

Aus der Klinik für Anästhesiologie mit Schwerpunkt operative
Intensivmedizin
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Verlässlichkeit von Angehörigen bei der Beurteilung der
prämorbidem Gesundheit kritisch kranker Patienten

„Reliability of proxies in the evaluation of the premorbid health
related quality of life of critically ill Patients“

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Lukas Schönagel

Datum der Promotion: 30.11.2023

Inhaltsverzeichnis

Tabellenverzeichnis	2
Abbildungsverzeichnis	3
Abkürzungsverzeichnis	5
1 Abstrakt	6
2 Einleitung	9
2.1 Hinführung zum Thema.....	9
2.2 ICF und WHODAS 2.0.....	11
2.3 Stand der Literatur	11
2.4 Fragestellung.....	15
3 Methodik	16
3.1 Studiendesign.....	16
3.2 Ethikvotum	16
3.3 Rekrutierung	16
3.4 Einschluss- und Ausschlusskriterien	17
3.5 Erhebungsinstrumente	17
3.6 Statistische Analyse.....	18
4 Ergebnisse	25
4.1. Populationscharakteristika.....	25
4.2 Ergebnisse zur Fragestellung 1	26
4.3 Ergebnisse zur Fragestellung 2.....	38
4.4 Ergebnisse zur Fragestellung 3.....	50
5 Diskussion	56
5.1 Diskussion der Ergebnisse.....	56
5.2 Stärken und Limitationen	64
6 Anhang	66
7 Literaturverzeichnis.....	69
Eidesstattliche Versicherung	73
Lebenslauf	74
Danksagung	76
Bescheinigung des akkreditierten Statistikers.....	77

Tabellenverzeichnis

Tabelle 1: Stand der Literatur	13
Tabelle 2: In anderen Studien untersuchte Einflussfaktoren auf die Übereinstimmung zwischen Patient*innen und deren Stellvertreter*innen	14
Tabelle 3: In Fragestellung 3 untersuchte Charakteristika und Gruppeneinteilung	24
Tabelle 4: Mediane Punktwerte und Differenzen in den Fragen, Gesundheitsdomänen und der Gesamtpunktzahl des WHODAS 2.0	28
Tabelle 5: Populationscharakteristika der postoperativ auf Normal- und Intensivstation behandelten Patient*innen und deren Stellvertreter*innen	39
Tabelle 6: Mediane Punktwerte der Patient*innen und derer Stellvertreter*innen in den Gesundheitsdomänen und der Gesamtpunktzahl des WHODAS 2.0	40
Tabelle 7: Vergleich der medianen Differenz der Punktwerten der Patient*innen und deren Stellvertreter*innen in den Gesundheitsdomänen und der Gesamtpunktzahl des WHODAS 2.0	41
Tabelle 8: Gruppencharakteristika der beiden Gruppen im Fallkontrollabgleich	46
Tabelle 9: Mediane Punktzahl der Patient*innen in der Gesamtpunktzahl des WHODAS 2.0 bei den verschiedenen Gruppen beim Fallkontrollabgleich zur Fragestellung 2	47
Tabelle 10: Mediane Differenzen zwischen den Patient*innen und deren Stellvertreter*innen in den Gesundheitsdomänen und der Gesamtpunktzahl des WHODAS 2.0 beim Fallkontrollabgleich zur Fragestellung 2	47
Tabelle 11: Anzahl in den jeweiligen Gruppen der Fragestellung 3	50
Tabelle 12: Mediane Punktwerte der Patient*innen in der Gesamtpunktzahl des WHODAS 2.0 in den jeweiligen Gruppen der Fragestellung 3	51
Tabelle 13: Differenz zwischen den Punktwerten der Patient*innen und deren Stellvertreter*innen in der Gesamtpunktzahl des WHODAS 2.0 bei den jeweiligen Gruppen der Fragestellung 3	52
Tabelle 14: Gruppencharakteristika der Patient*innen und derer Stellvertreter*innen im Fallkontrollabgleich der Fragstellung 3	55

Abbildungsverzeichnis

Abbildung 1: Linearen Gewichtung für die Fragen 1 bis 12 des WHODAS 2.0	22
Abbildung 2: Verteilung der Antworten in den Fragen 1 bis 12 des WHODAS 2.0.....	26
Abbildung 3: Verteilung der Punktwerte in den Gesundheitsdomänen des WHODAS 2.0.....	27
Abbildung 4: Anteil der gleichen, um 1 Kategorie und um mind. 2 Kategorien abweichender Antworten in den Fragen 1 bis 12 des WHODAS 2.0	29
Abbildung 5 Anteil gleichen, um 1 bis 2 und um mind. 3 Punkte abweichender Punktwerte der Patient*innen und deren Stellvertreter*innen in den Gesundheitsdomänen des WHODAS 2.0.....	30
Abbildung 6: Anteil um max. 0, 10 und 25 Punkte abweichender Punktwerte in der Gesamtpunktzahl des WHODAS 2.0	31
Abbildung 7: Bland-Altman-Diagramm für die Übereinstimmung zwischen den Patient*innen und derer Stellvertreter*innen in der Gesamtpunktzahl des WHODAS 2	32
Abbildung 8: Gewichtete Kappa-Werte in den Fragen 1 bis 12 Fragen des WHODAS 2.0	33
Abbildung 9: Gewichteten Kappa-Werte in den Gesundheitsdomänen des WHODAS 2.0	34
Abbildung 10: PABAK-Werte in den Fragen 1 bis 12 des WHODAS 2.0	35
Abbildung 11: PABAK-Werte in den Gesundheitsdomänen des WHODAS 2.0	36
Abbildung 12: Zusammenhang zwischen dem Prozentsatz an Patient*innen die angegeben haben „keine Einschränkungen“ zu haben und den gewichteten Kappa-Werten	37
Abbildung 12: Zusammenhang zwischen dem Prozentsatz an Patient*innen die angegeben haben „keine Einschränkungen“ zu haben und den PABAK-Werten	38
Abbildung 14: Bland-Altman-Diagramm für die Übereinstimmung in der Gesamtpunktzahl des WHODAS 2.0, zwischen den Patient*innen, die postoperativ nicht intensivmedizinisch behandelt wurden und deren Stellvertretern	42
Abbildung 15: Bland-Altman-Diagramm für die Übereinstimmung in der Gesamtpunktzahl des WHODAS 2.0, zwischen den Patient*innen, die postoperativ intensivmedizinisch behandelt wurden und deren Stellvertretern	43
Abbildung 16: Vergleich der gewichteten Kappa-Werte in den Gesundheitsdomänen des WHODAS 2.0 zwischen den verschiedenen Gruppen	44

Abbildung 17: Vergleich der PABAK-Werte in den Gesundheitsdomänen des WHODAS 2.0 zwischen den verschiedenen Gruppen	45
Abbildung 18: Vergleich der gewichteten Kappa-Werte in den Gesundheitsdomänen des WHODAS 2.0 zwischen den verschiedenen Gruppen im Fallkontrollabgleich	49
Abbildung 19: Vergleich der PABAK-Werte in den Gesundheitsdomänen des WHODAS 2.0 zwischen den verschiedenen Gruppen im Fallkontrollabgleich	49
Abbildung 20: ICC der verschiedenen Gruppen der Fragstellung 3 in der Gesamtpunktzahl