 Prozent der Teilnehmer (n=4) eine Strecke zwischen 56 und 110 Meter, 80 Prozent (n=65) zwischen 110 und 210 Metern und 16 Prozent (n=13) mehr als 210 Meter zurück (Abb. 11). Im Mittel betrug die Gehstrecke 172 Meter ( $\pm 37,6$ ).

Für den *Treppenlaufen-Test* benötigen 78 Prozent (n=63) zwischen 20 und 70 Sekunden, keiner dagegen mehr als 70 Sekunden (Abb. 12). 22 Prozent (n=18) bewältigten die beiden Treppenabsätze in weniger als 20 Sekunden. Im Durchschnitt benötigten Patienten 25 Sekunden ( $\pm 8,2$ ) für das Treppenlaufen bei einem Maximum von 133 Sekunden und einem Minimum von elf Sekunden. Einen Überblick über die deskriptive Statistik der gemessenen Tests gibt Tabelle 38.

Tab. 38. Die deskriptive Statistik zu den gemessenen Tests (t<sub>2</sub>).

N=81	Min	Max	M	s
<b>1-Minute-Up &amp; Go-Test<sup>1</sup></b>	5,00	15,00	9,19	2,23
<b>2-Minuten-Gehtest<sup>2</sup></b>	84,00	266,00	171,72	37,61
<b>Treppenlaufen-Test<sup>3</sup></b>	11,00	62,00	25,14	8,21
<b>2-Minuten-Lauftest<sup>2</sup> (N=4)</b>	0	340,00	12,74	57,26

Anm.: Angaben in <sup>1</sup>Wegstrecken; <sup>2</sup>Metern; <sup>3</sup>Sekunden.

Da der *Laufstest* nur von vier Personen durchgeführt werden konnte, wird an dieser Stelle auf eine graphische Darstellung verzichtet.

### 6.2.2 Deskriptive Statistik zu den Summenskalen

Fehlende Werte gegen Untersuchungsende traten zur Körperlichen Funktionsfähigkeit in 0,7 Prozent und zur Selbsteinschätzung in 1,7 Prozent der Fälle auf. Die Verteilungen der Summenskalen weisen für alle Instrumente für ein Viertel der Teilnehmer 71-80 Prozent der maximal erreichbaren Funktionskapazität auf (Abb. 13). Im Mittel werden 63-70 Prozent der Funktionskapazität auf den drei Skalen erreicht. Tabelle 39 zeigt die Mittelwerte und Standardabweichungen der Summenskalen.

Tab. 39. Die Mittelwerte und Standardabweichungen der Summenskalen ( $t_2$ ).

N=81	M	s
Polla	62,71	16,64
Selbsteinschätzung	70,23	17,50
Körperliche Funktionsfähigkeit	63,81	19,63

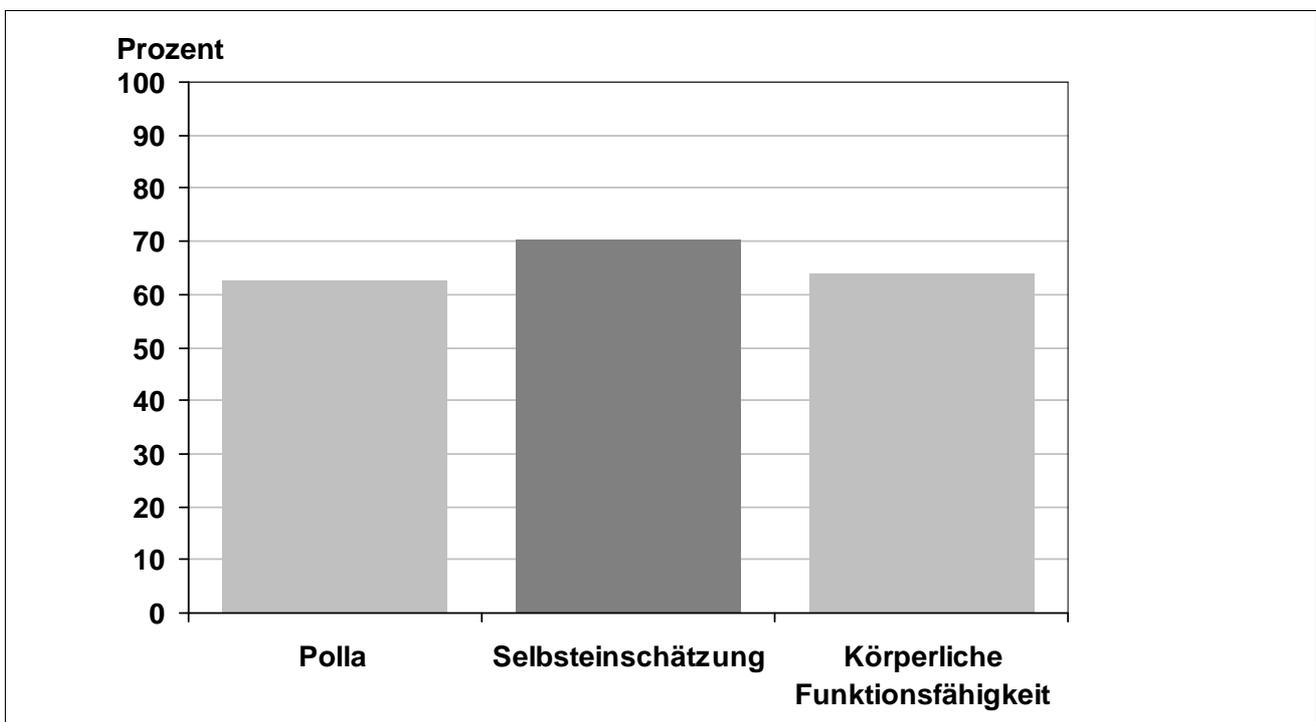


Abb. 13. Die Mittelwerte der Summenskalen zu  $t_2$ . (N=81).

### 6.2.3 Ergebnisse der spezifischen Zielkriterien

Gegen Therapieende beschrieben noch 11,8 Prozent (n=9 von 76), dass das zu Beginn angegebene *Hauptproblem* weiterhin ständig bestehe, 35,5 Prozent (n=27) gaben dagegen nun an, das Hauptproblem sei nicht mehr vorhanden. Täglich oder mehrmals tägliches Auftreten beschrieben 30,3 Prozent (n=23) und wöchentliches bzw. mehrmals wöchentliches Auftreten 22,4 Prozent (n=17) der Patienten. Zur Medikation (n=71) äußerten 8,4 Prozent eine regelmäßige Einnahme oder eine Einnahme nach Bedarf. 91,5 Prozent (n=65) nahmen keine Medikamente ein.

Ein *Giving-Way Phänomen* (n=73) beschrieben gegen Ende 4,1 Prozent (n=3) täglich und 36,9 Prozent (n=27) wöchentlich. Frei vom Giving-Way Phänomen waren dagegen 58,9 Prozent (n=43).

Hinsichtlich der *Schmerzdauer* (n=72) gaben 2,8 Prozent ständige Schmerzen, 13,9 Prozent (n=10) tägliche Schmerzen zwischen einer und mehreren Stunden, 58,4 Prozent (n=42) kurze bis zu einigen Minuten anhaltende und 25 Prozent (n=18) keine Schmerzen an. Die Mittelwerte zeigen für die Schmerzintensität 1,95, für die schmerzfreie Kniebeuge 96 Grad und die aktive Flexion 115 Grad (Tab. 40).

Tab. 40. Die deskriptive Statistik zur Schmerzintensität, Flexion und Kniebeuge zu t<sub>2</sub>.

	N	Min	Max	M	s
Schmerzintensität (NRS)	76	,00	8,00	1,95	1,65
Aktive Flexion in Grad	78	12,00	145,00	114,50	20,72
Schmerzfreie Kniebeuge in Grad	48	60,00	145,00	95,88	21,54

### 6.2.4 Deskriptive Statistik zur Vergleichsgruppe

Die Datenerhebung zum Polla und den gemessenen Tests in der Vergleichsgruppe von Gesunden konnte ohne fehlende Werte abgeschlossen werden. Lediglich zwei Personen wiesen je zwei Einschränkungen beim Strümpfe anziehen und Hocken auf; alle weiteren Aktivitäten waren ohne Einschränkungen möglich. Wie Tabelle 41 zeigt, benötigten die Personen der Vergleichsgruppe im Durchschnitt 15 Sekunden ( $\pm 2,7$ ) für das Laufen der beiden Treppenabsätze.

Tab.41. Deskriptive Statistik zu den gemessenen Tests der Vergleichsgruppe.

N=30	Min	Max	M	s
1-Minute-Up & Go Test (Angabe in Wegstrecken)	8,00	15,50	11,87	2,11
2-Minuten-Gehtest (Angabe in Metern)	160,00	252,00	220,43	22,06
Treppenlaufen-Test (Angabe in Sekunden)	11,50	21,41	14,82	2,66
2-Minuten-Lauftest (Angabe in Metern)	158,00	376,00	247,87	65,16

Die Anzahl der zurückgelegten Wegstrecken zwischen dem Hocker und der drei Meter entfernten Wand lag im Mittel bei zwölf ( $\pm 2,1$ ), die Gehstrecke zum 2-Minuten-Gehtest betrug 220 Meter ( $\pm 22,1$ ) sowie die Laufstrecke innerhalb von zwei Minuten 376 Meter ( $\pm 65,1$ ).

### 6.2.5 Zusammenhänge der spezifischen Zielkriterien

Einige der zu Beginn und gegen Ende der ambulanten Behandlung erhobenen spezifischen Zielkriterien sind zur Betrachtung von Zusammenhängen zu den Summenskalen des Polla, der Selbsteinschätzung und der Körperlichen Funktionsfähigkeit herangezogen worden.

Eine Analyse nach dem Korrelationskoeffizienten Spearman-Rho macht außer zur Selbsteinschätzung und zur Schmerzintensität (NRS) Zusammenhänge aller Skalen zur aktiven Flexion deutlich (Tab. 42).

Tab. 42. Die Zusammenhänge zu spezifischen Merkmalen nach Spearman-Rho.

MZP	Zielkriterium	N	Polla	KÖFU	SCHM <sup>2</sup>	Selbst <sup>1</sup>	Up & Go	Gehtest	Treppenlaufen
t <sub>1</sub>	SCHM	N=81	,260*	,426**	1	,370**	,174	,304**	-,315**
t <sub>2</sub>			,399**	,703**	1	,733**	,229*	,414**	-,428**
t <sub>1</sub>	NRS	N=76	-,322**	-,255*	-,608**	-,201	-,146	-,248*	,212
t <sub>2</sub>			-,433**	-,549**	-,669**	-,570**	-,286*	-,470**	,428**
t <sub>1</sub>	Flexion	N=78	,292*	,351**	,275*	,324**	,438**	,492**	-,476**
t <sub>2</sub>			,285*	,279*	,105	,206	,500**	,474**	-,432**
t <sub>1</sub>	Kniebeuge	N=48	,216	,188	,195	,103	-,068	,206	-,201
t <sub>2</sub>			,655**	,556**	,394**	,632**	,423**	,475**	-,507**
t <sub>1</sub>	Giving-Way	N=73	-,063	-,138	-,066	-,125	-,062	-,019	-,151
t <sub>2</sub>			,205*	,214*	-,228*	-,224*	-,040	,007	-,167

Anm.: <sup>1</sup>Selbsteinschätzung; <sup>2</sup>Schmerz; Signifikanzniveau: \*\*p ≤ 0.01; \*p ≤ 0.05.

Eine größere Beweglichkeit geht demnach mit einer höheren Punktschätzung im Polla und höheren Ergebniswerten bei den gemessenen Tests einher. Positive Korrelationen im Hinblick auf die NRS und den Treppenlaufen-Test bedeuten bei einer Verringerung der Schmerzintensität, dass sich die Zeit für die Bewältigung der beiden Treppenabsätze ebenfalls verringert.

Weiterhin stabile Zusammenhänge zeigen sich zur schmerzhaften Kniebeuge zu t<sub>2</sub>, da Patienten in den Knien zu Therapiebeginn nur wenig beugen konnten. Kaum Zusammenhänge zeigen sich dagegen zum Giving-Way Phänomen (Tab. 42).

Zusammenhänge zu den SF-36 Dimensionen weisen vereinzelt sowohl die aktive Flexion wie auch die schmerzhaft Kniebeuge zu Therapiebeginn, dagegen durchgehend die schmerzhaft Kniebeuge zum Therapieende auf (Tab. 43).

Tab. 43. Korrelationen zwischen der aktiven Flexion, der schmerzhaften Kniebeuge und den SF-36 Dimensionen.

N=81		SF-36 Skalen							
MZP	Kriterium	KÖFU <sup>1</sup>	KÖRO <sup>2</sup>	SCHM <sup>3</sup>	AGES <sup>4</sup>	VITA <sup>5</sup>	SOFU <sup>6</sup>	EMRO <sup>7</sup>	PSYC <sup>8</sup>
t <sub>1</sub>	Flexion	,351**	,350**	,275*	,027	,004	-,030	,224*	,099
t <sub>1</sub>	Kniebeuge	,188	-,134	,195	,438**	,481**	,020	,103	,357**
t <sub>2</sub>	Flexion	,279*	,182	,105	-,054	-,170	-,011	-,046	-,159
t <sub>2</sub>	Kniebeuge	,556**	,353*	,394**	,388**	,443**	,301*	,479**	,499**

Anm.: Signifikanzniveau: \*\*p ≤ 0.01; \*p ≤ 0.05; <sup>1</sup>Körperliche Funktionsfähigkeit, <sup>2</sup>Körperliche Rollenfunktion; <sup>3</sup>Schmerz, <sup>4</sup>Allgemeiner Gesundheitszustand; <sup>5</sup>Vitalität; <sup>6</sup>Soziale Funktionsfähigkeit, <sup>7</sup>Emotionale Rollenfunktion; <sup>8</sup>Psychisches Wohlbefinden.

### 6.2.6 Nebenwirkungen von Tests

Die Forderung zur Beschreibung von Testnebenwirkungen erheben Faller, Kohlmann Löschemann, Maurischat, Petermann, Schulz und Zwingmann (1999). Bei großen Nebenwirkungen eines Tests stellt sich die Frage nach seiner Sinnhaftigkeit. Patienten haben längere Leidenswege hinter sich gebracht, die nicht zusätzlich durch Beschwerden über Tests hervorgerufen oder verstärkt werden sollten, vor allem wenn diese über mehrere Tage andauern.

Insgesamt wurden 247 Tests zu t<sub>1</sub>, t<sub>2</sub>, an Gesunden, als Retest und zum Follow-up durchgeführt. Beschwerden in Form von Schmerz, Druck oder Ziehen wurde in 29 Fällen (6%) registriert (Tab. 44). Die häufigsten Beschwerden traten zu Therapiebeginn bei 14 Versuchspersonen auf (17,3% von n=81). Dies war allerdings auch während des Alltags ohne längere Nachwirkungen üblich. Kein Patient beschrieb aufgrund der Testdurchführungen länger anhaltende Beschwerden.

Tab. 44. Nebenwirkungen in Form von Schmerz, Druck oder Ziehen im Anschluss an Testdurchführung.

MZP	N	Häufigkeit	In Prozent von jeweiliger Stichprobengröße	In Prozent von 247 Tests
t <sub>1</sub>	81	14	17,28	5,67
t <sub>2</sub>	81	12	14,81	4,86
Retest	29	1	3,45	0,40
Gesunde	30	1	3,33	0,40
Follow-up	26	1	3,85	0,40
Insgesamt	247	29	-	2,35

## 6.3 Schließende Statistik

### 6.3.1 Überprüfungen auf Normalverteilung

Die Summenskalen der Messinstrumente wurden mit dem Kolmogorov-Smirnov-Anpassungstest auf *Normalverteilung* überprüft. Sowohl die gemessenen Tests (Ausnahme: Lauftest) als auch die Selbsteinschätzung zu Aktivitäten präsentieren zu beiden Zeitpunkten Normalverteilung. Die Polla Summenskala ist lediglich zum Messzeitpunkt Therapieende normal verteilt. Die SF-36 Skalen zeigen Normalverteilung für die *Körperliche Funktionsfähigkeit*, den Schmerz, die *Allgemeine Gesundheitswahrnehmung*, die *Vitalität* sowie die *Soziale Funktionsfähigkeit* zu Therapiebeginn und zu Therapieabschluss für die Dimensionen des Schmerzes, der *Gesundheitswahrnehmung* und der *Vitalität*.

In der *Vergleichsgruppe von Gesunden* zeigen die gemessenen Tests, die Selbsteinschätzung und die *Körperliche Funktionsfähigkeit*, die *Allgemeine Gesundheitswahrnehmung*, die *Vitalität* und das *Psychische Wohlbefinden* Normalverteilung auf. Der Polla dagegen ist nicht normal verteilt.

### 6.3.2 Veränderungen der Polla Items

Die *Veränderungen der ordinalskalierten Polla Items* („ja-eingeschränkt-nein“)  $t_2$  gegenüber  $t_1$  wurden mit dem parameterfreien Prüfverfahren für verbundene Stichproben nach Wilcoxon berechnet (Tab. 45).

Tab. 45. Die Veränderungen der Polla Items  $t_2$  gegenüber  $t_1$ .

N=81	Neg. Ränge	Pos. Ränge	Bindungen	Z <sup>1</sup>	Sig. <sup>2</sup>
Strümpfe anziehen	4	43	34	-5,60**	,000
Aufstehen vom Hocker	1	46	34	-6,56**	,000
Einbeinstand	0	56	25	-6,80**	,000
Aufheben	0	13	68	-3,61**	,000
Auf den Boden hinlegen	2	41	38	-5,60**	,000
Treppauf eine Stufe	1	58	22	-6,92**	,000
Hocken	1	29	51	-4,86**	,000
Hinknien	2	42	37	-5,66**	,000
Auf der Stelle laufen	2	59	20	-6,80**	,000
Einbeinhüpfen	0	15	66	-3,59**	,000
1-Miunte-Up & Go-Test	5	26	50	-3,77**	,000
2-Minuten-Gehtest	7	17	57	-2,04**	,041
Treppenlaufen-Test	1	28	52	-4,95**	,000
2-Minuten-Lauftest	-	-	-	-	-

Anm.: <sup>1</sup> Basiert auf negativen Rängen; <sup>2</sup> Asymptotische Signifikanz (2-seitig) nach Wilcoxon.

Demnach zeigen 13 von 14 Tests bedeutsame Veränderungen, wobei die größten das dreimalige Auf- und Absteigen auf einen 20cm hohen Stepp (F = -6,92; p < ,001), der Einbeinstand, das Laufen auf der Stelle (F = -6,80; p < ,001) sowie das Aufstehen vom Hocker (F = -6,56; p < ,001) aufzeigen. Die geringsten Veränderungen ergeben sich für den 2-Minuten-Gehtest (F = -2,04; p = ,04).

Die *Veränderungen der gemessenen Tests im Polla* wurden aufgrund der metrischen Datenerhebung über den T-Test für gepaarte Stichproben berechnet. Die Tabelle 46 zeigt die deskriptive Statistik zu den einzelnen Zeitpunkten. Demnach hat sich die Anzahl der zurückgelegten Wegstrecken gegen Therapieabschluss zum 1-Minute-Up & Go-Test im Mittel von 6,6 auf 9,2 erhöht, was einer Verbesserung im Mittel um 40 Prozent entspricht. Die Gehstrecke erhöhte sich im Durchschnitt von 132 auf 172 Meter (+30%) und die Zeit für das Treppensteigen verringerte sich im Mittel von 48 auf 25 Sekunden (-48%).

Tab. 46. Die deskriptive Statistik der gemessenen Tests zu den Zeitpunkten  $t_1$  und  $t_2$ .

N=81	Mzp	Min	Max	M	s
1-Minute-Up & Go-Test <sup>1</sup>	t <sub>1</sub>	3	13	6,6	1,98
	t <sub>2</sub>	5	15	9,2	2,23
2-Minuten Gehtest <sup>2</sup>	t <sub>1</sub>	56	240	131,7	35,54
	t <sub>2</sub>	84	266	171,7	37,61
Treppenlaufen-Test <sup>3</sup>	t <sub>1</sub>	14	133	48,4	24,78
	t <sub>2</sub>	11	62	25,1	8,21

Anm.: Angaben in: <sup>1</sup>Wegstrecken; <sup>2</sup>Metern; <sup>3</sup>Sekunden.

Der T-Test zeigt für alle drei gemessenen Performance-Tests signifikante Unterschiede zum Therapieende gegenüber dem Beginn, wobei die größten Veränderungen beim 1-Minute-Up & Go-Test und 2-Minuten-Gehtest auftreten (F = -14,32; p < ,001; Tab. 47).

Tab. 47. Die Ergebnisse des T-Tests auf Stichprobenunterschiede bei gepaarten Tests  $t_1/t_2$ .

N=81	Gepaarte Differenzen bei 95% KI <sup>1</sup>				T-Wert	df	Sig. <sup>2</sup>
	M	s	Untere	Obere			
1-Minute-Up & Go-Test	-2,59	1,63	-2,95	-2,23	<b>-14,32**</b>	80	,000
2-Minuten-Gehtest	-39,85	25,18	-45,42	-34,28	<b>-14,25**</b>	80	,000
Treppenlaufen-Test	23,22	21,93	18,37	28,07	<b>9,53**</b>	80	,000

Anm.: <sup>1</sup>95% Konfidenzintervall der Differenz, <sup>2</sup>Signifikanz 2-seitig.

Abbildung 14 veranschaulicht abschließend die Veränderungen im Vergleich zur Gruppe von Gesunden (Ges), ausgehend von den mittleren Prozentwerten der 0-100 transformierten Skalen.

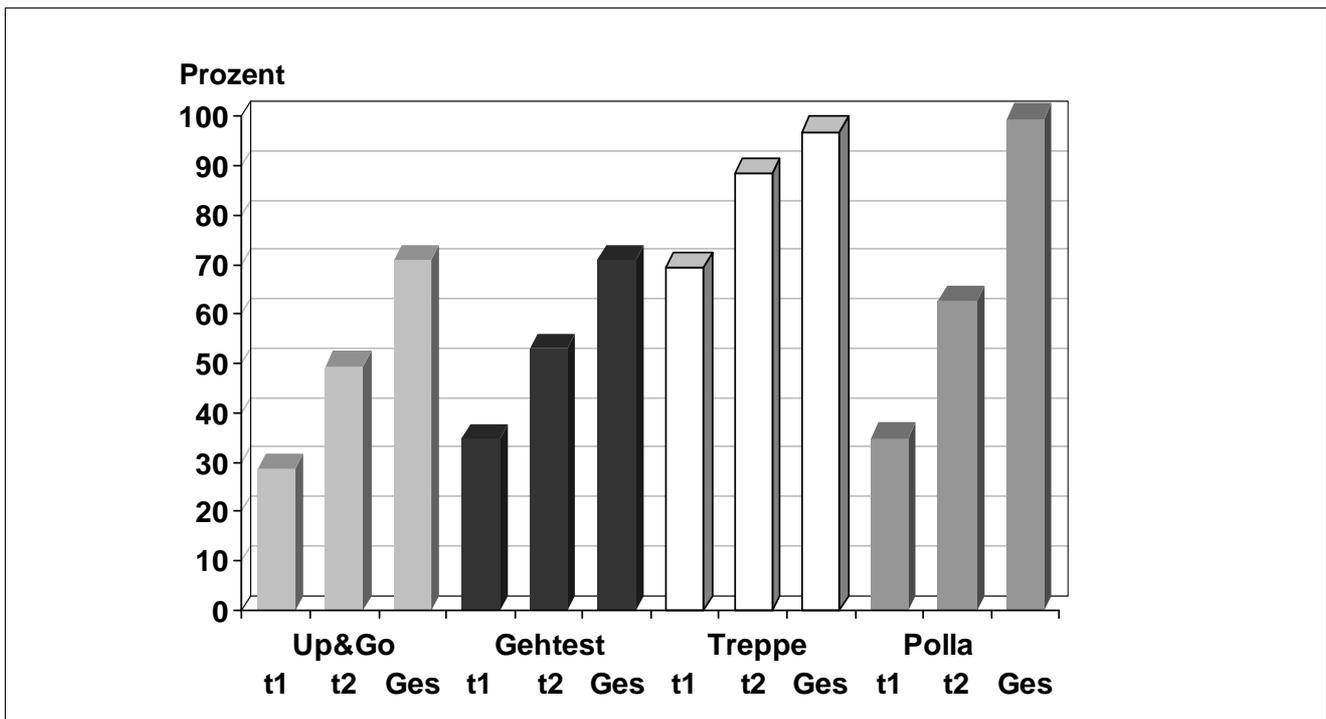


Abb. 14. Die Mittelwerte der Summenskalen zu  $t_1$  und  $t_2$  (N=81) im Vergleich zur Gruppe von Gesunden (Ges; N=30).

Es wird deutlich, dass die Ergebnisse der Patienten zum Therapieabschluss noch deutlich unter denen der Vergleichsgruppe liegen.

### 6.3.3 Veränderungen der SF-36-Skalen

Die Dimensionen der gesundheitsbezogenen Lebensqualität (SF-36): Die *Körperliche Funktionsfähigkeit* und *Rollenfunktion*, der *Schmerz*, der *Allgemeine Gesundheitszustand*, die *Vitalität*, die *Soziale Funktionsfähigkeit* und die *Emotionale Rollenfunktion* sowie das *Psychische Wohlbefinden* zeigen über die Mittelwerte der Prozentskalen gegen Therapieabschluss Veränderungen gegenüber dem Beginn in Richtung höhere Lebensqualität an (Tab. 48).

Abbildung 15 veranschaulicht die in Tabelle 48 dargestellten deskriptiven Werte. Dabei spiegelt die waagerechte Achse das Prozentniveau wider. Die Gruppenunterschiede jeder Summenskala zwischen  $t_1$  und  $t_2$  sind nach dem T-Test bei gepaarten Stichproben für intervallskalierte Daten berechnet und zeigen bedeutende Unterschiede ( $p \leq ,05$ ). Zum Vergleich ist das Ergebnis der Gruppe von Gesunden zur Körperlichen Funktionsfähigkeit in die Graphik eingebracht.

Tab. 48. Die deskriptive Statistik zu den SF-36 Skalen.

N=81	MZP	N	Min	Max	M	s
<b>KÖFU</b>	t <sub>1</sub>	N=81	0	88,89	36,38	19,19
	t <sub>2</sub>		20	100	63,81	19,63
<b>KÖRO</b>	t <sub>1</sub>	N=81	0	100	14,09	25,06
	t <sub>2</sub>		0	100	55,35	44,26
<b>SCHM</b>	t <sub>1</sub>	N=81	0	100	44,80	22,67
	t <sub>2</sub>		10	100	63,64	20,80
<b>AGES</b>	t <sub>1</sub>	N=80	10	97	63,79	18,23
	t <sub>2</sub>		20	97	67,84	17,21
<b>VITA</b>	t <sub>1</sub>	N=81	10	95	53,87	16,55
	t <sub>2</sub>		30	95	63,31	15,85
<b>SOFU</b>	t <sub>1</sub>	N=81	0	100	66,98	26,61
	t <sub>2</sub>		25	100	87,65	17,51
<b>EMRO</b>	t <sub>1</sub>	N=80	0	100	60,08	47,26
	t <sub>2</sub>		0	100	78,33	37,16
<b>PSYC</b>	t <sub>1</sub>	N=80	24	100	74,39	18,60
	t <sub>2</sub>		20	100	79,012	15,97

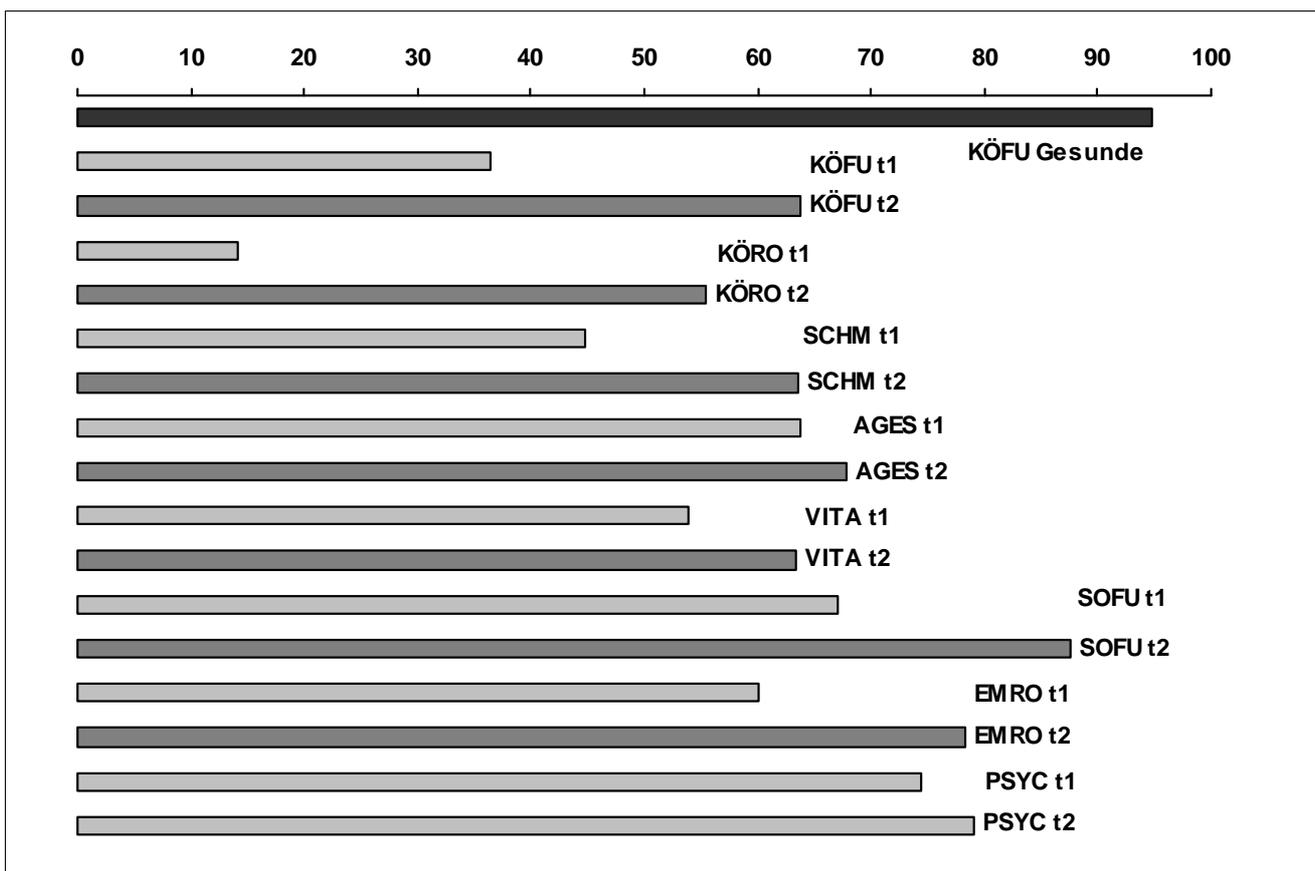


Abb. 15. Die Mittelwerte der SF-36 Skalen zu t<sub>1</sub> und t<sub>2</sub> (Angaben in Prozent). KÖFU: Körperliche Funktionsfähigkeit; KÖRO: Körperliche Rollenfunktion; SCHM: Schmerz, AGES: Allgemeiner Gesundheitszustand; VITA: Vitalität; SOFU: Soziale Funktionsfähigkeit; EMRO: Emotionale Rollenfunktion; PSYC: Psychisches Wohlbefinden (N=81).

### 6.3.4 Analyse von Unterschieden in Altersklassen

In der praktischen Untersuchungsdurchführung zeigten Patienten höheren Alters insgesamt langsamere Testausführungen als jüngere. Statistische Unterschiede aufgrund der Altersstreuung von 19 bis 76 Jahren waren daher denkbar. Eine Bildung von Altersklassen orientierte sich an der Voraussetzung vergleichbarer mittlerer Altersdifferenzen zwischen den Gruppen. Die Tabelle 49 zeigt drei Altersklassen mit einem mittleren Altersabstand von jeweils 20 Jahren, wobei die Fallzahlen unterschiedlich ausfallen.

Tab. 49. Die deskriptive Statistik in den Altersklassen und Verteilungen der Diagnosen.

N=81	N	Diagnosen nach Altersklassen				Geschlecht		Alter		
		KTE <sup>1</sup>	HTE <sup>2</sup>	Knie <sup>3</sup>	Knie-/ Usch <sup>4</sup>	w	m	Min	Max	M
19-38 Jahre	20	-	-	16	4	10	10	19	38	26,75
40-56 Jahre	15	2	1	9	3	7	8	40	56	46,80
58-76 Jahre	46	13	28	3	2	22	24	58	76	66,74

Anm.: <sup>1</sup>Knieendoprothesen, <sup>2</sup>Hüftendoprothesen, <sup>3</sup>Kniediagnosen, <sup>4</sup>Knie-/Unterschenkeldiagnosen.

Die Diagnoseverteilungen in den Altersklassen machen deutlich, dass in der Gruppe höheren Alters Hüft- und Knieendoprothesen überwiegen.

Als statistisches Verfahren für die Auswertung von Gruppenunterschieden diene nach dem allgemeinen linearen Modell die univariate Varianzanalyse mit dem festen Faktor Alter und den abhängigen Variablen *Polla* und *gemessene Tests* zu den Zeitpunkten Therapiebeginn und -abschluss.

#### 6.3.4.1 Ergebnisse der Varianzanalysen

Nahezu durchgängig signifikante Unterschiede trennen die jüngeren Patienten von den beiden Gruppen höheren Lebensalters gegen Therapieende, weniger dagegen zu Therapiebeginn, wobei die Unterscheidungen in Häufigkeit und Höhe besonders über den 1-Minute-Up & Go-Test und den 2-Minuten-Gehtest geleistet werden (Tab. 50). Zur Bestimmung der Größe signifikanter Unterschiede wird die Effektstärke *Eta* beschrieben und ergibt sich als Wurzel aus *Eta*-Quadrat. Die Beurteilung von *Eta* nach einer Varianzanalyse ergibt kleine Effekte bei Werten von 0,10, mittlere bei 0,25 und große Effekte bei 0,4 (Bortz & Döring, 2006, S. 606).

Mit überwiegend großen Effekten wird die *Gruppe der Jüngeren* (19-38-jährigen) von den Gruppen der Älteren (58-76-jährigen) zu beiden Zeitpunkten unterschieden (Tab. 50).

Tab. 50. Ergebnisse der Varianzanalysen zu Unterschieden zwischen Altersklassen (95% KI).

N=81	MZP	40-56 / 58-76			19-38 / 58-76			19-38 / 40-56		
		F	Sig. <sup>1</sup>	Eta	F	Sig. <sup>1</sup>	Eta	F	Sig. <sup>1</sup>	Eta
		,049	,825	,032	,984	,325	,122	1,32	,259	,195
<b>Polla</b>	<b>t<sub>1</sub></b>	11,40	<b>,001**</b>	<b>,402</b>	29,36	<b>,000**</b>	<b>,561</b>	2,85	,101	,281
<b>Up &amp; Go</b>		5,92	<b>,018**</b>	<b>,302</b>	17,14	<b>,000**</b>	<b>,459</b>	1,20	,280	,187
<b>Gehstest</b>		1,67	,201	,167	1,22	,274	,138	,066	,799	,045
<b>Treppe<sup>1</sup></b>										
<b>Polla</b>	<b>t<sub>2</sub></b>	5,10	<b>,028**</b>	<b>,281</b>	1,75	,191	,164	10,83	<b>,002**</b>	<b>,497</b>
<b>Up &amp; Go</b>		11,29	<b>,001**</b>	<b>,401</b>	74,71	<b>,000**</b>	<b>,734</b>	11,81	<b>,002**</b>	<b>,514</b>
<b>Gehstest</b>		3,22	,078	,228	47,56	<b>,000**</b>	<b>,653</b>	12,10	<b>,001**</b>	<b>,518</b>
<b>Treppe<sup>1</sup></b>		,049	,825	,032	12,56	<b>,001**</b>	<b>,405</b>	8,47	<b>,006**</b>	<b>,452</b>

Anm.: <sup>1</sup>Treppenlaufen-Test.

Eta-Quadrat beinhaltet den Anteil der Gesamtvarianz in der abhängigen Variable, die auf den jeweiligen Faktor zurückgeht (Willimczik, 1999) und ergibt sich als Quotient aus der Quadratsumme des abhängigen Faktors durch die Gesamtvarianz (gesamte Quadratsumme). Dabei leistet der Faktor Alter mittlere bis große Varianzaufklärungen (Bortz & Döring, 2006, S. 606, Werte um 0,01 = klein, 0,1 = mittel; 0,25 groß) von 16-54 Prozent an der abhängigen Variablen des 1-Minute-Up & Go-Test, 21-43 Prozent an der des 2-Minuten-Gehstests und 16-20 Prozent an der abhängigen Variablen des Treppenlaufen-Tests.

### 6.3.5 Analyse von Unterschieden in Diagnosegruppen

Eine Analyse von Unterschieden zu Diagnosegruppen erfolgte basierend auf der Annahme, dass sich nach verhältnismäßig leichteren operativen Eingriffen am Kniegelenk (z.B. Kreuzbandersatzplastik, Meniskusresektion) die Funktionsfähigkeit der unteren Extremität im Allgemeinen zügiger entwickelt als nach komplexeren Eingriffen (z.B. Knieendoprothesen). Die Gruppe der Kniediagnosen präsentierte eine große Streubreite von Krankheitsbildern, von denen einige Patienten Zusatzdiagnosen am betroffenen Gelenk aufwiesen, die aus physiotherapeutischer Erfahrung zeitliche Verzögerungen der funktionellen Entwicklung nach sich ziehen (Weinhold, 2002). Um Gruppenunterschiede herauszuarbeiten wurde die vierte Gruppe der Knie-/Unterschenkeldiagnosen bei der Überprüfung ausgeschlossen und nur 72 Patienten in die Auswertung aufgenommen.

Als statistisches Verfahren für die Auswertung von Gruppenunterschieden diente nach dem allgemeinen linearen Modell die univariate Varianzanalyse mit dem festen Faktor Diagnose und den abhängigen Variablen Polla und gemessene Tests zu den Zeitpunkten Therapiebeginn und -abschluss.

### 6.3.5.1 Merkmale der Stichprobe

Die Verteilungen von Alter und Geschlecht in den Diagnosegruppen zeigt Tabelle 51. Bei gleichmäßiger Geschlechtsverteilung weisen die Kniediagnosen ein mittleres Alter von 37 Jahren (Range: 52) auf. Die beiden anderen Gruppen zeigen im Vergleich einen Altersdurchschnitt von 65/67 Jahren bei einer Spannweite von 22/27 Jahren.

Tab. 51. Die deskriptive Statistik zu Alter und Geschlecht in den Diagnosegruppen.

N=72	N	Geschlecht		Alter			
		W	m	Min	Max	M	s
Kniediagnosen	28	15	13	19	71	36,61	14,12
Knieendoprothesen	15	7	8	54	76	64,93	6,05
Hüftendoprothesen	29	14	15	48	75	66,72	5,85

### 6.3.5.2 Ergebnisse der Varianzanalysen

Am deutlichsten wird die Gruppe der älteren Patienten mit Knieendoprothesen von der Gruppe der Kniegelenkdiagnosen mit großen Effekten über fast alle Tests unterschieden.

Die Gruppe mit Hüftendoprothesen wird über drei Tests von der gleichaltrigen Gruppe der Knieendoprothesen über große Effekte unterschieden (Tab. 52). Zusätzlich wird sie ebenfalls durch drei Verfahren über große Effekte von der durchschnittlich 20 Jahre jüngeren Gruppe mit Kniegelenksdiagnosen unterschieden (Tab. 52).

Durchgängige Unterschiede zwischen allen Diagnosegruppen zu  $t_1$  und  $t_2$  zeigen lediglich der 1-Minute-Up & Go-Test sowie der 2-Minuten-Gehtest (Tab. 52).

Tab. 52. Ergebnisse der Varianzanalysen zu Unterschieden zwischen Diagnosegruppen (95% KI).

N=72		KTE / Knie			HTE / KTE			HTE / Knie		
MZP		F	Sig. <sup>1</sup>	Eta	F	Sig. <sup>1</sup>	Eta	F	Sig. <sup>1</sup>	Eta
Polla	$t_1$	2,97	,092	,261	2,66	,110	,245	,033	,857	,032
Up & Go		12,10	<b>,001**</b>	<b>,477</b>	,457	,503	,105	15,31	<b>,000**</b>	<b>,467</b>
Gehtest		5,72	<b>,021*</b>	<b>,349</b>	2,44	,126	,235	2,98	,090	,226
Treppe <sup>2</sup>		4,91	<b>,032*</b>	<b>,327</b>	9,07	<b>,004**</b>	<b>,422</b>	,290	,593	,071
Polla	$t_2$	8,91	<b>,005**</b>	<b>,423</b>	26,41	<b>,000**</b>	<b>,621</b>	3,90	,053	,257
Up & Go		22,48	<b>,000**</b>	<b>,595</b>	3,46	,070	,276	20,57	<b>,000**</b>	<b>,522</b>
Gehtest		15,70	<b>,000**</b>	<b>,526</b>	3,87	,056	,290	11,60	<b>,001**</b>	<b>,417</b>
Treppe <sup>1</sup>		10,97	<b>,002**</b>	<b>,459</b>	9,12	<b>,004**</b>	<b>,422</b>	,992	,324	,134

Anm.: <sup>1</sup>Treppenlaufen-Test.

Die großen Varianzaufklärungen im Hinblick auf die Gruppen der Knieendoprothesen und Kniediagnosen liegen bei 23 und 35 Prozent für den 1-Minute-Up & Go-Test, bei 12 und 28 Prozent für den 2-Minuten-Gehtests sowie 11 und 21 Prozent für den Treppenlaufen-Test.

Für eine spätere Diskussion zu Unterschieden von Diagnosegruppen ist erwähnenswert, dass die gleichaltrigen Gruppen der Hüft- und Knieendoprothesen nur über den Treppenlaufen-Test und den Polla unterschieden werden. Die erklärten Varianzanteile bezüglich des Treppenlaufen-Tests liegen bei 18 Prozent zu beiden Zeitpunkten und bei 39 Prozent für den Polla zu Therapieende.

### 6.3.6 Veränderungen der spezifischen Zielkriterien

Die drei spezifischen Zielkriterien *Schmerzintensität*, *aktive Flexion* gemessen in Rückenlage sowie die *schmerzfremie Kniebeuge* gemessen am Kniegelenk in Grad sind auf Unterschiede zu Therapieende gegenüber dem Untersuchungsbeginn mit T-Test für gepaarte Stichprobenunterschiede überprüft worden. Bedeutsame Veränderungen weisen alle drei Kriterien auf (Tab. 53). Die Schmerzintensität reduzierte sich im Mittel von 3,8 auf zwei zum Therapieende ( $T = 8,61$ ), die aktive Flexion verbesserte sich von 94 auf 115 Grad ( $T = -13,84$ ) und die schmerzfremie Kniebeuge von 63 auf 96 Grad ( $T = -8,89$ ).

Tab. 53. Die Veränderungen der spezifischen Zielkriterien  $t_1$  gegenüber  $t_2$  nach dem T-Test.

T-Test	N	Gepaarte Differenzen mit 95% KI				t	Sig.
		M	s	Untere	Obere		
Schmerzintensität (NRS)	N=76	1,87	1,89	1,44	2,30	<b>8,61</b>	,000
Aktive Flexion	N=78	-20,33	12,98	-23,26	-17,40	<b>-13,84</b>	,000
Kniebeuge	N=48	32,83	25,58	-40,26	-25,40	<b>-8,89</b>	,000

Bei der Frage nach den Größenordnungen von Veränderungen im Hinblick auf Schmerz und Beweglichkeit wurden die Effekte von Unterschieden über den Standardized Response Mean ermittelt (SRM, Leonhardt, 2004). Die Klassifikation der Effekte orientiert sich an den Empfehlungen von Cohen (1988, S. 26 f.). Danach werden Werte von 0,2 als kleine, von 0,5 als mittlere und von 0,8 als große Effektstärken beschrieben. Für alle drei Parameter zeigen sich große Effekte, die für die aktive Flexion bei 0,99, für die NRS bei 1,0. und für die Kniebeuge bei 1,21 liegen.

Der Wilcoxon-Test als nichtparametrisches Verfahren wurde zur Überprüfung von Gruppenunterschieden auf rangskaliertem Niveau für das Giving-Way Phänomen,

die Häufigkeit der Medikation, die Schmerzdauer sowie das Hauptproblem in seiner Häufigkeit angewandt. Signifikante Veränderungen zum Ende gegenüber dem Therapieanfang werden bei allen Kriterien deutlich (Tab. 54).

Tab. 54. Die Veränderungen der spezifischen Zielkriterien  $t_2$  gegenüber  $t_1$  nach Wilcoxon.

Wilcoxon-Test	N	Neg. Ränge	Pos. Ränge	Bindungen	Z <sup>2</sup>	Sig. <sup>1</sup>
Hauptproblem	N=76	0	61	15	<b>-6,82</b>	,000
Medikation	N=71	4	45	22	<b>-5,17</b>	,000
Schmerzdauer	N=72	5	46	21	<b>-5,3</b>	,000
Giving-Way	N=73	5	16	52	<b>-2,08</b>	,037

Anm.: <sup>1</sup>Asymptotische Signifikanz (2-seitig), <sup>2</sup>Wilcoxon-Test, basiert auf negativen Rängen.

## 6.4 Überprüfung von Testgütekriterien zum Polla zu $t_1$ und $t_2$

### 6.4.1 Objektivität (Interrater-Reliabilität)

Die Ergebnisse der Interrater-Reliabilität zum Polla zum Messzeitpunkt Therapieabschluss deuten außer beim Einbeinstand auf überwiegend sehr gute Werte hin (Tab. 55). Im arithmetischen Mittel der Kappa-Werte zeigt sich eine gute Übereinstimmung von ,747. Die schlechten Ergebnisse zum Einbeinstand resultieren aus einem Standardisierungsproblem.

Tab. 55. Das Kappa-Maß zu den drei Interratern (it) zu den Polla Items.

N=32	it <sub>1</sub> /it <sub>2</sub>	It <sub>1</sub> /it <sub>3</sub>	it <sub>2</sub> /it <sub>3</sub>	Mittelwerte
Strümpfe anziehen	,688	,750	,803	,747
Aufstehen v. Hocker	,803	,680	,871	,785
Einbeinstand	,753	-	-	,753
Hinlegen/Aufstehen	1	1	1	1
Treppauf eine Stufe	,817	,822	,881	,84
Hocken	,824	,885	,707	,805
Hinknien	,950	,950	1	,967
Laufen auf der Stelle	,573	,583	,581	,579
Einbeinhüpfen	,862	,862	1	,608
1-Minute-Up & Go-Test	,614	,649	,810	,691
2-Minuten-Gehtest	,500	,333	,500	,444
Treppenlaufen-Test	,811	,680	,747	,746
<b>Gesamt (gemittelt)</b>	<b>,766</b>	<b>,745</b>	<b>,727</b>	<b>,747</b>

Die Beurteilung der Standsicherheit sollte im Seitenvergleich der Beine erfolgen, was nicht immer ganz einfach war und häufig zu differierenden Ansichten der Therapeuten und damit zu unterschiedlichen Ergebnissen führte. In der Korrektur der Standardisierung wird demnach auch nicht mehr die Standsicherheit beurteilt, sondern lediglich ob der Einbeinstand für 10 Sekunden möglich ist (ANHANG 8).

Insgesamt zeigt der Polla gute Ergebnisse hinsichtlich Interrater-Reliabilität. Allerdings ist die Anzahl von drei Beurteilern gering und sollte zusätzlich geprüft werden. Auch bei einer vorliegenden standardisierten Ausführungsbeschreibung werden Beurteiler in unterschiedlichen Einrichtungen immer wieder zu differierenden Ergebnissen kommen. Ein intensiver Erfahrungsaustausch über die Durchführungsweise der Tests lässt unter Therapeuten in der jeweiligen Einrichtung möglicherweise spezifische Interpretationen der Ergebnisse entstehen.

## 6.4.2 Reliabilität

### 6.4.2.1 Intrarater-Reliabilität

Die Ergebnisse zur *Intrarater-Reliabilität* – ein und derselbe Beurteiler überprüft zu zwei Zeitpunkten die drei gemessenen Tests an einem Patienten – zeigen über den Intra-Class-Korrelationskoeffizienten (ICC) für Intervallskalierungen Werte von 0,98. Nach den Kriterien zur Beurteilung von Objektivitätskoeffizienten ergibt sich nach Clarke (1976, S. 27) eine ausgezeichnete Übereinstimmung (Tab. 56).

Tab. 56. Korrelationskoeffizient in Klassen ( $t_2$ /Retest; durchschnittliche Maße).

N=29	$\alpha$	Korrelation innerhalb der Klasse	95%-KI		F-Test mit wahrem Wert 0			
			Untergrenze	Obergrenze	Wert	Freiheitsgrade 1	Freiheitsgrade 2	Sig.
<b>1-Min.-Up &amp; Go</b>	,980	,981	,959	,991	<b>51,43</b>	28	29	,000
<b>2-Min.-Gehtest</b>	,982	,980	,958	,991	<b>50,55</b>	28	29	,000
<b>Treppenlaufen</b>	,981	,973	,943	,987	<b>36,77</b>	28	29	,000

Anm.: Modell mit Ein-Weg-Zufallseffekten bei dem die Personeneffekte zufällig sind.

Schoppen, Boonstra, Groothoff, de Vries, Goeken und Eisma (1999) zeigten ebenso wie Podsiadlo und Richardson (1991) für den Timed-Up & Go-Test eine gute Intrarater- und Interrater-Reliabilität im zwei Wochenabstand. Podsiadlo und Richardson (1991) wiesen eine Intrarater-Reliabilität an 20 Patienten im Abstand von 2 Monaten mit einem ICC von 0,99 und eine Reliabilität zwischen drei Interratern an 22 Patienten an einem Tag mit einem ICC von 0,98 nach.

Von der Beurteilung der Intrarater-Reliabilität im Hinblick auf den ordinalskalierten Polla wurde aufgrund möglicher Erinnerungseffekte und nicht geplanter Verblindung der Therapeuten abgesehen.

### 6.4.2.2 Konsistenzanalyse

Die Ergebnisse der Konsistenzanalyse über Cronbachs Alpha zeigen für den Polla nach den Beurteilungskriterien von Weise (1975, S. 219) zu beiden Zeitpunkten eine mittlere Reliabilität um 0,8 und eine hohe Reliabilität von über 0,9 für die Selbsteinschätzung (Tab. 57).

Unterschiede haben sich auch hinsichtlich der Zusammenhänge zwischen Selbsteinschätzungsinstrumenten und Performance-Tests ergeben (s. Kap. 6.4.4).

Tab. 57. Interne Konsistenz zum Polla und zur Selbsteinschätzung.

N=81	Messzeitpunkt	Cronbachs Alpha
Polla	t <sub>1</sub>	0,79
	t <sub>2</sub>	0,81
Selbsteinschätzung	t <sub>1</sub>	0,91
	t <sub>2</sub>	0,94

### 6.4.3 Konstruktvalidität

Die Konstruktvalidität klärt die Beziehungen zwischen den in einem Messinstrument gemessenen Verhaltensweisen und einem zu erfassenden Merkmal (= Konstrukt). In der vorliegenden Untersuchung wird das Konstrukt der rangskalierten Datenaufnahme zum Polla („ja-eingeschränkt-nein“) für den Up & Go-Test, den Gehstest und den Treppenlaufen-Test mit dem Konstrukt den inhaltlich entsprechenden Items der gemessenen Tests zu t<sub>1</sub> und t<sub>2</sub> unter korrelativen Vergleichen betrachtet.

Ein weiterer Zusammenhang auf der Grundlage des Korrelationskoeffizienten Spearman-Rho erfolgt zwischen der Polla Summenskala und den einzelnen gemessenen Tests sowie zu ihrer gemeinsamen Summenskala. Die Ergebnisse zeigen Korrelationen auf Itemebene lediglich für den gemessenen Treppenlaufen-Test mit dem entsprechenden Polla Item („ja-eingeschränkt-nein“, Tab. 58).

Überwiegend mittlere Zusammenhänge weist jedoch die Summenskala zum Polla zu beiden Zeitpunkten zu den drei gemessenen Tests wie auch zu ihrer Summenskala auf ( $r = 0,415^{**} - 0,707^{**}$ ; Tab. 58).

Tab. 58. Korrelationen nach Spearman-Rho im Patientenkollektiv (n=81).

MZP	Polla Items („ja-eingeschränkt-nein“)	Gemessene Tests <sup>2</sup>	Korrelation	Sig. <sup>1</sup>
t <sub>1</sub>	Up & Go-Test	1-Minute-Up & Go-Test	,143	,201
	Gehtest	2 Minuten Gehtest	,129	,250
	Treppenlaufen-Test	Treppenlaufen-Test	<b>,353</b>	<b>,001</b>
t <sub>2</sub>	Up & Go-Test	1-Minute-Up & Go-Test	,135	,228
	Gehtest	2 Minuten Gehtest	,109	,331
	Treppenlaufen-Test	Treppenlaufen-Test	<b>,477</b>	<b>,000</b>
t <sub>1</sub>	Polla Summenskala	1-Minute-Up & Go-Test	<b>,534</b>	<b>,000</b>
		2-Minuten Gehtest	<b>,579</b>	<b>,000</b>
		Treppenlaufen-Test	<b>,645</b>	<b>,000</b>
t <sub>2</sub>	Polla Summenskala	1-Minute-Up & Go-Test	<b>,415</b>	<b>,000</b>
		2 Minuten Gehtest	<b>,479</b>	<b>,000</b>
		Treppenlaufen-Test	<b>,707</b>	<b>,000</b>
t <sub>1</sub>	Polla Summenskala	Summenskala (gemessene Tests)	<b>,629</b>	<b>,000</b>
t <sub>2</sub>	Polla Summenskala	Summenskala (gemessene Tests)	<b>,506</b>	<b>,000</b>

Anm.: <sup>1</sup>Signifikanzniveau (zweiseitig); <sup>2</sup>angewandt: 0-100 Skalen.

#### 6.4.4 Kriteriumsvalidität (Korrelationshypothese)

Die Bestimmung der Kriteriumsvalidität als Gradmesser für die Übereinstimmung der Testergebnisse des Polla und der Selbsteinschätzung zu Aktivitäten mit dem Außenkriterium des SF-36 erfolgt über korrelative Vergleiche nach dem Korrelationskoeffizienten Spearman-Rho und dem Produkt- Moment- Korrelationskoeffizienten nach Pearson zu den Messzeitpunkten Therapiebeginn und Therapieabschluss.

Die Ergebnisse weisen für den *Polla* Zusammenhänge zu fast allen *SF-36 Dimensionen* zu beiden Messzeitpunkten auf (Tab. 59). Weniger Zusammenhänge weisen der Treppenlaufen-Test mit zehn, der Gehtest mit neun und der 1-Minute-Up & Go-Test mit fünf Korrelationen auf. Die *Selbsteinschätzung* zu Aktivitäten zeigt gegenüber allen SF-36 Dimensionen höhere und häufigere Korrelationen als der Polla auf. Sie weist vergleichbare Zusammenhänge wie die Körperliche Funktionsfähigkeit (KÖFU) zu der Mehrzahl der SF-36 Skalen auf. Die Ergebnisse der *dimensionsähnlichen Aktivitätsebene* zeigen zwischen der KÖFU einerseits, dem Polla und den gemessenen Tests andererseits, moderate Zusammenhänge zu Therapiebeginn, die gegen Therapieende zum Polla und zu der Selbsteinschätzung ansteigen und bei den gemessenen Tests geringer ausfallen. Eine Ausnahme bildet hier die mittlere Korrelation zum Treppenlaufen-Test (Tab. 59).

Tab. 59. Die Korrelationen zwischen den Messinstrumenten zu Therapiebeginn.

N=81		Pearson	Spearman-Rho				
MZP	Dimension	KÖFU	Polla	Selbst <sup>9</sup>	Up & go	Gehtest	Treppenlaufen
t <sub>1</sub>	KÖFU <sup>1</sup>	1	,547**	,783**	,539**	,549**	,581**
	KÖRO <sup>2</sup>	,515**	,394**	,440**	,437**	,425**	,447**
	SCHM <sup>3</sup>	,415**	,243*	,370**	,174	,304**	,315**
	AGES <sup>4</sup> N=80	,163	,246*	,218	,095	,097	,183
	VITA <sup>5</sup>	,279*	,180	,177	-,038	,064	,066
	SOFU <sup>6</sup>	,392**	,258*	,470**	,156	,087	,230*
	EMRO <sup>7</sup>	,202	,065	,231*	,352**	,234*	,233*
	PSYC <sup>8</sup> N=80	,247*	,247*	,250*	,127	,037	,146
t <sub>2</sub>	KÖFU	1	,621**	,820**	,282*	,372**	,511**
	KÖRO	,644**	,387**	,596**	,159	,266*	,332**
	SCHM	,697**	,399**	,733**	,229*	,414**	,428**
	AGES	,357**	,269*	,368**	,087	,143	,178
	VITA	,423**	,238*	,366**	-,021	,74	,121
	SOFU	,431**	,260*	,497**	,202	,258*	,264*
	EMRO N=80	,292**	,260*	,317**	,071	,279*	,242*
	PSYC	,319**	,213	,379**	,034	,126	,110

Anm.: Berechnungsgrundlage: 0-100 transformierte Skalen; Signifikanzniveau (zweiseitig): \*\*p ≤ 0,01; \*p ≤ 0,05; <sup>1</sup>Körperliche Funktionsfähigkeit, <sup>2</sup>Körperliche Rollenfunktion, <sup>3</sup>Schmerz, <sup>4</sup>Allgemeiner Gesundheitszustand, <sup>5</sup>Vitalität, <sup>6</sup>Soziale Funktionsfähigkeit, <sup>7</sup>Emotionale Rollenfunktion, <sup>8</sup>Psychisches Wohlbefinden, <sup>9</sup>Selbsteinschätzung.

Zur Körperlichen Rollenfunktion (KÖRO) und zum Schmerz (SCHM) finden sich durchweg Zusammenhänge für den Polla, das Treppenlaufen und den Gehtest.

Zu den wenig verwandten Dimensionen wie der Vitalität (VITA) mit Fragen zu Schwung, Energie, Erschöpfung und Müdigkeit oder dem Psychischen Wohlbefinden (PSYC) weist nur der Polla geringe und instabile Zusammenhänge auf, stabil zeigen sich dagegen Zusammenhänge zum Allgemeinen Gesundheitszustand auf geringem Niveau (t<sub>1</sub>: r = 0,246\*, t<sub>2</sub>: r = 0,269\*, Tab. 59).

Von den gemessenen Tests hebt sich der Treppenlaufen-Test durch die Häufigkeit und Stärke von Zusammenhängen zur Körperlichen Rollenfunktion, zum Schmerz, zur Sozialen Funktionsfähigkeit als auch zur Emotionalen Rollenfunktion gegenüber den beiden anderen Tests hervor. Das Treppenlaufen korreliert zwar nicht so häufig mit den Dimensionen des SF-36, dafür aber höher als der Polla, insbesondere bei Fragen nach Schmerzen, danach wie körperliche bzw. seelische Schwierigkeiten im Alltag bewältigt werden und bei Fragen nach der Beeinträchtigung von Kontakten (SOFU).

### 6.4.5 Unterschiede in der Parallelisierung

Es wird davon ausgegangen, dass Patienten gegen Therapieende im Performance-Test bessere Ergebnisse zeigen und sich das Aktivitätsniveau an das von Gesunden annähert. Unter diesem Gesichtspunkt ist von Interesse, ob die Tests zu den verschiedenen Zeitpunkten und insbesondere zum Follow-up in der Lage sind, Patienten von Gesunden zu unterscheiden.

Tab. 60. Die Unterschiede in den Gruppen „Gesunde“ und „Patienten“ (Varianzanalyse, 95% KI).

MZP		Par <sup>2</sup>	N	M	s	df	Mittel der Quadrate	F	Sig. <sup>1</sup>	Eta
t <sub>1</sub>	Polla	Pat	30	36,11	13,94	1	60166	<b>612,57</b>	,000	<b>,956</b>
		Ges	30	99,44	1,44					
	Up & Go	Pat	30	6,90	2,25	1	370,02	<b>77,71</b>	,000	<b>,757</b>
		Ges	30	11,87	2,11					
	Gehetest	Pat	30	136,60	35,33	1	105420	<b>121,51</b>	,000	<b>,823</b>
		Ges	30	220,43	22,07					
	Treppe	Pat	30	45,65	20,53	1	14258	<b>66,53</b>	,000	<b>,731</b>
		Ges	30	14,82	2,66					
	Lauftest	Pat	30	,00	,00	1	921568	<b>434,10</b>	,000	<b>,939</b>
		Ges	30	247,87	65,16					
t <sub>2</sub>	Polla	Pat	30	64,03	15,76	1	18815	<b>150,22</b>	,000	<b>,849</b>
		Ges	30	99,44	1,44					
	Up & Go	Pat	30	9,87	2,60	1	63,04	<b>11,24</b>	,001	<b>,402</b>
		Ges	30	11,87	2,11					
	Gehetest	Pat	30	179,53	45,66	1	25092	<b>19,51</b>	,000	<b>,502</b>
		Ges	30	220,43	22,07					
	Treppe	Pat	30	24,33	10,03	1	1356	<b>25,19</b>	,000	<b>,550</b>
		Ges	30	14,82	2,66					
	Lauftest	Pat	30	26,80	83,49	1	733057	<b>130,71</b>	,000	<b>,832</b>
		Ges	30	247,87	65,16					

Anm.: <sup>1</sup>Signifikanzniveau (zweiseitig); <sup>2</sup>Parallelisierte Gruppen.

Die Ergebnisse univariater Varianzanalysen nach dem allgemeinen linearen Modell mit dem festen Faktor „Patient“ und „Gesunder“ sowie den abhängigen Variablen des Polla und der gemessenen Tests weisen große Effekte von Unterschieden von Eta über 0,4 für alle Instrumente zu Therapiebeginn und -abschluss nach (Tab. 60).

Einschränkend muss erwähnt werden, dass die Probanden der Vergleichsgruppe von Gesunden vorwiegend aus der Klientel rekrutiert wurde, welche das private Bewegungsangebot des *Ambulanten Rehasentrums Kiel* nutzt. Es kann daher von einer insgesamt sportlichen Gruppe ausgegangen werden, die nicht repräsentativ für diese Altersgruppe erscheint.

### 6.4.6 Änderungssensitivität (Unterschiedshypothese)

Die Veränderungen von Therapieergebnissen zwischen verschiedenen Messzeitpunkten werden über die Änderungssensitivität bestimmt. Ihre Betrachtung erfolgt zwischen Therapiebeginn und -abschluss für die Gesamtstichprobe (n=81) über den Standardized Response Mean (SRM, Kap. 6.3.6).

Die Änderungssensitivität für die Gesamtstichprobe zeigt für alle Instrumente große Effekte von größer als 0,8 zwischen  $t_1$  und  $t_2$ . Besonders große Effekte zeigt der Polla (1,79, Tab. 61), etwas geringere die *Körperliche Funktionsfähigkeit* und die Selbsteinschätzung (1,41).

Tab. 61. Die Änderungssensitivität über die SRM in der Gesamtstichprobe.

N=81	MZP	M	S	SRM $t_1/t_2$
<b>KÖFU</b>	$t_1$	36,38	19,19	1,41
	$t_2$	63,81	19,63	
<b>Polla</b>	$t_1$	34,83	14,38	1,79
	$t_2$	62,71	16,64	
<b>Selbsteinschätzung</b>	$t_1$	46,19	16,56	1,41
	$t_2$	70,23	17,50	
<b>1-Minute-Up &amp; Go-Test</b>	$t_1$	28,77	15,81	1,23
	$t_2$	49,49	17,84	
<b>2-Minuten-Gehtest</b>	$t_1$	34,80	16,30	1,11
	$t_2$	53,08	17,25	
<b>Treppenlaufen-Test</b>	$t_1$	69,37	20,32	1,26
	$t_2$	88,41	6,73	

Anm.: Grundlage der Berechnungen sind die 0-100 transformierten Skalen.

### 6.5 Auswertungen zum Follow-up mit 26 Patienten

Als sich während der Durchführung eine über den Untersuchungszeitraum hinaus währende Behandlung für einen Teil von Patienten abzeichnete, wurde nachträglich ein Follow-up in das Studiendesign aufgenommen. Von Interesse war die Frage, ob der Polla über den Beobachtungszeitraum von sechs Wochen hinaus Veränderungen des Aktivitätsniveaus von Patienten abzubilden vermag. Die Zeitspanne zum Follow-up war mit durchschnittlich 80 Tagen ( $\pm 39$ , Spannweite: 166) nach Therapieende doppelt so lang wie zwischen Therapiebeginn und -abschluss mit 42 Tagen ( $\pm 3,4$ ). Betrachtung finden an dieser Stelle der Polla, die Selbsteinschätzung zu Aktivitäten sowie die Körperliche Funktionsfähigkeit (SF-36).

### 6.5.1 Deskriptive Statistik zum Follow-up

In die Auswertung kamen jeweils 13 männliche und weibliche Versuchspersonen mit einem Durchschnittsalter von 51 Jahren ( $\pm 18$ ) und einem BMI von 29 ( $\pm 6$ ). Die Datenerhebung zum Polla konnte ohne Fehlwerte abgeschlossen werden. Fehlende Werte traten dagegen zur *Körperlichen Funktionsfähigkeit* und zur Selbsteinschätzung in 0,2-0,3 Prozent der Fälle auf.

Die Aktivitäten des Hockens, Hinkniens sowie des Einbeinhüpfens waren von acht bis elf Personen und der Lauftest von 19 Personen (73%) nicht durchführbar – dagegen waren acht Aktivitäten von allen 26 Personen zumindest eingeschränkt auszuführen (Tab. 62).

Tab. 62. Die Häufigkeiten der Polla Items (Follow-up).

N=26	Gültige Prozente		
	Nein	Eingeschränkt	Ja
Strümpfe anziehen	0	19,2	80,8
Aufstehen vom Hocker	0	19,2	80,8
Einbeinstand	11,5	7,7	80,8
Aufheben	0	0	100
Hinlegen/Aufstehen	0	30,8	69,2
Treppauf eine Stufe	3,8	23,1	73,1
Hocken	38,5	42,3	19,2
Hinknien	30,8	19,2	50,0
Laufen auf der Stelle	11,5	7,7	80,8
Einbeinhüpfen	42,3	15,4	42,3
Up & Go-Test	0	15,4	84,6
2-Minuten-Gehtest	0	19,2	80,8
Treppenlaufen-Test	0	19,2	80,8
2-Minuten-Lauftest	73,1	0	26,9

Bezüglich der drei gemessenen Tests erreichten 23 Patienten (88%) im Treppenlaufen-Test, dagegen nur je vier Patienten im Geh- und 1-Minute-Up & Go-Test mehr als 81 Prozent des auf der jeweiligen Skala erreichbaren Wertes.

Tab. 63. Mittelwerte und Standardabweichungen zu den Summenskalen (Follow-up).

N=26	M	s
1-Minute-Up & Go-Test	56,15	17,91
2-Minuten-Gehtest	59,74	19,88
Treppenlaufen-Test	92,07	6,19
Polla	78,53	20,61
Selbsteinschätzung	83,29	17,13
Körperliche Funktionsfähigkeit	77,83	17,79

17-20 Patienten (65-77%) erreichten beim Polla, der Selbsteinschätzung und der KÖFU Werte von mehr als 81 Prozent der möglichen Funktionskapazität. Tabelle 63 und Abbildung 16 zeigen die Mittelwerte und Standardabweichungen zu den Summenskalen.

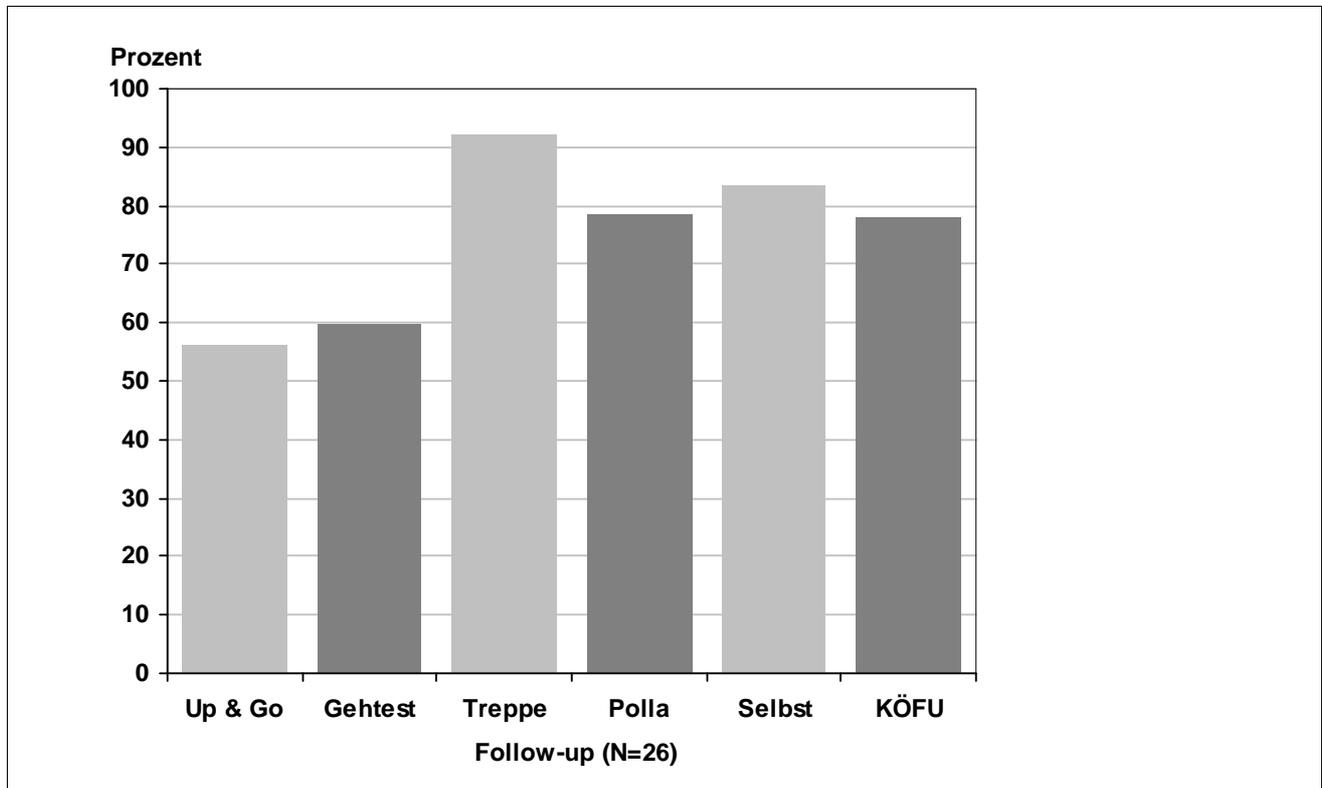


Abb. 16. Die Mittelwerte der Summenskalen zum Follow-up.

### 6.5.1.1 Itemstatistik und Reliabilität

Der Schwierigkeitsindex (Dahl, 1971) zu den Polla Items in der Gruppe mit Follow-up verdeutlicht zu den Zeitpunkten  $t_1$  und  $t_2$  vergleichbare Ergebnisse zu der Itemanalyse der Gesamtstichprobe (Kap. 5.1.1).

Zum Zeitpunkt Therapiebeginn weisen sechs von 14 Items große Schwierigkeiten von 0,2 oder geringer auf, leicht ist dagegen nur das Aufheben mit einem Wert größer als 0,9 (Tab. 64). Gegen Ende der Behandlung sind lediglich das Hocken und Einbeinhüpfen schwierige Items ( $< 0,2$ ), das Aufstehen vom Hocker und das Aufheben mit einem Wert von größer als 0,8 dagegen als leichte Items zu bezeichnen. Zum Follow-up stellt sich kein Item als schwer mit einem Wert unter 0,2 dar.

Tab. 64. Schwierigkeitsindex zu den Polla Items zu  $t_1$ ,  $t_2$  und zum Follow-up.

N=26	$t_1$		$t_2$		Fol		Trennschärfe
	Summe	Index	Summe	Index	Summe	Index	
Strümpfe anziehen	28,00	0,54	39,00	0,75	47,00	<b>0,90</b>	,299
Aufstehen vom Hocker	31,00	0,60	44,00	<b>0,85</b>	47,00	<b>0,90</b>	,695
Einbeinstand	16,00	0,31	40,00	0,77	44,00	0,85	,434
Aufheben	48,00	<b>0,92</b>	52,00	<b>1,00</b>	52,00	<b>1,00</b>	- <sup>1</sup>
Hinlegen	24,00	0,46	41,00	0,79	44,00	<b>0,85</b>	,658
Treppauf eine Stufe	11,00	0,21	34,00	0,65	44,00	<b>0,85</b>	,751
Hocken	1,00	<b>0,02</b>	10,00	<b>0,19</b>	21,00	0,40	,604
Hinknien	4,00	<b>0,08</b>	24,00	0,46	31,00	<b>0,60</b>	,747
Auf der Stelle laufen	9,00	<b>0,17</b>	29,00	0,56	44,00	<b>0,85</b>	,647
Einbeinhüpfen	0,00	<b>0,00</b>	7,00	<b>0,13</b>	26,00	0,50	,777
1 Minute-Up & Go Test	31,00	0,60	37,00	0,71	48,00	<b>0,92</b>	,545
2-Minuten-Gehtest	32,00	0,62	35,00	0,67	47,00	<b>0,90</b>	,634
Treppenlaufen-Test	27,00	0,52	39,00	0,75	47,00	<b>0,90</b>	,736
2-Minuten-Lauftest	0,00	<b>0,00</b>	4,00	0,08	14,00	0,27	,466

Anm.: <sup>1</sup> Varianz gleich null, kann daher nicht berechnet werden.

Das arithmetische Mittel der 14 Itemschwierigkeiten liegt zu Beginn bei 0,36, bei Abschluss bei 0,6 und zum Follow-up bei 0,76. Mit fortschreitendem Genesungsprozess werden die Aktivitäten also leichter bewältigt. Die Ergebnisse unterstützen die Entscheidung, die schwierigeren Items des Hockens, Hinknien, Einbeinhüpfens und des Lauftests mit Schwierigkeiten im Bereich von 0,6 bis 0,27 in der Testbatterie zu belassen, da sie über längere Zeiträume ihre Berechtigung erhalten, indem sie Deckeneffekte verringern. Die Trennschärfe zeigt außer bei dem Aufheben insgesamt gute Eigenschaften (Tab. 64).

Die interne Konsistenz über Alpha Cronbach zeigt eine mittlere Reliabilität um 0,89 für die Summenskala des Polla und eine hohe Reliabilität für die Selbsteinschätzung von 0,96.

## 6.5.2 Schließende Statistik zum Follow-up

### 6.5.2.1 Verteilungen und Gruppenunterschiede

Die Überprüfung von Gruppenunterschieden im Hinblick auf Alter und BMI ergab keine signifikanten Unterschiede zwischen der Gruppe mit Follow-up (n=26) und der Gesamtstichprobe (n=81). Das Kollektiv zeigt hinsichtlich der Diagnosen vergleichbare Verteilungen zur Gesamtstichprobe. Zwölf Patienten finden sich in der Gruppe der Kniediagnosen, acht in der Gruppe von Hüftendoprothesen, fünf in

der Gruppe der Knieendoprothesen und eine Diagnose in der der entwicklungsverzögerten Krankheitsbilder wieder.

Zum Follow-up ist der Polla, die *Körperliche Funktionsfähigkeit*, der *Schmerz*, die *Allgemeine Gesundheitswahrnehmung*, die *Vitalität* sowie das *Psychische Wohlbefinden* normal verteilt, die Selbsteinschätzung dagegen nicht.

### 6.5.2.2 Veränderungen der Items zum Polla

Zu Therapieende zeigen in der Gruppe mit Follow-up gegenüber dem Untersuchungsbeginn elf von 14 Polla Items („ja-eingeschränkt-nein“) signifikante Veränderungen mit F-Werten -0,91 bis -2,0 ( $p < ,05$ , Tab. 65).

Tab. 65. Veränderungen der Polla Items zum Follow-up gegenüber dem Therapieende.

N=26	Neg. Ränge	Pos. Ränge	Bindungen	Z <sup>1</sup>	Sig. <sup>2</sup>
<b>Strümpfe anziehen</b>	1	9	16	<b>-2,53</b>	<b>,011</b>
<b>Aufstehen vom Hocker</b>	1	4	21	-1,34	,180
<b>Einbeinstand</b>	3	8	15	-1,07	,285
<b>Aufheben</b>	0	0	26	,000	1,00
<b>Auf den Boden hinlegen</b>	1	4	21	-1,34	,180
<b>Treppauf eine Stufe</b>	0	9	17	<b>-2,89</b>	<b>,004</b>
<b>Hocken</b>	1	10	15	<b>-2,65</b>	<b>,008</b>
<b>Hinknien</b>	2	7	17	-1,81	,070
<b>Auf der Stelle laufen</b>	0	12	14	<b>-3,22</b>	<b>,001</b>
<b>Einbeinhüpfen</b>	1	12	13	<b>-3,07</b>	<b>,002</b>
<b>1-Miunte-Up &amp; Go-Test</b>	0	11	15	<b>-3,32</b>	<b>,001</b>
<b>2-Minuten-Gehtest</b>	0	12	14	<b>-3,46</b>	<b>,001</b>
<b>Treppenlaufen-Test</b>	0	8	18	<b>-2,83</b>	<b>,005</b>
<b>2-Minuten-Lauftest (N=4/7)</b>	0	5	21	<b>-2,24</b>	<b>,025</b>

Anm.: <sup>1</sup>Basiert auf negativen Rängen; <sup>2</sup>Asymptotische Signifikanz (2-seitig) nach Wilcoxon.

Ausnahme bilden der 1-Minute-Up & Go-Test ( $F = -1,73$ ;  $p = ,08$ ), der Lauftest ( $F = -1,41$ ;  $p = ,16$ ) und der 2-Minuten-Gehtest ( $F = -0,91$ ;  $p = ,37$ ), der bereits in der Gesamtstichprobe nur mit  $p = 0,04$  signifikant war (Tab. 65).

Die *gemessenen Tests im Polla* zeigen zum Follow-up gleichfalls Ergebnisverbesserungen auf (Tab. 66). Der Lauftest weist zwar die höchste mittlere Zunahme von 17,8 auf 73,5 Meter auf, ist aber aufgrund der unterschiedlichen Anzahl von Personen, die den Test ausführen konnten, mit sieben Patienten zum Follow-up gegenüber vier zum Therapieabschluss, eingeschränkt vergleichbar.

Tab. 66. Die deskriptive Statistik zu den Messzeitpunkten  $t_1$ ,  $t_2$  und zum Follow-up.

N=26	Mzp	N	Min	Max	M	s
1-Minute-Up & Go-Test <sup>1</sup>	$t_1$	26	4	11	6,6	1,88
	$t_2$	26	5	15	9,1	2,42
	fol	26	6	14	10,0	2,24
2-Minuten-Gehtest <sup>2</sup>	$t_1$	26	62	198	131,0	33,52
	$t_2$	26	88	266	173,9	42,36
	fol	26	102	274	186,2	43,31
Treppenlaufen-Test <sup>3</sup>	$t_1$	26	19	134	48,9	23,52
	$t_2$	26	12	62	25,9	11,10
	fol	26	11	39,4	20,7	7,55
2-Minuten-Lauftest <sup>2</sup>	$t_1$	0	-	-	-	-
	$t_2$	4	0	247	17,8	63,19
	fol	7	0	370	73,5	126,9

Anm.: Angaben in: <sup>1</sup>Wegstrecken; <sup>2</sup>Metern; <sup>3</sup>Sekunden.

Tabelle 67 zeigt zum Follow-up gegenüber dem Untersuchungsende bedeutsame Stichprobenunterschiede, überprüft mit dem T-Test bei gepaarten Stichproben, für alle gemessenen Tests. Die deskriptive Statistik ist der Tabelle 68 zu entnehmen. Einen anschaulichen Überblick über die Veränderungen der gemessenen Tests gibt Abbildung 17.

Tab. 67: Ergebnisse des T-Tests auf Stichprobenunterschiede bei gepaarten Tests ( $t_2$ /Follow-up).

N=26	Gepaarte Differenzen bei 95% KI <sup>1</sup>				T-Wert	df <sup>3</sup>	Sig. <sup>2</sup>
	M	s	Untere	Obere			
1-Minute-Up & Go-Test	-,90	1,11	-1,35	-,450	<b>-4,14</b>	25	,000
2-Minuten-Gehtest	-12,4	19,29	-20,14	-4,55	<b>-3,27</b>	25	,003
Treppenlaufen-Test	5,2	5,88	2,87	7,62	<b>4,55</b>	25	,000
2-Minuten-Lauftest	-55,7	103,3	-97,36	-13,95	<b>-2,75</b>	25	,011

Anm.: <sup>1</sup>95% Konfidenzintervall der Differenz; <sup>2</sup>Signifikanz (2-seitig), <sup>3</sup>Freiheitsgrade.

Tab. 68. Die deskriptive Statistik bei gepaarten Stichproben.

N=26		Mittelwert	s
Paar 1	$t_2$ Up & Go-Test	9,12	2,42
	fol Up & Go-Test	10,02	2,24
Paar 2	$t_2$ Gehtest	173,88	42,36
	fol Gehtest	186,23	43,31
Paar 3	$t_2$ Treppenlaufen-Test	25,91	11,10
	fol Treppenlaufen-Test	20,67	7,55
Paar 4	$t_2$ Lauftest	17,85	63,19
	fol Lauftest	73,50	126,86

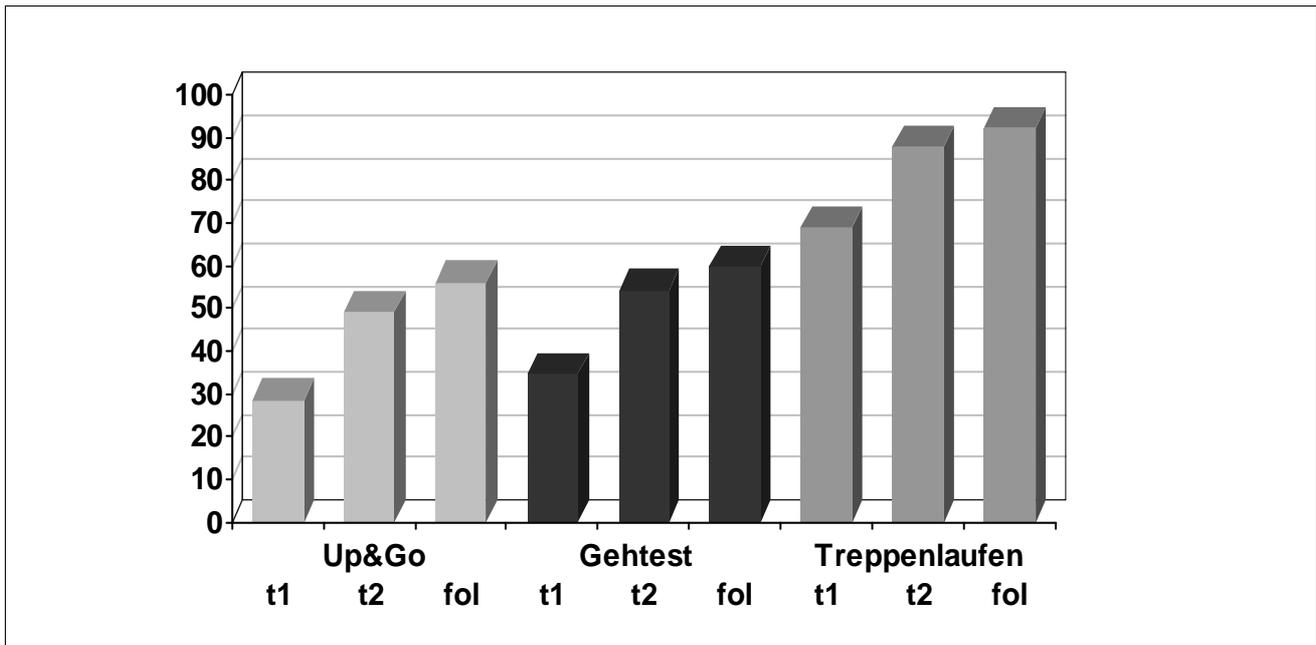


Abb. 17. Die Veränderungen der 0-100 transformierten Tests von t1 bis zum Follow-up.

### 6.5.3 Testgütekriterien zum Polla im Follow-up

#### 6.5.3.1 Konstruktvalidität

Die Konstruktvalidität präsentiert zum Follow-up mit 26 Versuchspersonen ein vergleichbares Bild wie zu Therapiebeginn und -abschluss (Tab. 69). Bei geringeren Fallzahlen ergeben sich insgesamt mittlere bis hohe Korrelationen zwischen der Polla Summenskala, den gemessenen Tests und ihrer Summenskala.

Tab. 69. Korrelationen nach Spearman-Rho zum Follow-up (n=26).

N=26	Polla Items („ja-eingeschränkt-nein“)	Gemessene Tests <sup>1</sup>	Korrelation	Sig. <sup>2</sup>
fol	Up & Go-Test	1-Minute-Up & Go-Test	,307	,127
	Gehstest	2-Minuten-Gehstest	,326	,105
	Treppenlaufen-Test	Treppenlaufen-Test	<b>,579**</b>	<b>,002</b>
fol	Polla Summenskala	1-Minute-Up & Go-Test	<b>,700**</b>	<b>,000</b>
		2-Minuten-Gehstest	<b>,633**</b>	<b>,001</b>
		Treppenlaufen-Test	<b>,762**</b>	<b>,000</b>
fol	Polla Summenskala	Summenskala (gemessene Tests)	<b>,719**</b>	<b>,000</b>

Anm.: <sup>1</sup> Grundlage der Berechnungen: 0-100 transformierte Skalen; <sup>2</sup> Signifikanzniveau (zweiseitig).

### 6.5.3.2 Kriteriumsvalidität

Die Konstanz von Zusammenhängen zum Follow-up mit 26 Patienten setzt sich zwischen dem Polla, der Selbsteinschätzung und den gemessenen Tests gegenüber der *Körperlichen Funktionsfähigkeit* und der *Rollenfunktion* des SF-36 fort (Tab. 70). Gleiches gilt für die Zusammenhänge zwischen den SF-36 Dimensionen *Schmerz* und *Allgemeiner Gesundheitszustand* und dem Polla.

Tab. 70. Korrelationen zwischen den Erhebungsinstrumenten.

N=26		Pearson	Spearman-Rho				
MZP	Dimen-sion	KÖFU	Polla	Selbst <sup>9</sup>	Up & go	Gehtest	Treppen-laufen
fol	KÖFU <sup>1</sup>	1	,800**	,859**	,694**	,537**	,718**
	KÖRO <sup>2</sup>	,517**	,610**	,594**	,495*	,409*	,551**
	SCHM <sup>3</sup>	,620**	,435*	,520*	,249	,232	,409*
	AGES <sup>4</sup>	,311	,253	,499*	,293	,268	,462*
	VITA <sup>5</sup>	,215	,209	,158	,223	,229	,289
	SOFU <sup>6</sup>	,034	,129	,237	-,058	-,060	,060
	EMRO <sup>7</sup>	,269	,289	,386	,142	,171	,281
	PSYC <sup>8</sup>	,080	,101	,150	,084	,084	,083

Anm.: Signifikanzniveau: \*\*p ≤ 0.01; \*p ≤ 0.05; <sup>1</sup>Körperliche Funktionsfähigkeit; <sup>2</sup>Körperliche Rollenfunktion; <sup>3</sup>Schmerz, <sup>4</sup>Allgemeiner Gesundheitszustand; <sup>5</sup>Vitalität; <sup>6</sup>Soziale Funktionsfähigkeit; <sup>7</sup>Emotionale Rollenfunktion; <sup>8</sup>Psychisches Wohlbefinden; <sup>9</sup>Selbsteinschätzung.

### 6.5.3.3 Unterschiede zwischen Follow-up und Gesunden

Auch zum Follow-up ist zu erwarten, dass sich das Funktionsniveau von Patienten an das von Gesunden im Verlauf des Heilungsprozesses angleicht. Fraglich ist, ob die Tests in der Lage sind, Patienten von Gesunden zu unterscheiden. Eine Analyse zu Gruppenunterschieden wies bereits (Kap. 6.4.5) große Effekte für alle Instrumente zu Therapiebeginn nach, die sich gegen Untersuchungsende abschwächen. An dieser Stelle muss allerdings erwähnt werden, dass sich die Patientengruppe im Follow-up von der Gruppe der Gesunden hinsichtlich des BMI signifikant unterscheidet (T: 3,4; p = ,001); nicht dagegen im Hinblick zum Alter (T: 0,24; p = ,808). Die deskriptiven Werte sind der Tabelle 71 zu entnehmen.

Tab. 71. Deskriptive Gruppenstatistiken.

	Gruppe	N	M	s
Alter	Follow-up	26	50,81	18,24
	Gesunde	30	49,67	16,78
Body-Mass-Index	Follow-up	26	29,08	6,10
	Gesunde	30	24,88	2,51

Tab. 72. Unterschiede zwischen den Gruppen „Gesunde“ und „Follow-up“ (Varianzanalyse, 95% KI).

MZP		Gruppe	N	M	s	df	Mittel der Quadrate	F	Sig.	Eta
fol	Polla	fol	26	78,55	20,60	1	6094	<b>30,83</b>	<b>,000</b>	<b>,602</b>
		Ges	30	99,44	1,44					
	Up & Go	fol	26	10,02	2,24	1	47,54	<b>10,08</b>	<b>,002</b>	<b>,396</b>
		Ges	30	11,87	2,11					
	Gehstest	fol	26	186,23	43,31	1	16293	<b>14,42</b>	<b>,000</b>	<b>,459</b>
		Ges	30	220,43	22,07					
	Treppe	fol	26	20,67	7,55	1	476,21	<b>15,78</b>	<b>,000</b>	<b>,475</b>
		Ges	30	14,82	2,67					
	Laufstest	fol	26	73,50	126,86	1	423480	<b>43,52</b>	<b>,000</b>	<b>,668</b>
		Ges	30	247,87	65,16					

Es zeigen sich weiterhin mittlere bis große Effekte zu Unterschieden für alle Instrumente zwischen Patienten mit Follow-up und Gesunden. Dies gilt insbesondere mit 0,668 für den Laufstest und 0,602 für den Polla (Tab. 72), was die Bedeutung des Laufstests insbesondere bei fortschreitender Genesung unterstreicht.

#### 6.5.3.4 Änderungssensitivität

Die Betrachtung der Änderungssensitivität für die Patientengruppe mit Follow-up (n=26) erfolgt zu den Zeitpunkten  $t_1$ ,  $t_2$  und zum Follow-up. Berechnet wurde die Änderungssensitivität über den Standardized Response Mean (SRM, Kap. 6.3.6). Die nachgewiesenen großen Änderungseffekte im Aktivitätsstatus zu Therapieabschluss werfen die Frage nach Veränderungen über Zeiträume von drei und mehr Monaten auf. Sind im Genesungsprozess mit erfahrungsgemäß zunehmend kleineren Fortschritten diese statistisch noch belegbar? Kritisch zu erwähnen sei an dieser Stelle die kleine Fallzahl von 26 Personen mit den Möglichkeiten der Selektion.

Die Änderungssensitivität in der Gruppe mit Follow-up weist mit einer SRM von 1,1 bis 1,7 für alle Instrumente zum Untersuchungsende – 42 Tage nach Therapiebeginn – vergleichbare Effekte zur Gesamtstichprobe auf (Tab. 73; vgl. Kap. 6.4.6). Verständlicherweise größere Effekte von 1,5 bis 2,5 ergeben sich zwischen Therapiebeginn und dem Follow-up nach durchschnittlich 122 Tagen.

Tab. 73. Die Änderungssensitivität in der Gruppe mit Follow-up zu  $t_1$ ,  $t_2$  und zum Follow-up.

N=26				SRM		
				42 Tage	80 Tage	122 Tage
Merkmal	Mzp <sup>1</sup>	M	S	$t_1/t_2$	$t_2/t_{fol}$	$t_1/_{fol}$
KÖFU	$t_1$	38,82	20,16	1,35	0,69	2,05
	$t_2$	65,24	18,94			
	fol	77,83	17,79			
Polla	$t_1$	34,29	14,40	1,68	0,94	2,49
	$t_2$	60,74	17,00			
	fol	78,55	20,61			
Selbsteinschätzung	$t_1$	49,28	19,20	1,38	0,55	1,87
	$t_2$	74,10	16,55			
	fol	83,29	17,13			
1-Minute-Up & Go-Test	$t_1$	28,46	14,94	1,18	0,39	1,69
	$t_2$	48,92	19,38			
	fol	56,15	17,91			
2-Minuten-Gehtest	$t_1$	34,42	15,38	1,12	0,29	1,46
	$t_2$	54,06	19,43			
	fol	59,74	19,87			
Treppenlaufen-Test	$t_1$	68,89	19,28	1,25	0,55	1,62
	$t_2$	87,77	9,10			
	fol	92,07	6,19			

Anm.: Grundlage der Berechnungen sind die 0-100 transformierten Skalen.

Geringer sind die Veränderungen zum Follow-up gegenüber dem Therapieende nach im Mittel 80 Tagen (0,29-0,94). Der Polla zeigt weiterhin große Effekte, die Körperliche Funktionsfähigkeit, die Selbsteinschätzung und der Treppenlaufen-Test mittlere Effekte und die beiden anderen Tests geringe Effekte (Tab. 73).

## 6.6 Zusammenfassende Darstellung der Testgütekriterien

Einen Überblick auf die erhobenen Gütekriterien zum Polla gibt Tabelle 74.

Tab. 74. Überblick über die Testgütekriterien zum Polla und den gemessenen Tests zu  $t_1$  und  $t_2$ .

Testgütekriterium	N	Werte	Beurteilung	Angewandte statistische Verfahren	Bemerkungen
Interrater-Reliabilität	32	0,747	gut	Kappa-Werte gemittelt	Polla/ drei Therapeuten
Intrarater-Reliabilität	29	0,98	ausgezeichnet	ICC <sup>2</sup>	gemess. Tests
Interne Konsistenz	81	0,79	mittelmäßig	Alpha Cronbach	Polla
Konstruktvalidität	81	0,42-0,71	gering-mittel	Spearman-Rho	Polla/gemess. Tests
Kriteriumsvalidität	81	0,55-0,62	mittel	Spearman-Rho	Polla/KÖFU
Änderungssensitivität t2 gegenüber t1	81	1,79	groß	SRM <sup>1</sup>	Polla
		1,4			KÖFU
		1,4			Selbsteinschätzung
Änderungssensitivität zum Follow-up	26	0,9	groß		Polla
		0,7	mittel		KÖFU
		0,5	klein		Selbsteinschätzung
Unterscheidungsfähigkeit zwischen Gesunden und Betroffenen	30	0,956-0,849	groß	Varianzanalyse	Polla
Effekte zu Unterschieden	81	16-54%	mittel-groß	Varianzaufklärung	Altersklassen Treppenlaufen-Test Knie-/Hüftendoprothesen
	72	18%	mittel		

Anm.: <sup>1</sup>Standardized Response Mean, <sup>2</sup>Intra-Class-Korrelationskoeffizient.

Die Ergebnisse zur Objektivität (*Interrater-Reliabilität*) deuten mit einem gemittelten Kappa-Wert von  $\kappa = 0,747$  sehr gute Werte und für die *Intrarater-Reliabilität* eine ausgezeichnete Übereinstimmung von  $ICC = 0,98$  an.

Bei insgesamt zufrieden stellender *interner Konsistenz* von Alpha Cronbach  $\alpha = 0,79$  (Tab. 74) weisen die Zusammenhänge zur *Konstruktvalidität* zwischen den einzelnen gemessenen Tests und dem Summenscore der qualitativ beurteilten Tests überwiegend mittlere Korrelation auf ( $r = 0,42 - 0,71$ ; Tab. 74).

Die Ergebnisse der *Kriteriumsvalidität* für die dimensionsähnliche Ebene von Aktivitäten zeigen moderate Zusammenhänge zwischen der KÖFU und dem Polla nach Spearman-Rho ( $r = 0,55 - 0,62$ ; Tab. 74).

Veränderungen im Therapieverlauf (*Änderungssensitivität*) vermag der Polla im Vergleich zu den anderen Instrumenten mit größeren Effekten ( $SD_{pool} = 1,79$ ; Tab.

74) zwischen t2 und t1 in der Gesamtstichprobe (N=81) sowie zum Follow-up gegenüber t2 mit 26 Patienten ( $SD_{\text{pool}} = 0,94$ ) darzustellen.

Der Polla ist sowohl zu Beginn wie auch zum Therapieabschluss in der Lage, Gesunde von Patienten mit großen Effekten von  $E > 0,8$  zu unterscheiden (Tab. 74). Mittlere bis große Varianzaufklärungen von 16-54 Prozent des 1-Minute-Up & Go-Tests, des 2-Minuten-Gehtests sowie des Treppenlaufen-Tests am Faktor Alter deuten auf die Fähigkeit der gemessenen Tests hin, leistungsabhängige *Altersunterschiede* herauszuarbeiten (Tab. 74). Unterscheidungen zwischen *Diagnosegruppen* werden konsequent nur über den Treppenlaufen-Test zwischen den Gruppen der Patienten mit Knieendoprothesen und den beiden anderen Gruppen mit mittleren Varianzaufklärungen von 18 Prozent getroffen.

## 7 DISKUSSION

Den Ausgangspunkt der vorliegenden Arbeit bildet ein Forschungsdefizit im Bereich von Performance-Tests, das insbesondere von der Arbeitsgruppe um Bührlen et al. (2002) gekennzeichnet wird. Das Ziel besteht daher in der Konstruktion und Validierung eines Performance-Tests für die untere Extremität. Die Kriterien für die Ableitung zentraler Aktivitäten für den Polla ergaben sich aus einer Literatursichtung, den in den Domainen *Mobilität* und *Selbstversorgung* der ICF beschriebenen Aktivitäten sowie aus den Empfehlungen nach Bührlen et al. (2002).

In den folgenden Ausführungen werden die Ergebnisse zusammenfassend und im Hinblick auf ihre Auswirkungen diskutiert. Die Diskussion der Korrelationshypothese als eine der zentralen Fragestellungen nimmt ebenso wie die Diskussion der Änderungssensitivität besonderen Raum ein. Weiterhin finden die Ergebnisse zur Konstruktvalidität und zu Gruppenunterschieden sowie die Veränderungen zu spezifischen Zielkriterien kritische Betrachtung.

### 7.1 Diskussion der Kriteriumsvalidität (Korrelationshypothese)

Die Ergebnisse zur Kriteriumsvalidität spiegeln die durchweg bedeutsamen Zusammenhänge des Polla zu der verwandten Dimension der *Körperlichen Funktionsfähigkeit* und darüber hinaus zu fast allen anderen SF-36-Dimensionen wider. Entsprechend geht ein höheres Ergebnis im Polla mit einer größeren Lebensqualität einher. Damit bestätigen sich die in der Korrelationshypothese angenommenen Zusammenhänge des Performance Tests zur gesundheitsbezogenen Lebensqualität ebenso wie die zur Selbsteinschätzung von Aktivitäten, die auch von Guralnik et al. (1994) bestätigt werden. Zudem zeigen die Ergebnisse in einer Untersuchung von Strand, Ljunggren, Haldorsen und Espehaug (2001) an Patienten mit Rückenschmerzen ein Jahr danach, dass die Patienten mit einer Verbesserung des Funktionsstatus, gemessen über Performance-Tests, eher wieder zur vollen Arbeitsfähigkeit zurückgekehrt sind als in der Kontrollgruppe.

Die insgesamt schwächeren Zusammenhänge zwischen den gemessenen Tests und den SF-36 Dimensionen verglichen mit dem Polla lassen sich durch ihre weniger differenzierte Abbildung von Alltagsaktivitäten erklären. Zu den Aktivitäten des täglichen Lebens zählen eben nicht nur der Up & Go-, der Geh- oder Treppenlaufen-Test, sondern in gleichem Maße auch die Aktivitäten des Einbeinstands, des Hinlegens, Hockens und Hinkniens sowie des Einbeinhüpfens. Je weniger Einschränkungen in Form von Hilfen, Hilfsmitteln, Schmerzen oder Ausweichbewegungen auftreten, desto höher fallen die Bewertungen des individuellen Gesundheitszustandes aus. Auch seelische Schwierigkeiten bei der Alltagsbewältigung

(EMRO) durch Tätigkeiten, die nicht so lange oder nicht mit entsprechender Sorgfältigkeit ausgeführt werden, können verringern sich, wenn die Aktivitäten des Polla besser bewältigt werden. Gleiches gilt für die Beeinträchtigungen von sozialen Kontakten zum Bekannten- und Freundeskreis (SOFU).

Der gemessene Treppenlaufen-Test weist gegenüber den beiden anderen gemessenen Tests leichte Vorteile in der Stabilität von Zusammenhängen zu der gesundheitsbezogenen Lebensqualität auf. Die einbeinige Beugebelastung und die Mobilität des Treppensteigens sind wichtige Aktivitäten des Alltags.

Verschiedene Autoren konnten zeigen, dass zwischen Instrumenten der Selbsteinschätzung und Fremdeinschätzungen in Form von Performance-Tests lediglich geringe bis moderate Zusammenhänge auftreten (Lin, Davey & Cochrane, 2001; Rejeski, Ettinger, Schumaker, James, Burns & Elam, 1995; Lee, Simmonds, Novy & Jones, 2001; Kennedy, Stratford, Pagura, Walsh & Woodhouse, 2002). Lin, Davey und Cochrane (2001) begründen die geringen Zusammenhänge mit Mess- und Auswertungsfehlern (Guralnik et al., 1994) sowie sich ständig verändernden Symptomatiken (Bellamy, 1993). Da Performance-Tests objektiv funktionelle Einschränkungen erfassen, während Instrumente zur Messung der gesundheitsbezogenen Lebensqualität eher Behinderungen aufspüren, seien keine perfekten Übereinstimmungen bei Korrelationen zu erwarten (Lin et al., 2001). Auch könnten Patienten in der Funktionsüberprüfung ein besonders gutes Ergebnis erreichen wollen, um dem Untersucher oder Therapeuten zu gefallen (Lin et al., 2001).

Weitere Diskrepanzen ergeben sich aus der schriftlichen Abfrage solcher Aktivitäten, die nach Operationen häufig noch nicht oder auch nie ausgeführt werden, weil beispielsweise die Aktivität des Joggens für einen 75-jährigen entweder im Moment nicht möglich ist oder unter Umständen nicht mehr relevant sein wird. Häufig unterschätzen Patienten ihre Leistungsfähigkeit in der schriftlichen Selbsteinschätzung. Ein Patient gibt an, er könne sich noch nicht hinlegen, beweist aber im anschließenden Performance-Test das Gegenteil. Solche Fälle wurden nicht selten beobachtet. So waren beispielsweise zu Therapiebeginn 80 von 81 Personen nicht in der Lage das Einbeinhüpfen auszuführen, dagegen äußerten dies in der Selbsteinschätzung zum selben Zeitpunkt nur 61 Personen (vgl. Kelly-Hayes, Jette, Wolf, D'Agostino & Odell, 1992). Fehleinschätzungen resultieren aus Autorensicht zudem aus Diskrepanzen zwischen der Patientenvorstellung über die Art der Durchführung alltäglicher Bewegungen und den vorgegebenen Testanforderungen. In diesem Zusammenhang wäre interessant festzustellen, wie sich die Ergebnisse der Selbsteinschätzung verändern, würde der Zeitpunkt der Selbsteinschätzung einmal vor oder nach der Durchführung der Tests gewählt.

Die Unterschiede zwischen den Ergebnissen der Fremdbeurteilungen und Selbsteinschätzungen veranlassen verschiedene Autoren zu der Empfehlung, beide Erhebungsarten einzusetzen, um ein sich ergänzendes Abbild von Therapieergebnissen zu erhalten (Guralnik et al., 1994; Kelly-Hayes et al., 1992; Reuben, Siu & Kimpau, 1992).

Kohlmann (2004b) sieht den wesentlichen Vorteil patientenbezogener Selbsteinschätzungen in deren hoher Praktikabilität und Ökonomie. Dennoch sollten in den Erhebungsperspektiven sowohl die patientenbezogene als auch die klinische Perspektive berücksichtigt werden, da nur eine komplementäre Betrachtungsweise dieser Parameter ein ausgewogenes Bild der Krankheitsverläufe und Therapieergebnisse ermöglicht.

## 7.2 Diskussion der Änderungssensitivität

Die über alle Instrumente zu Therapieende gemessenen großen Effekte zu Veränderungen gegenüber dem Beginn waren in dem frühen postoperativen Zeitraum, mit erfahrungsgemäß schnellen Entwicklungen, zu erwarten. Geringere Effekte der körperlichen Funktionsfähigkeit (KÖFU) gegenüber dem Polla werden in der Tatsache vermutet, dass in die KÖFU anstrengende Tätigkeiten (z.B. schnell laufen, schwere Gegenstände heben, anstrengenden Sport treiben) und mittelschwere Tätigkeiten (z.B. einen Tisch verschieben, Staubsaugen, Kegeln, im Garten arbeiten) in zwei Items aggregiert sind und damit die Abbildung kleinerer Veränderungen erschweren.

Dass generische Fragebögen wie der SF-36 mit ihrem Anliegen zur umfassenden Abbildung psychometrischer Kriterien eindimensionale krankheitsspezifische Merkmale weniger sensibel dokumentieren, ist daher nachvollziehbar (Kohlmann, 2004a).

Insgesamt werden kleinere Effekte zum Follow-up aus den sich verringernden Adaptationen verständlich, je weiter der Genesungsprozess fortschreitet. Eine geringere Änderungssensitivität der gemessenen Tests gegenüber dem Polla wird aus den gegen Therapieende erreichten Leistungsgrenzen deutlich, wobei der Treppenlaufen-Test zum Follow-up noch das größte Veränderungspotential der gemessenen Tests aufweist. Entscheidende Verbesserungen in den Ergebnissen sind darüber hinaus nur durch konsequentes Üben der Testsituation möglich.

Der Polla ist über seine qualitativen Veränderungen durch die Reduktion von Schmerzen, Verringerung des Einsatzes von Hilfen/Hilfsmitteln insbesondere über die Aktivitäten Treppauf eine Stufe, Hocken, Laufen, Einbeinhüpfen, bei dem Up &

Go-Test sowie dem Geh- und Treppenlaufen-Test insgesamt besser in der Lage, Unterschiede von Veränderungen darzustellen.

### 7.3 Diskussion der Konstruktvalidität

Die Zusammenhänge zwischen dem Konstrukt der drei qualitativ beurteilten Polla Items („ja-eingeschränkt-nein“) und ihren inhaltlichen Entsprechungen der quantitativ gemessenen Tests: Der 1-Minute-Up & Go-Test, der 2-Minuten-Gehtest und der Treppenlaufen-Test finden an dieser Stelle ihre nähere Betrachtung.

Zunächst bedeuten die Ergebnisse der fehlenden Korrelationen im Hinblick auf den Up & Go- und den Gehtest, dass Patienten mit Schmerzen oder dem Einsatz von Hilfen/Hilfsmitteln genauso schnell oder schneller gehen können als Patienten ohne Einschränkungen. Die Ursachen ergeben sich aus der grundsätzlichen Verschiedenheit qualitativer und quantitativer Datenerhebung. Nach der Beurteilung, ob eine Aktivität möglich, eingeschränkt oder nicht möglich ist, sind Patienten zu Therapiebeginn mit 82-86% der Fälle nur eingeschränkt in der Lage oder unfähig diese Tests auszuführen. Durch die drei gesetzten Antwortalternativen sind die Antwortkategorien auf zwei beschränkt (eingeschränkt-nein) und reduzieren mögliche Unterscheidungen. Auf der anderen Seite unterliegen die quantitativen gemessenen Tests einem leistungsbezogenen, damit auch im Outcome altersabhängigen Überprüfungscharakter und beinhalten die Aufnahme einer großen Ergebnisvariabilität.

Im Zusammenhang mit Alterseinflüssen auf Testergebnisse kommt dem 1-Minute-Up & Go-Test in zweierlei Art eine wichtige Bedeutung zu. Zum einen zeigt er die größten Effekte zu Unterschieden. Sie sind dadurch begründet, dass der Test durch Zeitdruck hohe Anforderungen an die verschiedenen motorischen Grundeigenschaften wie Beweglichkeit, Kraft und Gleichgewicht sowie deren Koordination im Bewegungsablauf bei insgesamt höherem Sturzrisiko stellt. Der 1-Minute-Up & Go-Test erfordert nämlich das Aufstehen vom Hocker – Beschleunigen des Körpers über eine kurze Strecke von drei Metern – Abbremsen – Durchführen einer schnellen Umkehrbewegung – wieder Beschleunigen des Körpers und sich in leichter Rückwärtsbewegung aus einer halben Körperdrehung wieder auf den Hocker zu setzen, um dies bis zu 15 Mal in der Testzeit von einer Minute zu wiederholen. Eine unter Zeitmessung gerade für Ältere durchaus nicht einfache Aufgabe.

Die Anforderungen im Gehtest beschränken sich dagegen auf eine hohe Ganggeschwindigkeit bei wenig Umkehrbewegungen, fehlender Beugebelastung und geringerer Sturzgefahr.

Bei dem Treppenlaufen-Test reduzieren sowohl die Teststandardisierung – es dürfen nicht mehrere Stufen auf einmal genommen werden – als auch die Beugebelastung über das betroffene Bein sowie die insgesamt kürzere mittlere Testdauer die Möglichkeit größere altersbedingte Leistungsunterschiede zu erfassen. Gleichzeitig erhöhen sich aber die Zusammenhänge zwischen den Items Treppensteigen „ja-eingeschränkt-nein“ und dem gemessenen Treppenlaufen.

Die in der Summenskala zum Polla inhaltlich dagegen differenzierter angelegte Beurteilung von Aktivitäten demonstriert unter Ausschluss der drei hier behandelten Items zu beiden Zeitpunkten gleichmäßigere Verteilungen über die drei Beurteilungskategorien. Aus dieser Tatsache werden die mittleren bis hohen Zusammenhänge zwischen dem Polla und den gemessenen Tests verständlich.

#### 7.4 Diskussion zu Gruppenunterschieden

Die Ergebnisse der Varianzanalysen zu Unterschieden in Altersklassen machen z.T. große Alterseinflüsse auf die Ergebnisse zum 1-Minute-Up & Go-Test und 2-Minuten-Gehtest mit Varianzaufklärungen von 16 bis 54 Prozent v.a. über die Trennung der jüngeren 19-38-jährigen von den Gruppen der Älteren deutlich.

Eine Interpretation von Unterschieden zu Diagnosegruppen ist nur eingeschränkt möglich. Unter der weiten Fassung von Einschlusskriterien und der Bildung von Diagnosegruppen, die an große Altersunterschiede gebunden war, sind Ergebnisse zwangsläufig Alterseinflüssen ausgesetzt. Die Unterscheidung von Diagnosen war aber nicht primäres Ziel der vorliegenden Arbeit. Vielmehr sollte eine Testbatterie mit einer breiten Vielfalt an Diagnosen unter regulären physiotherapeutischen Arbeitsbedingungen auf ihre Gütekriterien überprüft werden. Hier könnten Untersuchungen mit einer engeren Fassung von Ein- und Ausschlusskriterien zu günstigeren Ergebnissen führen.

Die größten Unterschiede zu Diagnosegruppen werden zwischen den komplexeren Krankheitsbildern der Knieendoprothesen und den einfacheren Diagnosen am Knie herausgearbeitet (11-39% Varianzaufklärung). Die Unterschiede sind allerdings auf eine Addition von Effekten zurückzuführen, die sich einerseits durch länger währende funktionelle Wiederherstellung nach operativen Kniegelenkersatz, aber insbesondere durch große mittlere Altersunterschiede von über 30 Jahren zwischen den beiden Gruppen ergeben.

Bei einem mittleren Altersunterschied von nur 2 Jahren zeigen die beiden Gruppen der Knie- und Hüftendoprothesen die *deutlichsten diagnosebezogenen Unterschiede* und sind daher altersunabhängig zu diskutieren. Die Unterschiede werden im

besonderen Maße über den Treppenlaufen-Test und den Polla deutlich. Die Varianzaufklärung der Diagnosen am Treppenlaufen-Test und Polla beträgt 18-39 Prozent. Die Unterscheidungsfähigkeiten bestätigen den Eindruck, dass Aktivitäten wie das Treppensteigen, Hinknien, Hocken, Einbeinhüpfen und Laufen – wie sie im Polla vertreten sind – Patienten nach endoprothetischem Kniegelenkersatz schwerer fallen.

Die häufigere und etwas höhere Unterscheidungsfähigkeit des Treppenlaufen-Tests im Hinblick auf Diagnosegruppen als auf Altersklassen unterstützt die Vermutung, dass einbeinige Kniebeugebelastungen über die betroffene Extremität hohe funktionelle Anforderungen an den endoprothetischen Kniegelenksersatz stellen.

Stabilere Unterschiede zu Therapieabschluss finden ihre Begründung in einer Verwischung von Effekten zu Therapiebeginn. So gehen beispielsweise einerseits jüngere Patienten aufgrund des Schweregrads ihrer Diagnose häufig noch mit Schmerzen an Stützen, andererseits Ältere nach endoprothetischem Hüftersatz nicht selten ohne Hilfsmittel und beschwerdefrei.

## 7.5 Diskussion zu spezifischen Zielkriterien

Neben den Performance-Tests, die Veränderungen im Therapieverlauf darlegen, geben die in der Krankengymnastik üblichen Zielkriterien ebenfalls Auskunft über bedeutsame Unterschiede gegenüber dem Therapiebeginn. Sowohl die *aktive Flexion*, die *schmerzhafte Kniebeuge* als auch die Messung der *Schmerzintensität* (NRS) dokumentieren dies mit großen Änderungssensitivitäten (SRM).

Die Zusammenhänge zur Selbsteinschätzung, zur *Körperlichen Funktionsfähigkeit* wie auch zu den gemessenen Tests unterstreichen die Notwendigkeit zur regelmäßigen *aktiven Beweglichkeitsmessung* in der postoperativen Nachbehandlung. Diese findet allerdings dort ihre Grenzen, wo Funktionskrankheiten auftreten – denn sie gehen nur bedingt oder gar nicht mit Bewegungseinschränkungen einher – oder dort, wo sich Probleme aufgrund einer Beweglichkeitsmessung unter Entlastung des Körper (hier: in Rückenlage) nicht widerspiegeln. So weist die aktive Beweglichkeit im Gegensatz zur schmerzhaften Kniebeuge zu  $t_2$  keine Zusammenhänge mehr zur NRS auf. Eine Untersuchung von Simmonds et al. (1998) an konservativ behandelten chronischen Rückenpatienten stützt diese Annahme. Sie konnten zeigen, dass sich Performance-Tests als bessere Anzeiger möglicher Behinderungen darstellen, als eine Beweglichkeitsmessung des Rückens. Zwar weist die aktive Flexionsmessung zu Beginn höhere Zusammenhänge zu allen gemessenen Tests, zum Polla sowie zur KÖFU und Selbsteinschätzung auf, diese

sind aber gegen Therapieende zu keiner der SF-36 Dimension, sondern deutlich nur noch zu den gemessenen Tests vorhanden.

Dagegen beweist die *Messung der schmerzfreien Kniebeuge* in der Untersuchung ihre Qualität gegenüber der aktiven Flexionsmessung. Sie macht Zusammenhänge zu allen SF-36 Skalen und damit zur subjektiven Gesundheit zu  $t_2$  deutlich und fällt insgesamt stärker zu den gemessenen Tests, dem Polla und der Selbsteinschätzung zu Aktivitäten aus.

Konstante und *höhere Zusammenhänge* zu  $t_2$  gegenüber  $t_1$  zum Schmerz (NRS) bezüglich der Selbsteinschätzungen und Performance-Tests begründen sich aus einer insgesamt eher geringeren Belastungsfähigkeit und aus dem Einsatz von Hilfsmitteln (z.B. Stützen). Infolgedessen können weniger Schmerzen bei höheren Testergebnissen auftreten, was auch die Berechnungen zur Konstruktvalidität bestätigen. Mit einer Belastungssteigerung nimmt dann die Schmerzintensität unter einer Steigerung der Testergebnisse ab, womit sich die Zusammenhänge verstärken.

*Zusammenfassend* sind angesichts alltagsorientierter Funktionseinsätze der Extremitäten im Hinblick auf die frühe postoperative Bewegungsbeurteilung die aktive Beweglichkeitsmessung in Rückenlage und die Messung der schmerzhaften Kniebeuge empfehlenswert. Im fortgeschrittenen Genesungszustand und bei Funktionskrankheiten sollte auf die Messung der *schmerzhaften Kniebeuge* zurückgegriffen werden. Sie stellt ein schnell durchzuführendes zusätzliches Maß für die Beweglichkeitsmessung der unteren Extremität dar.

## 8 SCHLUSSFOLGERUNGEN UND EMPFEHLUNGEN

Als zeitökonomisches und kostenloses Testverfahren ist der Polla mit hoher praktischer Relevanz in jeder physiotherapeutischen Einrichtung einsetzbar. In Form einer Fremdbeurteilung als Therapeutenurteil liefert er Informationen, die über die von Selbsteinschätzungen hinausgehen. Die Ergebnisse von Performance-Tests schaffen bei Therapeuten genauere Vorstellungen über die Qualität der Bewegungsdurchführung und geben Therapeuten im Clinical-Reasoning Prozess wichtige diagnostische Hinweise, um eine anliegende Behandlungsstrategie abzusichern. Beispielsweise könnten die bei Testdemonstrationen von Aktivitäten festgestellten Defizite (z.B. koordinative Schwächen, Ängstlichkeit, mangelnde Übung, oder Schmerzen durch statische Fehlbelastungen) dazu genutzt werden, entsprechende Behandlungsinhalte aufzugreifen.

Für den physiotherapeutischen Praxisalltag empfiehlt sich aus zeitökonomischen und aus Gründen vergleichbarer Gütekriterien der gemessenen Tests die Anwendung einer modifizierten Form des Polla (ANHANG 7). Die Durchführung des qualitativen Polla auf der Ebene „ja-eingeschränkt-nein“ ist aufgrund des breiten Anforderungscharakters an die Aktivitäten empfehlenswert, denn nach Operationen ist die Mehrzahl der Tests nicht durchführbar und eine Dokumentation daher rasch zu vollziehen. Die modifizierte Version des Polla enthält in der qualitativen Beurteilung nun lediglich zehn Items. Neben dem Aufheben, wurden der Up & Go-, der Lauf- sowie der Treppenlaufen-Test entfernt. Das Treppensteigen ist in Form des dreimaligen treppauf auf eine Stufe im Polla bereits vorhanden.

Zusätzlich wird lediglich einer der gemessenen Tests als ausreichend erachtet, wobei dessen Auswahl zum einen von den räumlichen und materiellen Gegebenheiten einer therapeutischen Einrichtung abhängt. Zum anderen sollte der Treppenlaufen-Test zur Unterscheidung zwischen Diagnosegruppen und der 1-Minute-Up & Go-Test zur Unterscheidung zwischen altersspezifischen Leistungsunterschieden sowie der Lauftest im späteren postoperativen Zeitraum angewandt werden.

Zur Outcome-Messung in der orthopädischen Physiotherapie wird zur Vermeidung von Deckeneffekten ein anspruchsvoller Test dort empfohlen, wo Patienten lediglich mit Funktionsstörungen oder jüngere Patienten in einer späten postoperativen Phase – im Übergang zur sportlichen Aktivität – getestet werden sollen. Ein einbeiniger Dreisprung (Triple-Jump-Test) erfüllt die Kriterien einer im besonderen Maße erhöhten Aktivitätsanforderung. Der Test trifft nach Risberg und Ekeland (1994) noch im späteren postoperativen Zeitraum von über einem halben Jahr Unterscheidungen zwischen dem gesunden und dem betroffenen Bein.

## 9 ZUSAMMENFASSUNG

Die vorliegende Arbeit hatte die Konstruktion und Evaluation eines einfachen, nützlichen und zeitökonomischen *Performance-Test of lower limb activities* (Polla) für den Routineeinsatz im physiotherapeutischen Praxisalltag zum Ziel.

Die *zentralen Fragestellungen* zielten auf die Klärung von Zusammenhängen (Kriteriumsvalidität) zwischen der objektiven Überprüfung von Aktivitäten der unteren Extremität durch den Polla einerseits und der subjektiven Selbsteinschätzung der gesundheitsbezogenen Lebensqualität (SF-36) andererseits. Des Weiteren sollte die Frage nach der Abbildung von Veränderungen zwischen verschiedenen Messzeitpunkten im Therapieverlauf (Änderungssensitivität), geklärt werden. Zusätzliche Betrachtungen zielten auf die Testgütekriterien von Objektivität, Reliabilität und Validität.

*Ausgangspunkt* für eine Ableitung von Aktivitäten für den Polla war die Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF) aus den Domänen *Mobilität* und *Selbstversorgung*. Die Aktivitäten sollten während der physiotherapeutischen Behandlung von Therapeuten direkt überprüfbar sein. Nicht aufgenommen wurden daher Aktivitäten wie das Fortbewegen in verschiedenen Umgebungen oder das Fortbewegen unter Verwendung von Geräten und mit Transportmitteln, ebenso wie Aktivitäten, die nicht im Zusammenhang mit der unteren Extremität standen (z.B. feinmotorischer Handgebrauch). Die Ableitung der Aktivitäten erfolgte gekoppelt mit den in Testbatterien der Literatur häufig angeführten Aktivitäten, den Empfehlungen zu einem Performance-Test nach Bührlen et al. (2002) und einer anschließenden Pilotstudie, in der 12 Tests auf ihre Praktikabilität überprüft wurden. Entfernt wurden der Lauftest als schwieriges Item sowie das Aufheben eines leichten Gegenstandes als leichtes Item. Auf der Grundlage der Ergebnisse wurde der Polla konstruiert und anschließend validiert.

In einer *Längsschnittstudie* wurden zu den Zeitpunkten Therapiebeginn und Therapieende die Veränderungen der Ergebnisse einer sechswöchigen physiotherapeutischen Behandlungsserie an einem Kollektiv von 81 Patienten mit verschiedenen Erhebungsinstrumenten erfasst. Sie umfassten den Polla, einen klinikintern entwickelten Selbsteinschätzungsfragebogen zu Einschränkungen von Aktivitäten und den SF-36 Fragebogen zur gesundheitsbezogenen Lebensqualität. Zielgruppe war ein 19-75-jähriges orthopädisch-traumatologisch orientiertes Patientenkollektiv, das sich nach operativem Eingriff an der unteren Extremität im *Ambulanten Reha-zentrum Kiel der Lubinus Gruppe* zur Therapie anmeldete. Zusätzlich wurden 26 Patienten zu einem Follow-up im Durchschnitt 80 Tage ( $\pm 39$ ) nach Therapieabschluss erneut überprüft.

Die *Ergebnisse* zum Polla zeigen bei geringen Nebenwirkungen der Tests hinsichtlich der Objektivität eine *ausgezeichnete Intrarater-Reliabilität* (n=29) und eine *gute Interrater-Reliabilität* bei allerdings nur drei Beurteilern (n=32).

Die Konsistenzanalyse weist eine *zufrieden stellende Zuverlässigkeit über Cronbachs Alpha* (interne Konsistenz) nach.

Eine überwiegend *mittlere bis hohe Konstruktvalidität* – unter Betrachtung der Summenskalen – ergibt sich aus den statistischen Zusammenhängen der Konstrukte qualitativer Beurteilung im Polla und quantitativer Messung der Tests.

Die *statistische Hypothese* zu vermuteten Zusammenhängen des Polla zur gesundheitsbezogenen Lebensqualität, als einem zentralen Anliegen der Untersuchung (Kriteriumsvalidität), bestätigte sich. Dies vor allem auf der Ebene der verwandten Dimension der Aktivitäten, zwischen der *Körperlichen Funktionsfähigkeit* (SF-36), dem Polla und der Selbsteinschätzung zu Aktivitäten, aber auch im Hinblick auf die *Körperliche Rollenfunktion* und den *Schmerz*. Stabile Zusammenhänge zu allen Dimensionen des SF-36 sind nicht vorhanden. Die in der Literatur beschriebenen geringen bis moderaten Zusammenhänge zwischen den Erhebungsarten von Fremdbeurteilungen in Form von Performance-Tests und Selbsteinschätzungen wurden in dieser Studie bestätigt. Eine Beurteilung von Therapieergebnissen könnte daher grundsätzlich aus unterschiedlicher Sicht erfolgen – aus der Patientenperspektive einerseits und aus der Sicht von Therapeuten oder Ärzten andererseits. Die Benutzung von Performance-Tests in der Outcome-Beschreibung physiotherapeutischer Maßnahmen stellt sich als eine sinnvolle Ergänzung von Selbsteinschätzungsinstrumenten dar.

In der *zweiten statistischen Hypothese* belegt der Polla eindrucksvoll, dass er nach sechs Wochen Behandlung und nach durchschnittlich weiteren drei Monaten Veränderungen mit großen Effekten nachweisen kann.

Die Evaluation *altersspezifischer* und damit *leistungsbedingter* Unterschiede wird im besonderen Maße über den 1-Minute-Up & Go-Test und den 2-Minuten-Gehtest möglich. Sollen Zusammenhänge zum Allgemeinen Gesundheitszustand dargestellt werden, sind diese eher über den ordinalskalierten Polla sowie eingeschränkt über den in der Zeit gemessenen Treppenlaufen-Test möglich. Letzterer bietet sich auch zur Unterscheidung zwischen Diagnosegruppen mit Knieproblematiken unterschiedlichen Ausprägungsgrades an.

Die *Zusammenhänge* des Polla sowohl zur *aktiven Flexion* als auch zur *schmerzhaften Kniebeuge* unterstreichen die Bedeutung der Dokumentation von Beweglichkeitsentwicklungen in der postoperativen Phase. Die Messung der *schmerzhaften Kniebeuge* bietet durch Zusammenhänge zur gesundheitsbezogenen Lebens-

qualität im späteren Genesungszeitraum darüber hinaus eine verbesserte funktionelle Aussage und ist daher der Beweglichkeitsmessung nach der Neutral-Null-Methode vorzuziehen.

Vergleichbare *Testgütekriterien zum Follow-up* mit 26 Patienten durchschnittlich 80 Tage nach Therapieende unterstreichen die Qualität des Polla auch über einen längeren Zeitraum hinweg. Bei insgesamt schwächeren Mittelwertänderungen durch erreichte Leistungsgrenzen ergibt sich für die gemessenen Tests eine geringere Änderungssensitivität gegenüber dem Polla.

Die *Ergebnisse* zum Polla zeigen eine insgesamt gute Absicherung der Teststatistik für die untersuchte Stichprobe von 81 Patienten zu den Zeitpunkten Therapiebeginn und -abschluss hinsichtlich aller Kriterien. Aufgrund der kleinen Fallzahl von 26 Personen zum Follow-up sind die Ergebnisse allerdings kritisch zu bewerten.

Für den *physiotherapeutischen Praxisalltag* empfiehlt sich aus zeitökonomischen Gründen die Anwendung einer modifizierten Form des Polla (ANHANG 7 und 8). Diese Testbatterie ist auf zehn Items reduziert. Es empfiehlt sich die Durchführung nur eines der gemessenen Tests und bietet zusätzlich zur Vermeidung von Deckeneffekten einen hoch anspruchsvollen einbeinigen Sprungtest zur Aktivitätsüberprüfung an (Triple-Jump-Test).

*Ausstehende Fragestellungen* könnten sich mit der Änderungssensitivität zum Polla über Zeiträume von mehr als 6 Monaten befassen. Es sollten bei eng begrenzten Krankheitsbildern die Unterscheidungsfähigkeiten zwischen verschiedenen Altersstufen und Diagnosen herausgearbeitet werden, um die Testbatterie im Vergleich unterschiedlicher Behandlungsmethoden einzusetzen. Insgesamt zeigt der Polla gute Ergebnisse hinsichtlich der Interrater-Reliabilität bei allerdings nur drei Beurteilern. Hier empfiehlt sich eine zusätzliche Überprüfung. Schließlich steht die Entwicklung einer Testbatterie für die obere Extremität noch aus, die aus Gründen des Umfangs an dieser Stelle nicht aufgenommen werden konnte.

## LITERATUR

- Ageberg, E., Zätterström, R. & Moritz, U. (1998). Stabilometry and one-leg hop test have high test-retest reliability. *Scand. J. Med. Sci. Sports*, 8, 198-202.
- Altenhöner, T., Leppin, A., Grande, G. & Romppel, M. (2001). Die Bedeutung von Patienteneinstellungen für die Erreichung subjektiver Rehabilitationsziele in der kardiologischen Rehabilitation. In Verband Deutscher Rentenversicherungsträger (VDR, Hrsg.), *10. Rehabilitationswissenschaftliches Kolloquium. Wissenstransfer zwischen Forschung und Praxis*. (12. bis 14. März in Halle/Saale. Tagungsband, S. 232-233.). Frankfurt am Main: Postverlagsort.
- Antonovsky, A. (1997). *Salutogenese: zur Entmystifizierung der Gesundheit*. Tübingen: DGVT.
- Antonovsky, A. (1993). Gesundheitsforschung versus Krankheitsforschung. In A. Franke & M. Broda (Hrsg.), *Psychosomatische Gesundheit. Versuch einer Abkehr vom Pathogenese-Konzept*. (S. 3-14). Tübingen: DGVT.
- Antonovsky, A. (1987). *Unraveling the mystery of health*. San Francisco: Jossey-Bass.
- AWMF, ÄZQ (Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften, AWMF; Ärztliches Zentrum für Qualität in der Medizin, ÄZQ) (2005). *Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI)*. Fassung 2005/2006. Zugriff am 12.3.07 unter: <http://www.awmf-online.de>
- Auracher, M., Müller, M., Hörterer, H. (2002). Der MPH-Therapiescore – Eine Semiobjektive Möglichkeit der Therapieverlaufsdokumentation. *Zeitschrift für Physiotherapeuten*, 54 (2), 247-249.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2006). *Multivariate Analysemethoden: eine anwendungsorientierte Einführung*. (11. überarb. Aufl.). Berlin: Springer.
- Barber, S. D., Noyes, F. R., Mangine, R. E., McCloskey, J. W. & Hartmann, W. (1990). Quantitative assessment of functional limitations in normal and anterior cruciate ligament-deficient knees. *Clin. Orth. Rel. Res.*, 255 (6), 204-214.
- Bäumer, F. (2007). Wie viel Professionalisierung müsste es denn sein? Einschätzungen und Meinungsbild zur Professionalisierung in der Physiotherapie – eine Fragebogenuntersuchung. *Zeitschrift für Physiotherapeuten*, 59 (2), 138-148.
- Bäumer, F. (2006). Was bewegt die Physiotherapie? – Wunschvorstellungen von einer zukünftigen Physiotherapie in Deutschland. *Zeitschrift für Physiotherapeuten*, 58 (4), 314-324.
- Bellamy, N. (1993). *Musculoskeletal Clinical Metrology* (S. 205-342). London: Kluwer Academic Publishers.
- Bellamy, N. & Buchanan, W. W. (1988). Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes following total hip or knee arthroplasty in osteoarthritis. *Journal of Orthopaedic Rheumatology* 1, 95-108.
- Bengel, J. & Koch, U. (2000). Definition und Selbstverständnis der Rehabilitationswissenschaften. In J. Bengel & U. Koch (Hrsg.), *Grundlagen der Rehabilitationswissenschaften* (S. 1-18). Berlin: Springer.
- Bengel, J. (2001). *Was hält Menschen gesund? Antonovskys Modell der Salutogenese – Diskussionsstand und Stellenwert*. Bundeszentrale für gesundheitliche Aufklärung (BZgA). (Band 06, PDF-Dokument). Zugriff am: 03.03.07 unter: [www.bzga.de](http://www.bzga.de)
- Beyer, H.-M. (2001/2002). *Heilmittelkatalog, Heilmittel der physikalischen Therapie. Nachschlagewerk zur Ermittlung der verordnungsfähigen Heilmittel entsprechend der Indikatoren*. Ludwigsburg: Intellimed.

- Beyer, H.-M. (2004). *Heilmittelkatalog, Heilmittelkatalog der Physikalischen Therapie*. Nachschlagewerk zur Ermittlung der verordnungsfähigen Heilmittel entsprechend der Indikationen. Intelimed: Ludwigsburg.
- Biefang, S. & Schuntermann, M. F. (2000). Diagnostik und Assessment in der Rehabilitation. In J. Bengel & U. Koch (Hrsg.), *Grundlagen der Rehabilitationswissenschaften* (S. 104-120). Berlin: Springer.
- Biefang, S., Potthoff, P. & Schliehe, F. (1999). *Assessmentverfahren für die Rehabilitation*. (S. 15-25). Göttingen: Hogrefe.
- Biefang, S., Birkner, B., Thien, U., Härtel, U. & Bullinger, M. (1997). Harmonisierung der Messung von Outcomes, Prädiktoren und Kosten sowie Prüfung geschlechtsspezifischer Unterschiede in der rehabilitationswissenschaftlichen Forschung. *Rehabilitation* (36), 213-223.
- Biefang, S. & Potthoff, P. (1996). Ein Konzept für die Messung von Reha-Outcomes im Rahmen des Qualitätsmanagements. In Verband Deutscher Rentenversicherungsträger (VDR, Hrsg.), *Evaluation in der Rehabilitation. 6. Rehabilitationswissenschaftliches Kolloquium*. (4.-6. März in Bad Säckingen, DRV-Schriften, Band 6, Tagungsband, S. 69.) Berlin: Springer.
- Bös, K. (Hrsg.) (2001). *Handbuch Motorische Tests. (2. vollständig überarb. und erw. Aufl.)*. Göttingen: Hogrefe.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler*. (4. überarb. Aufl.). Heidelberg: Springer.
- Bortz, J. & Döring, N. (1995). *Forschungsmethoden und Evaluation: für Sozialwissenschaftler*. (2. überarb. Aufl., S. 163 -174). Heidelberg: Springer.
- Briggs, S. R. & Cheek, J. M. (1986). The Role of Factor Analysis in the Development and Evaluation of Personality Scales. *Journal of Personality and Social Psychology*, 54 (1), 106-148.
- Broda, M. & Beckmann, U. (2000). Dokumentation und Katamnestic. In J. Bengel & U. Koch (Hrsg.), *Grundlagen der Rehabilitationswissenschaften* (S. 361-376). Berlin: Springer.
- Brügger, A. (2000). *Lehrbuch der funktionellen Störungen des Bewegungssystems: das neurale Szenario der Schmerzen und Behinderungen des Bewegungssystems*. Zollikon: Brügger.
- Brosseau, L., Yonge, K. A., Robinson, V., Marchand, S., Judd, M., Wells, G. & Tugwell, P. (2003). Thermotherapy for treatment of osteoarthritis. *Cochrane Database. Syst. Rev.*, 32. 527-534.
- Brosseau L, Milne S., Wells G., Tugwell, P., Robinson, V., Casimiro, L., Pelland, L., Noel, M. J., Davis, J. & Drouin, H. (2004). Efficacy of continuous passive motion following total knee arthroplasty: a metaanalysis. *J. Rheumatol.*, 31(11), 2251-64.
- Buckup, K. (2000). *Klinische Tests an Knochen, Gelenken und Muskeln. Untersuchungen – Zeichen – Phänomene* (2. überarb. Aufl.). Stuttgart: Thieme.
- Bühl, A., Zöfel, P. (2005). *SPSS 12. Einführung in die moderne Datenanalyse unter Windows* (9. Aufl.). München: Pearson Studium.
- Bühlren, B., Gerdes, N. Jäckel, W. H. (2002). *Die Operationalisierung von Therapiezielen für die Überprüfung der Zielerreichung in der Rehabilitation* (Abschlussbericht). Bad Säckingen: Hochrhein-Institut für Rehabilitationsforschung e.V.
- Bullinger, M., Morfeld, M., Kohlmann, T., Nantke, J., van den Bussche, H., Dodt, B., Dunkelberg, S., Kirchberger, I., Krüger-Bödecker, A., Lachmann, A., Lang, K., Mathis, C., Mittag, O., Peters, A., Raspe, H.-H. & Schulz, H. (2003). Der SF-36 in der rehabilitationswissenschaftlichen Forschung. Ergebnisse aus dem Norddeutschen Verbund für Rehabilitationsforschung (NVRF) im Förderschwerpunkt Rehabilitationswissenschaften. *Die Rehabilitation, Leitlinien-Rehabilitation*, 42 (4). 218-225.
- Bullinger, M. & Ravens-Sieberer, U. (2000). Indikatoren des Rehabilitationsergebnisses. In J. Bengel & U. Koch (Hrsg.), *Grundlagen der Rehabilitationswissenschaften* (S. 305-320). Berlin: Springer.

- Bullinger, M. & Kirchberger, I. (1998). *Der deutsche SF-36 Health Survey-Fragebogen zum Gesundheitszustand: Handbuch für die deutschsprachige Fragebogenversion*. Göttingen: Hogrefe.
- Bullinger, M. (1996). Erfassung der gesundheitsbezogenen Lebensqualität mit dem SF-36 Health Survey. *Die Rehabilitation* 35, XVII–XXX.
- Cabri, J. (2001). Testverfahren am Bewegungsapparat. In F. van den Berg, *Angewandte Physiologie. Therapie, Training, Tests* (S. 196-240). New York: Thieme.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2. Aufl.). Hillsdale. New-York: Erlbaum.
- Clarke, H. H. (1976). *Application of measurement*. New-York. S. 27.
- Creditor, M. C. (1993). Hazards of hospitalization of the elderly. *Ann. Intern. Med.*, 118, 219-223.
- Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin* 52, 281.
- Cronbach, L. J. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, 16, 297-334.
- Dahl, G. (1971). Zur Berechnung des Schwierigkeitsindex bei quantitativ abgestufter Aufgabenbewertung. *Diagnostica*, 17, 139-142.
- Daltroy, L. H., Phillips, C. B., Eaton H. M., Larson, M. G., Partridge, A. J., Logigian, M. & Liang, M. H. (1995). *Physical Capacity Evaluation (PCE)*. Kontakt: Multipurpose Arthritis Center / Pbb-B2 Brigham and Women's Hospital 75 Francis St., Boston, MA 02115-6195, USA.
- De Bie, R. (1998a). Die Notwendigkeit von Effektivitätsstudien als Grundlage für die Physiotherapie, Teil 1. *Manuelle Therapie*, (2), 61-65.
- De Bie, R. (1998b). Die randomisierte kontrollierte Studie in der Physiotherapie, Teil 2. *Manuelle Therapie*, (2), 131-137.
- Debrunner, H. U. (1971). *Gelenkmessung (Neutral-Null-Methode), Längenmessung, Umfangsmessung*. Bern: Bulletin der Schweizerischen Arbeitsgemeinschaft für Osteosynthesefragen.
- Deck, R. & Röckelein, E. (1999). Zur Erhebung soziodemographischer und sozialmedizinischer Indikatoren in den rehabilitationswissenschaftlichen Forschungsverbänden. In Verband Deutscher Rentenversicherungsträger (DRV, Hrsg.), *Förderschwerpunkt „Rehabilitationswissenschaften“*. Empfehlungen der Arbeitsgruppen "Generische Methoden", "Routinedaten" und "Reha-Ökonomie". (DRV-Schriften Band 16. Sonderausgabe, S. 84-102). Frankfurt am Main: Postverlagsort.
- Dorenburg, U. & Tiefensee, J. (2000). Qualitätssicherung in der medizinischen Rehabilitation. In J. Bengel & U. Koch (Hrsg.), *Grundlagen der Rehabilitationswissenschaften* (S. 197-213). Berlin: Springer.
- Dorfmueller-Kuechlin, S., Schlennstedt, D. & Voigt-Radloff, S. (1998). Das Physiotherapeutische Assessment. *Krankengymnastik – Zeitschrift für Physiotherapeuten*, 50, 1711-1723.
- Deutsche Vereinigung für Sportwissenschaft (DVS) (2006). Zugriff am 7.3.07 unter: <http://www.sportwissenschaft.de/index.php?id=6>
- DGQ & ZVK (Deutsche Gesellschaft für Qualität e.V. & Deutscher Verband für Physiotherapie – Zentralverband der Physiotherapie / Krankengymnasten (Hrsg.) (2003). *Qualitätsmanagement in physiotherapeutischen Einrichtungen*. Berlin: Beuth.
- Deutsches Institut für Medizinische Dokumentation und Information (DIMDI) (2005). Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit. Zugriff am 25.3.08 unter: <http://www.dimdi.de/dynamic/de/klassi/downloadcenter/icf/endaussage/>

- DIN (Deutsches Institut für Normierung) (1992). *DIN ISO 9004*, Teil 2. Berlin: Beuth.
- Engel, G. L. (1979). Die Notwendigkeit eines neuen medizinischen Modells: Eine Herausforderung der Biomedizin. In H. Keupp (Hrsg.), *Normalität und Abweichung. Fortsetzung einer notwendigen Kontroverse* (S. 123-143). München: Urban & Schwarzenberg.
- Esser, H., Chr. (2006). Die Zukunft des Gesundheitssystems. *Mitteilungen des Zentralverbandes. Deutscher Verband für Physiotherapie (ZVK)*, (8), 4-5.
- Faller, H., Haaf, H. G., Kohlmann, T., Löschmann, Ch., Maurischat, C., Petermann, F., Schulz, H. & Zwingmann, C. (1999). Orientierungshilfen und Empfehlungen für die Anlage, Durchführung und Interpretation von Studien in der Rehabilitationsforschung. In Verband Deutscher Rentenversicherungsträger (DRV, Hrsg.), *Förderschwerpunkt „Rehabilitationswissenschaften“. Empfehlungen der Arbeitsgruppen "Generische Methoden", "Routinedaten" und "Reha-Ökonomie"*. (DRV-Schriften, Band 16. Sonderausgabe, S. 9-51). Frankfurt am Main: Postverlagsort.
- Faltermaier, T. (1994). *Gesundheitsbewußtsein und Gesundheitshandeln*. Weinheim: Beltz
- Farin, E., Fleitz, A. & Follert, P. (2006). Entwicklung eines ICF-orientierten Patientenfragebogens zur Erfassung von Mobilität und Selbstversorgung. *Physikalische Medizin, Rehabilitationsmedizin, Kurortmedizin*, 16 (4), 197-211.
- Fransen, J. & de Bruin, E. D. (2000). Evidence Based Medicine in der Manuellen Therapie. *Manuelle Therapie*, (4), 95-102.
- Freud, A. (1959). *Das Ich und die Abwehrmechanismen*. München: Kindler.
- Friess, J. F., Spitz, P., Kraines, R. G. & Holman, H. R. (1980). Measurement of patient outcome in arthritis. *Arthritis Rheum.*, 23, 137-145.
- Filipp, S. H., Aymanns, P. (1987). Die Bedeutung sozialer und personaler Ressourcen in der Auseinandersetzung mit kritischen Lebensereignissen. *Z. Klien. Psychol.* 16, 1-14.
- Froböse, I. & Nellessen, G. (1998). *Training in der Therapie. Grundlagen und Praxis*. Wiesbaden: Ullstein.
- Geissner, E. (1996). *Die Schmerzempfindungs-Skala (SES). Handanweisung*. Göttingen: Hogrefe.
- Gerdes, N. (2006). Zielorientierung in der Ergebnismessung. 15. *Rehabilitationswissenschaftliches Kolloquium. Rehabilitation und Arbeitswelt*. 13. bis 15. März 2006 in Bayreuth (DRV-Schriften, Tagungsband, S. 111-112).
- Gerdes, N. & Weis, J. (2000). Zur Theorie der Rehabilitation, in J. Bengel & U. Koch (Hrsg.), *Grundlagen der Rehabilitationswissenschaften* (S. 41-68). Berlin: Springer.
- Gerdes, N., Jäckel, W. H. (1995). Der IRES-Fragebogen für Klinik und Forschung. *Die Rehabilitation*, 34, XIII-XXIV.
- Glatz, A., Anneken, V., Heipertz, W. Schüle, K., Mozdzanowski, M. & Schian, H. M. (2006). Der Beitrag des FCE-Assessments ERGOS Work Simulator zur arbeits- und sozialmedizinischen Leistungsbeurteilung - Erste Ergebnisse [Abstract]. In Deutscher Rentenversicherungsbund (DVR, Hrsg.), *Rehabilitation und Arbeitswelt - Herausforderungen und Strategien*. 15. *Rehabilitationswissenschaftliches Kolloquium*. 13. bis 15. März in Bayreuth. (DRV-Schriften, Band 64, Tagungsband, S. 136-138).
- Guralnik, J. M., Simonsick, E. M., Ferrucci, L., Glynn, R. J., Berkman, L. F., Blazer, D. G., Scherr P. A. & Wallace, R. B. (1994). A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol. Mar.*; 49 (2), M85-94.
- Guralnik, J. M., Branch, L. G., Cummings, S. R. & Curb, J. D. (1989). Physical Performance Measures in Aging Research. *Journal of Gerontology: Medical Sciences*, 44, M141–M146.
- Guyatt, G. H., Jaeschke, R., Feeny, D. H. & Patrick, D. L. (1996). Measurement in clinical trials: Choosing the right approach. In B. Spilker (Ed.), *Quality of life and pharmacoconomics in clinical trails* (2nd ed., pp. 41-48). Philadelphia: Lippincott-Raven.

- Haaf, H.-G., Schliehe, F. (2000). Zur Situation der Rehabilitationsforschung: Stand und Bedarf. In: J. Bengel & U. Koch (Hrsg.) *Grundlagen der Rehabilitationswissenschaften* (S. 19-40). Heidelberg: Springer.
- Haan, N. (1977). *Coping and Defending*. New-York: Academic Press.
- Hallmann, U. (1999). Effizienznachweis in der Physiotherapie. *Zeitschrift für Physiotherapeuten*, 51 (10), 1744-1746.
- Harding, V. R., Williams, A. C., Richardson, P. H., Nicholas, M. K., Jackson, J. L., Richardson, I. C. H. & Pither, C. E. (1994). The development of a battery of measures for assessing physical functioning of chronic pain patients. *Pain*, 58, 367-375.
- Harris, W. H. (1969). Traumatic arthritis of the hip after dislocation and acetabular fractures: Treatment by Mold arthroplasty. An end result study using a new method of result evaluation. *J. Bone Jt. Surg.* 51-A, 737-755.
- Heim, E. (1988). Coping und Adaptivität. Gibt es ein geeignetes oder ungeeignetes Coping? *Psychother. Psychosom. Med. Psych.*, 38 (2), 8-18.
- Herrmann, T. (1969). *Lehrbuch der empirischen Persönlichkeitsforschung*. Göttingen: Hogrefe.
- Höppner, H. & Borgetto, B. (2007). Den Wandel mitgestalten. Zukunftsorientierung: Prävention und Gesundheitsförderung. Bedarf, Ansätze und Chancen für eine qualifizierte Physiotherapie. *Zeitschrift für Physiotherapeuten*, 59 (7), 565-569.
- Hollmann, W. & Hettinger, T. (1980). *Sportmedizin. Arbeits- und Trainingsgrundlagen* (2. neu bearb. Aufl.). Stuttgart, New-York: Schattauer.
- Hüter-Becker, A. (2002). Das "Neue Denkmodell": Ein integrativer Ansatz in der Physiotherapie. In: Hüter-Becker: *Lehrbuch zum Neuen Denkmodell der Physiotherapie, Band 1 Bewegungssystem*. Stuttgart-New York: Thieme.
- Hüter-Becker, A. (1997). Ein neues Denkmodell für die Physiotherapie. *Krankengymnastik*, 49 (4), 565-569.
- HVG e.V. (Hochschulverbund Gesundheitsfachberufe) (2007). Hochschulen schließen sich zusammen. *Zeitschrift für Physiotherapeuten*, 59 (4). S. 404.
- Igl, W., Zwingmann, C., Faller, H., Beutel, M., Beyer, W., Bischoff, C., Fritschka, E., Gustson, D., Hillert, A., Knickenberg, R., Kühn, A., Kulick, B., Lamprecht, F., Nübling, R., Rief, W., Schmidt, J., Schmitz, B., Vauth, R., Vogel, H., Wagner, R., Wallesch, C.-W. & Wittmann, M. (2006). Änderungssensitivität von generischen Patientenfragebogen – Ergebnisse einer verbundübergreifenden Reanalyse. *Physikalische Medizin, Rehabilitationsmedizin, Kurortmedizin*, 16 (2), 69-81.
- Insall, J. N., Dorr, L. D., Scott, R. & Scott, W. N. (1989). Rationale of the knee society clinical rating system. *Clin. Orthop.*, 248, 13-14.
- IFK (Bundesverband selbstständiger Physiotherapeuten) (2003). *Positionspapier der AG MTG zur Akademisierung der Medizinalfachberufe in der Therapie und Geburtshilfe*. Zugriff am 7.3.07 unter: [http://www.agmtg.de/\\_pdf/Positionspapier.pdf](http://www.agmtg.de/_pdf/Positionspapier.pdf)
- IGPTR (Interessengemeinschaft Physiotherapie Rehabilitation) (2006). Manueller Muskeltest. Zugriff am 2.12. 06 unter: <http://www.igptr.ch/pdf/Manueller%20Muskeltest.pdf>
- IQPR (Institut für Qualitätssicherung in Prävention und Rehabilitation GmbH). Zugriff am 3.12.06 unter: <http://www.assessment-info.de/assessment/seiten/default.asp>
- Israel, S. (1992) Die Problematik körperbezogener Normwerte bei Menschen nach dem so genannten Höchstleistungsalter. *Sport Praxis*. 33, 1, 37-39; 2, 35-39.
- Isernhagen, S. J. (1988). Functional Capacity Evaluation. In S. J. Isernhagen (ed), *Work injury: Management and prevention* (S. 139-174). Aspen Publishers: Gaithsburg.

- Jensen, J. E., Conn, R., Hazelrigg, G. & Hewett, J. (1984). Systematic evaluation of acute knee injuries. *Clin. Sports Med.*, 4, 295.
- Jensen, M. P., Karoly, P. & Braver, S. (1986). The measurement of clinical pain intensity: A comparison of six methods. *Pain*, 27 (1), 117-26.
- Kaiser, H., Kersting, M. & Schian, H. M. (2000). Der Stellenwert des Arbeitssimulationsgerätes ERGOS als Bestandteil der leistungsdiagnostischen Begutachtung. *Rehabilitation*, 39, 175-184.
- Katz, S. & Akpom, C. A. (1974). Index of ADL. *Medical Care*, 14, 116-118.
- Kelly-Hayes, M., Jette, A. M., Wolf, P. A., D'Agostino, R. B. & Odell, P. M. (1992). Functional limitations and disability among elders in the Framingham Study. *Am. J. Public Health*, 841-845.
- Kennedy, D., Stratford, P. W., Pagura, S. M., Walsh, M. & Woodhouse, L. J. (2002). Comparison of gender and group differences in self-report and physical performance measures in total hip and knee arthroplasty candidates. *J. Arthroplasty*, 17 (1). 70-77.
- Keitel, W., Hoffmann, H., Weber, G. & Krieger, U. (1993). *Bewegungsfunktionstest (BFT)*. Hrsg. Giesela Westhoff. Göttingen Hogrefe.
- Klemme, B., Geuter, G. & Willimczik, K. (2007). Physiotherapie – über eine Akademisierung zur Profession. *Physioscience*, 3, 80-87.
- König, A., Kirschner, S., Wahlter, M., Böhm, D. & Faller, H. (2000). Kulturelle Adaptation, Praktikabilitäts- und Reliabilitätsprüfung des Funktionsfragebogens Bewegungsapparat (SMFA-D) *Zeitschrift für Orthopädie*, (138), 295-301.
- Koch, U. & Buschmann-Steinhage, R. (2004). Zum Verständnis und zu den Voraussetzungen der Rehabilitationswissenschaften in Deutschland. *Deutsche Rentenversicherung* 5, 263-272.
- Koch, U., Schliehe, F. & Aufderheide, E. (1998). Stand und Entwicklung der rehabilitationswissenschaftlichen Forschung. *Die Rehabilitation*, 37, Suppl. 2, 66-70.
- Koch, U., Lehmann, D. & Morfeld, M. (2007). Bestandsaufnahme und Zukunft der Rehabilitationsforschung in Deutschland. *Rehabilitation*, 46, 127-144.
- Kohlmann, T., Ludwig, F.J., Daalman, H.-H., Kensy, E. & Hekler, J. (2004a). Entwicklung eines Fragebogens zur Messung der Funktionskapazität bei Patienten mit Arthrosen der Hüft- und Kniegelenke. Bericht. Gefördert aus den Mitteln des Vereins zur Förderung der Rehabilitationsforschung in Schleswig-Holstein (vffr-Projekt Nummer 44). Universität Greifswald.
- Kohlmann, T. (2004b). Patientenbezogene Assessments in der orthopädischen Rehabilitation-Konzept und Methodik am Beispiel des Funktionsfragebogens Hannover FFbH [Abstract]. In Verband Deutscher Rentenversicherungsträger (VDR, Hrsg.), *Selbstkompetenz – Weg und Ziel der Rehabilitation. 13. Rehabilitationswissenschaftliches Kolloquium. 8. bis 10. März 2004 in Düsseldorf*. (DRV-Schriften, Band 52, Tagungsband, S. 68-69).
- Kohlmann, Knahr, Krysylin-Exner, Heinrichs & Peschel (1996). FFb-H-OA – Funktionsfragebogen Hannover-Osteoarthrose; DRK Schmerz-Zentrum Mainz. Zugriff am 04.08.07 unter: [http://www.schmerz-zentrum.de/downloads/eqa\\_scales/ffbhoa2.pdf](http://www.schmerz-zentrum.de/downloads/eqa_scales/ffbhoa2.pdf)
- Kohlmann, T., Bullinger, M., Hunt, S.M. & McKenna, S.P. (1992). *Zur Messung von Dimensionen der subjektiven Gesundheit: Die Deutsche Version des "Nottingham Health Profile" (NHP)*. Lübeck (Arbeitsbericht).
- Kolster, B. & Ebel-Paprotny, G. (1996). *Leitfaden Physiotherapie. Befundtechniken, Behandlung, Rehabilitation* (2. neu bearb. Auflage). Stuttgart: Fischer.
- Kool, J. & de Bie, R. (2001). *Der Weg zum wissenschaftlichen Arbeiten* (S. 66). Stuttgart: Thieme.
- Krämer, K. L. & Maichl, F. P. (1993). *Scores, Bewertungsschemata und Klassifikationen in Orthopädie und Traumatologie*. Stuttgart: Thieme.
- Lazarus, R. S. & Folkman, S. (1984). *Stress, appraisal and coping*. New-York: Springer.

- Lee, C. E., Simmonds, M. J. & Novy, D. M. (2001). Selfreports and clinician-measured physical function among patients with low back pain: a comparison. *Arch. Phys. Med. Rehabil.*, 82 (2), 227-231.
- Lequesne, M. G., Mery, C., Samson, M. & Gerard, P. (1987). Indexes of severity for osteoarthritis of the hip and knee. Validation – Value in comparison with other assessment tests. *Scan. J. Rheumatology, Suppl.*, 65, 85-89.
- Leonhardt, R. (2004). Effektgrößenberechnung bei Interventionsstudien. *Rehabilitation*, 43 (4), 241-246.
- Lephart, S., Pincivero, D. M., Giraldo, J. L. & Fu, F. H. (1994). The role of proprioception in the management and rehabilitation of athletic injuries. *Am. J. Sports Med.*, 25 (1), 130-137.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Aufl.). Weinheim: Psychologische Verlagsunion.
- Lin, Y. C., Davey R. C. & Cochrane, T. (2001). Tests for physical function of the elderly with knee and hip. *Scandinavian journal of medicine & science in sports*, 11, (5), 280-286.
- Liow, R. Y. L., Walker, K., A., Wajid, M., Gurinder, B. & Lennox, C. M. E. (2000). The reliability of the American Knee Society Score. *Acta Orthop. Scand.*, 54 (1), 603-608.
- Löschmann, C. & Bengel, J. (2000). Implementierung und Umsetzung von Ergebnissen der Rehabilitationsforschung, in Verband Deutscher Rentenversicherungsträger (Hrsg.), *Förderschwerpunkt „Rehabilitationswissenschaften“*. Empfehlungen der Arbeitsgruppen "Generische Methoden", "Routinedaten" und "Reha-Ökonomie". (DRV-Schriften Band 16, Sonderausgabe, S. 525-536). Frankfurt am Main: Postverlagsort.
- Lüdtke, K. (2000). Qualitätskontrolle durch Evidence based Practice. *Zeitschrift für Physiotherapeuten*, 52 (8), 1325-1328.
- Ludwig, F. J., Melzer, C., Grimmig, H. & Daalman, H. H. (2002). Kulturelle Adaptation des Lequesne-Index für Hüft- und Kniegelenkerkrankungen im deutschen Sprachraum. *Rehabilitation*; 41, 249-257
- Lysholm, J. & Gillquist, J. (1982). Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. *Am. J. Sports Med.*, 10, 150-154.
- Marshall, J. L.; Fetto, J. F. & Botero, P. M. (1977). Knee ligament injuries: A standardized evaluation method. *Clin. Orthop.*, 115-129.
- Mahomed, N. N., Liang, M. H., Cook, E. F., Daltroy, L. H., Fortin, P. R., Fossel, A. H. & Katz, J. N. (2006). The importance of patient expectations in predicting functional out-comes after total joint arthroplasty. *J. Rheumatol.*, 29 (6), 1273-1279.
- Mahony, F. I. & Barthel, D. W. (1965). Functional evaluation: The Barthel Index. *Maryland State Medical Journal*, 14, 61-65.
- Maier-Riehle, B., Gerdes, N., Protz, W. & Jäckel, W. H. (1998). Übereinstimmung und Unterschiede zwischen Beurteilern bei einem Peer-Review-Verfahren. *Gesundheitswesen*, 60, 290-296.
- Mathias, S., Nayak, U. S. L. & Isaacs, B. (1986). Balance in elderly patient: The "Get-up and Go" test. *Arch. Phys. Med. Rehabil.*, 67, 387.
- McGill, M. (1994). Functional progression in rehabilitation. In W. E. Prentice, *Rehabilitation techniques in Sports Medicine*. St. Louis.
- Müller, W., Biedert R., Hefti, F., Jakob, R. P., Munzinger, U. & Stäubli, H. U. (1988). OAK knee evaluation. A new way to assess knee ligament injuries. *Clinical Orthopaedy*, 232, 37-50.
- Müller, H., Franke, A., Schuck, P. & Resch, K. L. (2001). Die Anwendung des SF-36 in Kliniken und resultierende Verzerrungen der Ergebnisqualität. In M. F. Schuntermann, F. Schliehe (Hrsg.), *Wissenstransfer zwischen Forschung und Praxis*. 10. *Rehabilitationswissenschaftli-*

- ches Kolloquium. (12. bis 14. März 2001 in Halle/Saale, Tagungsband, S. 98-100). Frankfurt am Main: Postverlagsort.
- Muthny, F. A., Bullinger, M. & Kohlmann, T. (1999). Variablen und Erhebungsinstrumente in der rehabilitationswissenschaftlichen Forschung – Würdigung und Empfehlungen. In Verband Deutscher Rentenversicherungsträger (DRV, Hrsg.), *Förderschwerpunkt „Rehabilitationswissenschaften“*. Empfehlungen der Arbeitsgruppen "Generische Methoden", "Routinedaten" und "Reha-Ökonomie". (DRV-Schriften, Band 16. Sonderausgabe, S. 84-102). Frankfurt am Main: Postverlagsort.
- Najman, J. M. & Levine, S. (1981). Evaluating the impact of medical care and technologies on the quality of life: a review and critique. *Soc. Sci. Med.*, 15, 107-115.
- Nübling, R. & Schmidt, J. (2000). Methodische Grundlagen der Ergebnisevaluation. In J. Bengel & U. Koch (Hrsg.), *Grundlagen der Rehabilitationswissenschaften* (S. 323-346). Berlin: Springer.
- Nübling, R. & Schmidt, J. (1998). Qualitätssicherung in der Psychotherapie – Grundlagen, Realisierungsansätze, künftige Aufgaben. In A.-R. Laireiter & H. Vogel (Hrsg.), *Qualitätssicherung in der Psychotherapie – Ein Werkstattbuch*. Tübingen: Dgvt.
- Oetiker-Streit, D. (1999). Qualitätsmanagement in der physiotherapeutischen Arbeit. *Zeitschrift für Physiotherapeuten*, 51 (10), 1747-1756.
- Patrick, D. L. & Erickson, P. (1993). *Health status and health policy: Allocating resources to health care*. New York: Oxford University Press.
- Pfeifer, K. (1996). *Bewegungsverhalten und neuromuskuläre Aktivierung nach Kniebandverletzungen*. Neu-Isenburg.
- Pfingsten, M. & Hildebrandt, J. (1995). Degenerative und andere nicht-entzündliche Erkrankungen der Haltungs- und Bewegungsorgane. In F. Petermann (Hrsg.), *Verhaltensmedizin in der Rehabilitation*. Göttingen: Hogrefe.
- Podsiadlo, D. & Richardson, S. (1991). The Timed "Up & Go": A Test of basic functional mobility for frail elderly persons. *American Geriatrics Society*, 39, 142-148.
- Protz, W., Gerdes, N., Maier-Riehle, B. & Jäckel, W. H. (1996). Therapieziele und ihre Messung in der medizinischen Rehabilitation: Erste Ergebnisse des Qualitätsscreenings. In Verband Deutscher Rentenversicherungsträger (VDR, Hrsg.), *Evaluation in der Rehabilitation. 6. Rehabilitationswissenschaftliches Kolloquium*. (4.- 6. März in Bad Säckingen, DRV-Schriften, Band 6, Tagungsband, S. 98-99) Berlin: Springer.
- Raspe, H. H., Hagedorn, U., Kohlmann, T. & Mattussek, S. (1990). Der Funktionsfragebogen Hannover (FFbH): Ein Instrument zur Funktionsdiagnostik bei polyartikulären Gelenkerkrankungen. In J. Siegrist (Hrsg.), *Wohnortnahe Betreuung Rheumakrankter. Ergebnisse sozialwissenschaftlicher Evaluation eines Modellversuchs*. 164-182.
- Reichel, H.-S. (2005). *Das PNF-Konzept: Prinzip-Methode-Technik; Lernprogramm propriozeptive neuromuskuläre Faszilitation* (4. unveränd. Aufl.). Stuttgart, New-York: Thieme.
- Reitman, R. D., Emerson, R. H., Higgins, L. L. & Tarbox, T. R. (2003). A multimodality regimen for deep venous thrombosis prophylaxis in total knee arthroplasty. *J. Arthroplasty*, 18 (2), 161-8.
- Rejeski, W. J., Ettinger, W. H., Schumaker, S., James, P., Burns, R. & Elam, J. T. (1995). Assessing performance-related disability in patients with knee osteoarthritis. *Osteoarthritis and Cartilage*, 3 (3), 157-167.
- Reuben, D. B. & Siu, A. L. (1990). An objective measure of physical function of elderly patients. The physical performance test. *Journal of the American Geriatrics Society*, 38, 1105-1112.
- Reuben, D. B., Siu, A. L. & Kimpau, S. (1992). The predictive validity of self-report and performance-based measures of function and health. *J. Gerontol. Med. Sci.*, (47), M106-110.
- Risberg, M. A. & Ekeland, A. (1994). Assessment of functional tests after anterior cruciate ligament surgery. *JOSPT*, 19 (4), 212-217.

- Rietz, I., Josenhans, J., Höder, J. & Arlt, A. (2003). Auswirkungen von Zieldiskrepanzen auf den Reha-Erfolg bei stationären Rückenschmerzpatienten. In Verband Deutscher Rentenversicherungsträger (VDR, Hrsg.), 12. *Rehabilitationswissenschaftliches Kolloquium. Rehabilitation im Gesundheitssystem*. (10. bis 12. März in Bad Kreuznach. DRV-Schriften, Tagungsband. S. 242-243).
- Roos, E. (1999). Knee Injury and Knee Osteoarthritis Development, evaluation and clinical application of patient-relevant questionnaires. Zugriff am 5.2.07 unter: <http://www.koos.nu>
- Sackett, D. J. & Kunz, R. (1999). *Evidenzbasierte Medizin (EBM-Umsetzung und Vermittlung)*. Deutsche Ausgabe. München: Zuckschwerdt.
- Sangha, O. & Stucki, G. (1997). Patientenzentrierte Evaluation der Krankheitsauswirkungen bei muskuloskelettalen Erkrankungen: Übersicht über die wichtigsten Outcome-Instrumente. *Zeitschrift für Rheumatologie*, 56, 322-333.
- Sayers, S. P., Guralnik, J. M., Newman, A. B., Brach, J. S. & Fielding, R. A. (2006). Concordance and discordance between two measures of lower extremity function: 400 meter self-paced walk and SPPB. *Aging Clin. Exp. Res.*, 18 (2), 100-106.
- Schämann, A. (2005). *Akademisierung und Professionalisierung der Physiotherapie: „Der studentische Blick auf die Profession“*. Dissertation. Zugriff am 7.3.07 unter: <http://edoc.hu-berlin.de/dissertationen/schaemann-astrid-2005-07-06/HTML/front.html>
- Schämann, A. (2003). Zur Bedeutung einer Forschungstiftung im Kontext des Professionalisierungsprozesses der Physiotherapie. *Zeitschrift für Physiotherapeuten*, (55) 10.
- Scherfer, E. (2003). Standardisierte Tests und Assessments: Bindeglied zwischen Forschung, Praxis, Qualitätssicherung und einer ganzheitlicheren Perspektive, Folge 6. *Zeitschrift für Physiotherapeuten*, 55 (2), 1178-1184.
- Scherfer, E. (2001). Evidenzbasierte Praxis in der Physiotherapie-Bedrohung oder Chance? *Zeitschrift für Physiotherapeuten*, 53 (6), 945-958.
- Schoppen, T., Boonstra, A., Groothoff, J. W., de Vries, J., Goeken, L. N. & Eisma, W. H. (1999). The Timed "up and go" test: reliability and validity in persons with unilateral lower limb amputation. *Arch. Phys. Med. Rehabil.*, 80 (7), 825-828.
- Schliehe, F. & Jäckel, W. H. (2004). Aktuelle Entwicklungen in der Rehabilitation. *Die Rehabilitation*, 43, 257-259.
- Schliehe, F. & Haaf, H.G. (1996). Zur Effektivität und Effizienz der medizinischen Rehabilitation. *Deutsche Rentenversicherung*, 10-11, 666-689.
- Schulte, D. (1993). Wie soll Therapieerfolg gemessen werden? *Zeitschrift für Klinische Psychologie*, 4, 374-393.
- Schuntermann, M. F. (2006). Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF) der Weltgesundheitsorganisation (WHO). Kurzeinführung. pdf-Version. Zugriff am 12.12.06 unter <http://www.deutsche-rentenversicherung.de>
- Schuntermann, M. F. (2002). Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF). Kurzversion zu Ausbildungszwecken mit Definitionen, Sachindex und Anhängen 2 und 3 der ICF: pdf-Version.
- Schuntermann, M. F. (1995). Hinweise auf Bewertungsinstrumente zur Qualitätssicherung in der Rehabilitation – Blatt 2: Der IRES-Fragebogen für Klinik und Forschung. Vorwort. *Rehabilitation*, 34, XIII. Vorwort. *Rehabilitation*, 36, XXXI.
- Simmonds, M. J., Olson, S. L., Jones, S. L., Hussein, S., Lee, T., Novy, C. E. & Radwan, D. (1998). Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. *Spine*, (23) 22, 2412-2421.

- Simmonds, M. J. (1997). Muscle strength. In J. von Deusen & D. Brunt (eds.), *Assessment in occupational therapy and physical therapy*. Philadelphia: W.B. Saunders.
- Södermann, P. & Malchau, M. (2000). Validity and reliability of Swedish WOMAC osteoarthritis index. *Acta Orthop. Scand.*, 71 (1), 39-46.
- Steffanowski, A., Lichtenberg, S., Schmidt, J., Huber, C., Wittmann, W. W. & Nübling, R. (2004). Ergebnisqualität psychosomatischer Rehabilitation: Zielerreichungsskalierung auf der Basis einer strukturierten Therapiezielliste. *Die Rehabilitation*, 43 (4), 219-232.
- Strand, L. I., Ljunggren, A. E., Haldorsen, E. M. & Espehaug, B. (2001). The impact of physical function and pain on work status at 1-year follow-up in patients with back pain. *Spine*, Vol 26., 800-808.
- Stucki, G., Stucki, S. & Sangha, O. (1997). Patienten-zentrierte Evaluation der Krankheitsauswirkungen bei muskuloskelettalen Erkrankungen: Modell der Krankheitsauswirkungen und Begriffsdefinition. *Zeitschrift für Rheumatologie*, 56, 245-254.
- Stucki, G., Meier, D., Stucki, S., Michel, B. A., Tyndall, A. G., Dick, W. & Theiler, R. (1996). Evaluation einer deutschen Version des WOMAC (Western Ontario und McMaster Universities) Arthroseindex. *Zeitschrift für Rheumatologie*, 55, 40-49.
- Tegner, Y., Lysholm, J., Lysholm, M. & Gillquist, J. (1986). A performance test to monitor rehabilitation and evaluate anterior cruciate ligament injuries. *Am. J. Sports Med.*, 14 (2), 156-159.
- Tiffkin, J. & Ascher, E. J. (1948). The Purdue Pegboard: Norms and studies of reliability and validity. *J. Appl. Psychol.*, 32, 234-247.
- Tinetti, M. E. (1986). Performance-oriented assessment of mobility problems in elderly patients. *J. Am. Geriatr. Soc.*, 34 (2).119-26.
- Vandenboorn, H. J. M. (2000). Wissenschaftliche Arbeiten bewerten: Studien lesen und interpretieren. *Krankengymnastik*, 52 (8), 7-12.
- Verdonck, A. (1998). Standardtestverfahren. In I. Froböse, G. Nellessen, (Hrsg.), *Training in der Therapie* (S. 129-146). Wiesbaden: Ullstein Medical.
- Verdonck, A. & Wilke, C. (1998). Screeningverfahren. In I. Froböse & G. Nellessen, (Hrsg.), *Training in der Therapie* (S. 115-128). Wiesbaden: Ullstein Medical.
- Wade, D. T. (1992). *Measurement in Neurological Rehabilitation*. Oxford University Press. 162f.
- Walker, S. R. & Rosser, R. M. (1988). *Quality of Life: Assessment and Application*. Lancaster: MTP Press.
- Ware, J. E. & Sherbourne, C. D. (1992). The MOS 36-item Short-Form Health Survey (SF-36): Conceptual framework and item selection. *Medical Care*, 30 (6). 473-483.
- Wegner, M. (2001). *Sport und Behinderung. Zur Psychologie der Belastungsverarbeitung im Spiegel von Einzelfallanalysen. Beiträge zur Lehre und Forschung im Sport*. Band 129. Schorn-dorf: Hoffmann.
- Weinhold, W. (2008). EDV-gestützte modularisierte Assessment-geleitete Therapiedokumentation. *Zeitschrift für Physiotherapeuten*, 60 (2), 160-171.
- Weinhold, W. (2002). Ergebnisse einer Befragung von 65 Patienten nach vorderer Kreuzbandplastik hinsichtlich ihrer Belastungsfähigkeit in Alltag und Sport 28-43 Monate nach Operation. *Krankengymnastik*, 54 (4), 734-738.
- Weis, J. & Koch, U. (1998). Forschung in der Rehabilitationsmedizin. In H. Delbrück & E. Haupt (Hrsg.), *Rehabilitationsmedizin*. 150-164.
- Weise, G. (1975). *Psychologische Leistungstests*. Göttingen: Hogrefe.
- Werner, S. (2007). Die Auswahl wächst. Studiengänge für Physiotherapie in Deutschland. *Zeitschrift für Physiotherapeuten*, 59 (3), 272-273.

- Wilk, K. E., Romaniello, W. T., Soscia, S. M., Arrigo, C. A. & Andrews, J. R. (1994). The relationship between subjective knee score, isokinetic testing and functional testing in the ACL-deficient knee. *JOSPT*, 20, (2), 60-73.
- Willimczik, K. (2003). *Sportwissenschaften interdisziplinär. Ein wissenschaftlicher Dialog*. Band 1. Hamburg: Czwalina.
- Willimczik, K. (1999). *Statistik im Sport. Grundlagen – Verfahren – Anwendungen*. Band 1. Hamburg: Czwalina.
- Wind, H., Gouttebauge, V., Kuijer P. P., Sluiter J. K. & Frings-Dresen, M. H. (2006). The utility of Functional Capacity Evaluation: the opinion of physicians and other experts in the field of return to work and disability claims. *Int. Arch. Occup. Environ. Health*, 79 (6), 528-534.
- Winograd, C. H., Lemsky, C. M, Nevitt, M. C., Nordstorm, T. M., Stewart, A. L., Miller, C. J. & Bloch, D. A. (1994). Development of a physical performance and mobility examination. *J. Am. Geriatr. Soc.*, 42 (7), 743-749.
- Winkler, J. (1998). Die Messung des sozialen Status mit Hilfe eines Index in den Gesundheitssurveys der DHP. In W. Ahrens, B. M. Bellach & K. H. Jöckel, *Messung soziodemographischer Merkmale in der Epidemiologie* (RKI-Schriften 1/98, S. 69-74). München: MMV Medizinverlag.
- Wirtz, M. (2004). Bestimmung der Güte von Beurteilereinschätzungen mittels der Intraklassenkorrelation und Verbesserung von Beurteilereinschätzungen. *Die Rehabilitation*, 43 (6). 384-389.
- Witte, F. (2005). Gewusst woher. Forschungsgelder für Physiotherapeuten. *Physiopraxis*, (3) 4, 58-59.
- Wollmerstedt, N., Faller, H., Ackermann, H., Schneider, J., Glatzel, M., Kirschner, S. & König, A. (2006). Evaluierung des XSMFA-D an Patienten mit Erkrankungen des Bewegungsapparates und operativer oder konservativer stationärer Therapie. *Rehabilitation*, 45, 78-87.
- World Health Organization (WHO) (2006). Internationale Klassifikation der Funktionsfähigkeit, Behinderung, und Gesundheit (ICF). Kurzeinführung. Zugriff am 5.2.07 unter: <http://www.deutscherentenversicherung.de>.
- World Health Organization (WHO) (2002). Einführung in die internationale Klassifikation der Funktionsfähigkeit, Behinderung, und Gesundheit (ICF) der Weltgesundheitsorganisation (WHO) unter besonderer Berücksichtigung der sozialmedizinischen Begutachtung und Rehabilitation. Erhalten anlässlich einer Arbeitstagung ICF in Bensheim am 30. April 2003. Michael F. Schuntermann, VDR, Rehabilitationswissenschaftliche Abteilung. Koordinator für die deutsche Fassung der ICF im deutschsprachigen Raum.
- World Health Organization (WHO) (1948). Präambel zur Satzung. Genf: WHO
- Wydra, G. (2004). Assessmentverfahren – eine Übersicht. In G. Wydra, S. Winchenbach, M. Schwarz & K. Pfeiffer (Hrsg.), *Assessmentverfahren in Gesundheitssport und Bewegungstherapie. Messen, Testen, Beurteilen, Bewerten*. Jahrestagung der dvs-Kommission Gesundheit vom 23.- 24. September 2004 in Saarbrücken (S. 10). Hamburg: Czwalina.
- Zentralverband der Physiotherapeuten (ZVK) (2005). Standardisierte Ergebnismessung in der Physiotherapiepraxis. Mitteilungen des ZVK e.V. Zugriff am 28.11.06 unter: <http://www.physio-akademie.de/Presseberichte.580.0.html>
- Zentralverband der Physiotherapeuten (ZVK) (2006a). *Handbuch – Standardisierte Ergebnismessung in der Physiotherapie-Praxis*. Zusammengestellt von der Physio-Akademie gGmbH im Auftrag des Deutschen Verbandes für Physiotherapie-Zentralverband der Physiotherapeuten / Krankengymnasten e.V.
- Zentralverband der Physiotherapeuten (ZVK) (2006b). *Entstehung und Gründung der Bildungswerk Physio-Akademie des ZVK gGmbH*. Zugriff am 7.3.07 unter: [http://www.physio-akademie.de/Unsere\\_Geschichte.398.0.html](http://www.physio-akademie.de/Unsere_Geschichte.398.0.html)

- Zimmermann, A. (2007). „Profis“ sind Querdenker. *Zeitschrift für Physiotherapeuten*, 59 (6). 552-562.
- ZIPT (Zukunftsinitiative Physiotherapie) (2006). „Infobörse“ – Neue Berufsfelder. Zugriff am 3.7.07 unter: <http://www.zipt.de/dokumente/Infob%C3%B6rse.pdf>
- ZIPT (Zukunftsinitiative Physiotherapie) (2005). *ZIPT – Basiskonsens – vierter Entwurf: Was ist Physiotherapie? – Unser Selbstverständnis*. Zugriff am 16.7.07 unter: <http://www.ZIPT.de>
- ZIPT (Zukunftsinitiative Physiotherapie) (2003). – ZIPT – Die Zukunftsinitiative in der Physiotherapie. *Zeitschrift für Physiotherapeuten*, 55 (7).
- Zwingmann, C., Metzger, D. & Jäckel, W. H. (1998). Short Form SF-36 Health Survey (SF-36): Psychometrische Analysen der deutschen Version bei Rehabilitanden mit chronischen Rückenschmerzen. *Die Rehabilitation*, 44, 209-219.

**ANHANG 1      Polla**

Test       Intrarater-Test       Interrater-Test       Gesunder-Test

			<b>Ja</b>	<b>Eg.</b>	<b>Nein</b>
1.	1x	<i>Strümpfe anziehen</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.	3x	<i>Aufstehen vom Hocker</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.	10s	<i>Einbeinstand</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.	3x	<i>Aufheben</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.	1x	<i>Hinlegen / Aufstehen</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.	3 Stufen	<i>mit betroffenem Bein</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7.	2x	<i>Hocken</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8.	1x	<i>Hinknien</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.	10s	<i>Laufen auf der Stelle</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.	10x	<i>Einbeinhüpfen</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11.	1-Min.-	<i>Up &amp; Go-Test</i>	Anzahl:	<input type="checkbox"/>	<input type="checkbox"/>
12.	2-Min.-	<i>Gehetest</i>	Meter:	<input type="checkbox"/>	<input type="checkbox"/>
13.	25 Stufen	<i>Treppenlaufen-Test</i>	Sek:	<input type="checkbox"/>	<input type="checkbox"/>
14.	2-Min.-	<i>Lauftest</i>	Meter:	<input type="checkbox"/>	<input type="checkbox"/>

**Wenn Eg. (eingeschränkt), wie?**

1.	1x	<i>Strümpfe anziehen</i>	S D Z AWB 1 2 Hm Hi
2.	3x	<i>Aufstehen vom Hocker</i>	S D Z AWB 1 2 Hm Hi
3.	10s	<i>Einbeinstand</i>	S D Z AWB 1 2 Hm Hi
4.	3x	<i>Aufheben</i>	S D Z Hm Hi
5.	1x	<i>Hinlegen / Aufstehen</i>	S D Z Hm Hi
6.	3 Stufen	<i>mit betroffenem Bein</i>	S D Z AWB 1 2 Hm Hi
7.	2x	<i>Hocken</i>	S D Z AWB 1 2 Hm Hi
8.	1x	<i>Hinknien</i>	S D Z AWB 1 2 Hm Hi
9.	10s	<i>Laufen auf der Stelle</i>	S D Z AWB 1 2
10.	10x	<i>Einbeinhüpfen</i>	S D Z AWB 1 2
11.	1-Min.-	<i>Up &amp; Go-Test</i>	S D Z AWB 1 2 Hm Hi
12.	2-Min.-	<i>Gehetest</i>	S D Z AWB 1 2 Hm Hi
13.	25 Stufen	<i>Treppenlaufen-Test</i>	S D Z AWB 1 2 Hm Hi
14.	2-Min.-	<i>Lauftest</i>	S D Z

S = Schmerz    D = Druck    Z = Ziehen    AWB: 1 = leicht; 2 = stark    Hm = Hilfsmittel    Hi = Hilfe

Testabbruch? Warum?

## ANHANG 2 Gesundheitsbezogene Lebensqualität (SF-36)

In diesem Fragebogen geht es um die Beurteilung Ihres Gesundheitszustandes. Der Bogen ermöglicht es, im Zeitverlauf nachzuvollziehen, wie Sie sich fühlen und wie Sie im Alltag zurechtkommen. Bitte beantworten Sie jede der Fragen, indem Sie bei den Antwortmöglichkeiten das ankreuzen, was am besten auf Sie zutrifft.

1.	Wie würden Sie Ihren Gesundheitszustand im Allgemeinen beschreiben?				
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	ausgezeichnet	sehr gut	gut	weniger gut	schlecht

2.	Im Vergleich zum vergangenen Jahr, wie würden Sie Ihren derzeitigen Gesundheitszustand beschreiben?				
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	derzeit viel besser	derzeit etwas besser	etwa wie vor einem Jahr	derzeit etwas schlechter	derzeit viel schlechter

Im Folgenden sind einige Tätigkeiten beschrieben, die Sie vielleicht an einem normalen Tag ausüben. Sind Sie durch Ihren derzeitigen Gesundheitszustand bei diesen Tätigkeiten eingeschränkt? Wenn ja, wie stark?				
3.		Ja, sehr stark eingeschränkt	Ja, etwas eingeschränkt	Nein, überhaupt nicht eingeschränkt
3.a	<b>Anstrengende Tätigkeiten</b> , z.B. schnell laufen, schwere Gegenstände heben, anstrengenden Sport treiben	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.b	<b>Mittelschwere Tätigkeiten</b> , z.B. einen Tisch verschieben, Staubsaugen, Kegeln, im Garten arbeiten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.c	Einkaufstaschen heben oder tragen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.d	<b>Mehrere</b> Treppenabsätze steigen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.e	<b>Einen</b> Treppenabsatz steigen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.f	Sich beugen, knien, bücken	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.g	<b>Mehr als 1 Kilometer</b> zu Fuß gehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.h	<b>Mehrere</b> Straßenkreuzungen weit zu Fuß gehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.i	<b>Eine</b> Straßenkreuzung weit zu Fuß gehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.j	Sich baden oder anziehen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Hatten Sie in der *vergangenen Woche* aufgrund Ihrer **körperlichen Gesundheit** irgendwelche Schwierigkeiten bei der Arbeit oder anderen alltäglichen Tätigkeiten im Beruf oder zu Hause?

	Ja	Nein
4.a Ich konnte nicht <b>so lange</b> wie üblich tätig sein.	<input type="checkbox"/>	<input type="checkbox"/>
4.b Ich habe <b>weniger geschafft</b> als ich wollte.	<input type="checkbox"/>	<input type="checkbox"/>
4.c Ich konnte <b>nur bestimmte Dinge</b> tun.	<input type="checkbox"/>	<input type="checkbox"/>
4.d Ich hatte <b>Schwierigkeiten</b> bei der Ausführung.	<input type="checkbox"/>	<input type="checkbox"/>

Hatten Sie in der *vergangenen Woche* aufgrund **seelischer Probleme** irgendwelche Schwierigkeiten bei der Arbeit oder anderen alltäglichen Tätigkeiten im Beruf bzw. zu Hause (z.B. weil Sie sich niedergeschlagen oder ängstlich fühlten)?

	Ja	Nein
5.a Ich konnte nicht <b>so lange</b> wie üblich tätig sein.	<input type="checkbox"/>	<input type="checkbox"/>
5.b Ich habe <b>weniger geschafft</b> als ich wollte.	<input type="checkbox"/>	<input type="checkbox"/>
5.c Ich konnte nicht so <b>sorgfältig</b> wie üblich arbeiten.	<input type="checkbox"/>	<input type="checkbox"/>

6. Wie sehr haben Ihre körperliche Gesundheit oder seelischen Probleme in der *vergangenen Woche* Ihre normalen Kontakte zu Familienangehörigen, Freunden, Nachbarn oder zum Bekanntenkreis beeinträchtigt?

<input type="checkbox"/>				
überhaupt nicht	etwas	mäßig	ziemlich	sehr

7. Hatten Sie in der *vergangenen Woche* Schmerzen?

<input type="checkbox"/>					
keine	sehr leichte	leichte	mäßige	starke	sehr starke

8. Inwieweit haben die Schmerzen Sie in der *vergangenen Woche* bei der Ausübung Ihrer Alltagsaktivitäten zu Hause und im Beruf behindert?

<input type="checkbox"/>				
überhaupt nicht	etwas	mäßig	ziemlich	sehr

In diesen Fragen geht es darum, wie Sie sich fühlen und wie es Ihnen in der vergangenen Woche gegangen ist. (Bitte kreuzen Sie in jeder Zeile die Zahl an, die Ihrem Befinden am ehesten entspricht). Wie oft waren Sie in der vergangenen Woche...

	immer	meist	ziemlich oft	manchmal	selten	nie
9.a	<input type="checkbox"/>					
9.b	<input type="checkbox"/>					
9.c	<input type="checkbox"/>					
9.d	<input type="checkbox"/>					
9.e	<input type="checkbox"/>					
9.f	<input type="checkbox"/>					
9.g	<input type="checkbox"/>					
9.h	<input type="checkbox"/>					
9.i	<input type="checkbox"/>					

10. Wie häufig haben Ihre körperliche Gesundheit oder seelischen Probleme in der vergangenen Woche Ihre Kontakte zu anderen Menschen (Besuche bei Freunden, Verwandten usw.) beeinträchtigt?

<input type="checkbox"/>				
immer	meistens	manchmal	selten	nie

	Inwieweit trifft <i>jede</i> der folgenden Aussagen auf Sie zu?	trifft ganz zu	trifft weitgehend zu	Weiß nicht	trifft weitgehend nicht zu	trifft überhaupt nicht zu
11.a	Ich schein(e) etwas leichter als andere krank zu werden	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11.b	Ich bin genauso gesund wie alle anderen, die ich kenne	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11.c	Ich erwarte, dass meine Gesundheit nachlässt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11.d	Ich erfreue mich ausgezeichnete(r) Gesundheit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

### ANHANG 3 Umkodierte und kalibrierte Items des SF-36

Item	Beschreibung	Antworten	1	2
1	Allgemeiner Gesundheitszustand	ausgezeichnet sehr gut gut weniger gut schlecht	1 2 3 4 5	5,0 4,4 3,4 2,0 1,0
11.b 11.d	Ich bin so gesund wie alle anderen, die ich kenne Ich erfreue mich ausgezeichneter Gesundheit	trifft ganz zu trifft weitgehend zu weiß nicht trifft weitgehend nicht zu trifft überhaupt nicht zu	1 2 3 4 5	5 4 3 2 1
7	Schmerzintensität	keine Schmerzen sehr leicht leicht mäßig stark sehr stark	1 2 3 4 5 6	6,0 5,4 4,2 3,1 2,2 1,0
8	Behinderung durch Schmerzen im Alltag	überhaupt nicht überhaupt nicht ein bisschen mäßig Ziemlich sehr	1 1 2 3 4 5	6 5 4 3 2 1
9a 9e	...voller Schwung? ...voller Energie?	immer meistens ziemlich oft manchmal selten nie	1 2 3 4 5 6	6 5 4 3 2 1
6	Beeinträchtigung von Kontakten in der Intensität	überhaupt nicht etwas mäßig ziemlich sehr	1 2 3 4 5	5 4 3 2 1
10	Beeinträchtigung von Kontakten in der Häufigkeit	immer meistens manchmal selten nie	1 2 3 4 5	5 4 3 2 1
2	Veränderung des Gesundheitszustandes			
	Derzeit viel besser als vor einem Jahr		1	5
	Derzeit etwas besser als vor einem Jahr		2	4
	Etwa wie vor einem Jahr		3	3
	Derzeit etwas schlechter als vor einem Jahr		4	2
	Derzeit viel schlechter als vor einem Jahr		5	1

Anm.: 1: vorkodierte Item Werte; 2: endgültige Item Werte

## ANHANG 4      Selbsteinschätzung von Aktivitäten

Durch diesen Fragebogen bekommen wir Informationen darüber, in welcher Form Beschwerden Ihren Alltag beeinflussen. Wir würden gerne wissen, ob es Ihnen in der vergangenen Woche aufgrund ihrer Beinbeschwerden Probleme bereitet hat, nachstehende Aktivitäten auszuführen. Bitte wählen Sie bei jeder Aktivität eine Antwort aus und umkreisen die entsprechende Zahl.

	Ich habe keine Probleme	Ich habe leichte Probleme	Ich habe mäßige Probleme	Ich habe große Probleme	Ich kann diese Aktivität nicht
Auf einer weichen Matte auf dem Boden knien	4	3	2	1	0
Strümpfe anziehen	4	3	2	1	0
Auf den Boden hinlegen und wieder aufstehen	4	3	2	1	0
Für eine Stunde auf einem Stuhl sitzen	4	3	2	1	0
Aus einem Auto aussteigen oder von einem tiefen Sessel aufstehen	4	3	2	1	0
Aufstehen nach längerem Sitzen	4	3	2	1	0
Einige Sekunden auf dem betroffenen Bein stehen	4	3	2	1	0
Ohne Unterbrechung 20 - 30 Min stehen	4	3	2	1	0
Eine Straßenkreuzung weit zu Fuß gehen	4	3	2	1	0
Mehrere Straßenkreuzungen zu Fuß gehen	4	3	2	1	0
Einen langen Spaziergang machen	4	3	2	1	0
Eine Einkaufstasche heben oder tragen	4	3	2	1	0
Auf den Zehenspitzen gehen	4	3	2	1	0
Auf den Hacken gehen	4	3	2	1	0
Ein Stockwerk hinaufgehen	4	3	2	1	0
Ein Stockwerk hinuntergehen	4	3	2	1	0
In die Hocke gehen	4	3	2	1	0
Auf dem betroffenen Bein hüpfen (5-10x)	4	3	2	1	0
Für einige Sekunden auf der Stelle laufen	4	3	2	1	0
Ein kleines Stück rennen	4	3	2	1	0
Eine Zeit lang joggen (zwei Minuten)	4	3	2	1	0
Schwimmen gehen	4	3	2	1	0
Einen Schuh vom Boden aufheben	4	3	2	1	0
Radfahren im Freien	4	3	2	1	0
Eine Weile über Kopfsteinpflaster gehen	4	3	2	1	0
Die Toilette benutzen	4	3	2	1	0
Eine Kiste Wasser vom Boden auf den Tisch stellen	4	3	2	1	0



## ANHANG 6      Patienteninformation und Einwilligungserklärung



**Sehr geehrter Patient, sehr geehrte Patientin**

das **Ambulante Rehasentrum Kiel GmbH** führt mit Unterstützung des **Vereins zur Förderung der Rehabilitationsforschung in Schleswig-Holstein e.V.** eine wissenschaftliche Studie zur Qualitätssicherung Ihrer Therapie durch.

Zu diesem Zweck sind zu Therapiebeginn neben der regulären Befunderhebung Fragebögen und praktische Überprüfungen (Tests) geplant. Ziel der Tests ist es herauszufinden, inwieweit Sie Schwierigkeiten bei der Ausführung von Aktivitäten des Alltages haben, wie sich diese während der Therapie verändern und welche Zusammenhänge zur gesundheitsbezogenen Lebensqualität (Fragebogen) bestehen. Um zu erfahren, wie der Stand der Therapie nach sechs Wochen ist, werden wir uns dann noch einmal mit den gleichen Fragen und Überprüfungen an Sie wenden.

Die Tests werden selbstverständlich nur dann durchgeführt, wenn Sie sich dazu im Stande fühlen, **die Bewegungen schmerzfrei möglich** und **keinerlei Schäden zu erwarten sind**. Die Tests beinhalten folgende Aktivitäten:

1. Strümpfe anziehen
2. Aufstehen von einem Hocker
3. Einbeinstand
4. Aufheben eines leichten Gegenstandes
5. Hinlegen und wieder aufstehen
6. Treppensteigen
7. Hocken
8. Hinknien
9. Laufen auf der Stelle
10. Einbeinhüpfen auf der Stelle
11. 1-Minute-Aufsteh-und-Geh-Test
12. 2-Minuten-Gehtest
13. 25 Stufen Treppenlaufen-Test
14. 2-Minuten-Lauftest auf dem Laufband.

**Wir bitten Sie herzlich um Ihre Mitarbeit in dieser Studie!**

Neben den Angaben aus dem Fragebogen werden keine weiteren Daten erfasst. Die Befragungen erfolgen anonym.

Selbstverständlich ist Ihre Teilnahme an der Befragung freiwillig. Wenn Sie nicht teilnehmen wollen, entstehen Ihnen dadurch keine Nachteile. Ihre Angaben werden vertraulich behandelt. Alle Angaben werden ausschließlich von den Mitarbeitern des **Ambulanten Rehasentrums Kiel GmbH** und dem **Verein zur Förderung der Rehabilitationsforschung in Schleswig-Holstein e.V.** mit Hilfe der EDV ausgewertet. Ihr Name und Ihre Einwilligungserklärung werden von den Angaben der Fragebögen getrennt und am Ende der Erhebungsphase vernichtet. Eine Weitergabe an staatliche oder private Einrichtungen ist ausgeschlossen.

Alle Unterlagen werden nach Abschluss der Studie vernichtet.

Die Befragung ist an Ihre ausdrückliche Einwilligung gebunden. Wenn Sie mit der Teilnahme an der Studie einverstanden sind, dann unterschreiben Sie bitte die nachfolgende Einverständniserklärung.

### EINVERSTÄNDNISERKLÄRUNG

**Ich bin mit der Teilnahme an der Untersuchung des *Ambulanten Rehasentrums Kiel GmbH* einverstanden.**

Die Untersuchungsergebnisse dürfen ausschließlich vom ***Ambulanten Rehasentrum Kiel GmbH*** wissenschaftlich und mit Hilfe der EDV ausgewertet werden.

DATUM

UNTERSCHRIFT

Wenn Sie Fragen haben, wenden Sie sich bitte an die Rezeption, den zuständigen Therapeuten, Dr. Beske, Herrn Naeve oder Herrn Weinhold.

Sie können sich aber auch an die wissenschaftliche Begleitung des Projektes mit folgender Adresse schriftlich oder mündlich wenden: Professor Dr. Manfred Wegner, Institut für Sport und Sportwissenschaft, Damaschkestraße 25 in 34121 Kassel, Telefon: 0561-804-5256, Email: [m.wegner@uni-kassel.de](mailto:m.wegner@uni-kassel.de).

Vielen Dank für Ihre Mitarbeit.

***Ihr Reha Team***

## ANHANG 7 Polla modifiziert

<b>Qualitativ</b>		<b>Test</b>	<b>Ja</b>	<b>Eg*</b>	<b>Nein</b>
1.	1x	<b>Strümpfe an-/ausziehen</b>	2	1	0
2.	3x	<b>Aufstehen vom Hocker</b>	2	1	0
3.	10s	<b>Einbeinstand</b>	2	1	0
4.	20m	<b>Gehen</b>	2	1	0
5.	1x	<b>Hinlegen / Aufstehen</b>	2	1	0
6.	3x	<b>Auf einen Stepp steigen</b>	2	1	0
7.	2x	<b>Hocken</b>	2	1	0
8.	1x	<b>Hinknien</b>	2	1	0
9.	10m	<b>Laufen durch den Raum</b>	2	1	0
10.	10x	<b>Einbeinhüpfen</b>	2	1	0
*Eingeschränkt (max. 22 Pkte.)					

S = Schmerz    D = Druck    Z = Ziehen    AWB: 1 = leicht; 2 = stark    Hm = Hilfsmittel    Hi = Hilfe

### Ein gemessener Test (Wahlpflicht)

<b>Quantitativ</b>	<b>Test</b>	<b>Ergebnis</b>
1.	<b>1-Minute-Up &amp; Go-Test</b>	<b>Anzahl:</b>
2.	<b>2-Minuten-Gehtest</b>	<b>Meter:</b>
3.	<b>Treppenlaufen-Test</b>	<b>Sek.:</b>
4.	<b>2-Minuten-Lauftest</b>	<b>Meter:</b>

### Ein hoch anspruchsvoller Test (Optional)

	<b>Test</b>	<b>Ergebnis</b>
1.	<b>Triple-Jump-Test</b>	<b>Links:</b>
		<b>Rechts:</b>
		<b>Differenz in cm:</b>

Testabbruch? Warum?

## ANHANG 8 Testanweisung Polla modifiziert

### Vorbemerkungen

Zunächst werden die 11 qualitativ beurteilten Tests zum Polla („ja-eingeschränkt-nein“) durchgeführt und die erreichte Punktzahl ermittelt. Darüber kommt ein quantitativer (gemessener) Test zum Einsatz, der vom Therapeuten ausgewählt wird. Zusätzlich kann der Triple-Jump-Test als Test mit erhöhtem Anforderungscharakter an die einbeinige Belastung (bei Funktionserkrankungen, Jüngeren oder Sportlern) angewandt werden. Alle Tests werden mit Schuhen absolviert. Je Test ist ein Versuch zugelassen und im Anschluss an jeden Test wird nach Schmerz, Druck oder Ziehen gefragt.

### Beurteilungskriterien der Tests im Polla

Die Durchführung jedes einzelnen qualitativen Tests erfolgt in der Beurteilung nach den Kriterien:

1. „nein, die Aktivität ist nicht möglich“ = 0
2. „die Aktivität ist eingeschränkt möglich“ = 1
3. „die Aktivität ist uneingeschränkt möglich“ = 2

Einschränkungen bei der Ausführung der Aktivitäten unterliegen der Beurteilung durch:

1. das Auftreten von Schmerz, Druckgefühl oder Ziehen,
2. das Auftreten von leichten oder starken Ausweichbewegungen,
3. die Inanspruchnahme von Hilfe (durch eine Person),
4. die Inanspruchnahme eines Hilfsmittels.

### Ausführungsbeschreibung – Polla qualitativ

#### 1. Strümpfe an-/ausziehen

Ausgangsposition: Sitz auf einem 45 cm hohen Hocker ohne Lehne, die Füße stehen parallel.

Therapeut fragt: „Können Sie ihre Strümpfe im Sitzen einmal aus- und wieder anziehen?“

#### 2. Aufstehen vom Hocker

Ausgangsposition: Sitz auf einem 45 cm hohen Hocker ohne Lehne, die Füße stehen parallel.

Therapeut fragt: „Können Sie dreimal von diesem Hocker ohne Zuhilfenahme der Arme aufstehen und sich wieder hinsetzen, Füße bitte parallel aufsetzen?“

#### 3. Einbeinstand

Therapeut fragt: „Können Sie für zehn Sekunden auf einem leicht gebeugten Bein so stehen, dass sich die Beine nicht berühren?“

Beurteilung uneingeschränkt: wenn 10 Sekunden freies Stehen möglich.

Beurteilung eingeschränkt: wenn mit Schmerz, Druck oder Ziehen sowie mit Hilfe/Hilfsmittel.

#### 4. Gehen

Therapeut fragt: „Können Sie bitte in diesem Raum auf- und abgehen (ca. 20 m)?“

---

## 5. Hinlegen/Aufstehen

Ausgangsposition: Parallelstand vor einer auf dem Fußboden liegenden Therapiematte.

Therapeut fragt: „Können Sie sich aus dem Stehen hier auf dieser Matte einmal flach auf den Rücken legen, die Beine ausstrecken und wieder aufstehen?“

---

## 6. Auf einen Stepp steigen

Ausgangsposition: Parallelstand vor einem 20 Zentimeter hohen Stepp.

Therapeut fragt: „Können Sie dreimal mit einem Bein auf diese Stufe steigen, das Knie dabei strecken und nach vorne absteigen?“

Hinweis: Erst dreimal das gesunde, anschließend dreimal das betroffene Bein.

---

## 7. Hocken

Ausgangsposition: Parallelstand.

Therapeut fragt: „Können Sie zweimal nacheinander in die Hocke gehen und wieder aufstehen?“

Beurteilung uneingeschränkt: - Waden-Oberschenkel-Kontakt und Seitengleichheit.

Beurteilung eingeschränkt: - Wenn kein Waden-Oberschenkelkontakt,  
- mit Ausweichbewegung, Druck, Schmerz, Ziehen  
- oder wenn Hilfe/Hilfsmittel zum Einsatz kommt.

---

## 8. Hinknien

Ausgangsposition: Parallelstand vor einer auf dem Fußboden liegenden Therapiematte.

Therapeut fragt: „Trauen Sie sich, einmal auf dieser Matte zu knien und 20x mit den Knien auf der Stelle zu tippeln?“

---

## 9. Laufen durch den Raum

Therapeut fragt: „Ist es Ihnen möglich, 10 Sekunden durch den Raum laufen?“ (Füße lösen)

---

## 10. Einbeinhüpfen

Therapeut fragt: „Trauen Sie sich, 10 Sekunden auf einem Bein auf der Stelle zu hüpfen?“ (Fuß löst ganz vom Boden).

### Ausführungsbeschreibung zu gemessenen Tests – quantitativ

#### 1-Minute-Up & Go-Test

Ausgangsposition: Sitz auf einem 45 cm hohen Hocker ohne Lehne, die Füße stehen parallel mit den Fußballen auf einer Dreimeter Markierung von der Wand entfernt.

Therapeut fragt: „Wie oft schaffen Sie es, innerhalb einer Minute so oft wie möglich vom Hocker aufzustehen, drei Meter bis zur Wand zu gehen, die Wand mit der Hand zu berühren und sich wieder auf den Hocker zu setzen?“ Achten Sie darauf, dass Sie nicht den Hocker verschieben“.

Sicherheitshinweis: Zur Sicherheit steht der Versuchsleiter hinter dem Patienten, die Füße des Patienten stehen auf der drei Meter Markierung. Gezählt wird jedes Hinsetzen oder halbe Strecken bei Abbruch an der Wand.

---

### Zwei-Minuten-Gehtest

Ausgangsposition: Der Patient steht an einer Markierung, eine zweite befindet sich im Abstand von 14 Meter zur ersten.

Therapeut fordert auf: *„Bitte gehen Sie zwischen den beiden Markierungen so schnell es Ihnen möglich ist hin und her. Wir halten die Anzahl der gegangenen Meter fest. Teilen Sie ihre Kräfte während der zwei Minuten so ein, dass Sie in gleichmäßiger Geschwindigkeit ohne Beschwerden (oder ohne zunehmende) die Übung durchführen können und sich selbst dabei nicht schädigen.“*

---

### Treppenlaufen-Test

Ausgangsposition: Patient steht an einem Treppenabsatz im Treppenhaus.

Therapeut fordert auf: *„Bitte gehen oder laufen Sie – unter Benutzung jeder Stufe – so schnell es ihnen möglich ist, ein Stockwerk tiefer und wieder hoch. Wenn möglich, benutzen Sie dabei bitte kein Geländer. Die dafür benötigte Zeit wird registriert.“*

Hinweis: Der Patient kann das Geländer anfassen oder auch Stützen einsetzen. Entsprechend wird dies als „eingeschränkt“ (Punktabzug) vermerkt.

---

### Zwei-Minuten-Lauftest

Ausgangsposition: Sofern sich der Patient damit einverstanden erklärt hat, steht er auf einem Laufband.

Therapeut fordert auf: *„Trauen Sie sich zu – so schnell wie es ihnen möglich ist – zwei Minuten auf dem Laufband zu laufen? Die Anzahl der Meter wird festgehalten.“*

Hinweis: Vor Testbeginn wird zur Sicherheit das Laufen auf dem Band erprobt. Ist sicheres Laufen gewährleistet, wird das Laufband vom Therapeuten neu gestartet. Er stellt dabei zunächst eine mäßige Laufgeschwindigkeit ein, um diese in Absprache mit dem Patienten zu steigern. Der Patient läuft so schnell wie möglich mit der Auflage, bei beginnendem Schmerz den Test abzubrechen.

Hinweis: Bei Gelenkersatz wird kein Lauftest durchgeführt.

**Test mit erhöhtem Anforderungscharakter an einbeinige Belastung**

### Triple-Jump-Test

Ausgangsposition: Der Patient steht im Einbeinstand.

Therapeut fordert auf: *„Trauen Sie es sich zu, auf dem gesunden Bein in drei Sprüngen entlang der Linie auf dem Boden vorwärts zu springen?“*

Im Anschluss wird der Dreibeinsprung mit dem betroffenen Bein durchgeführt. Die Differenz beider Sprünge wird in Zentimetern erfasst.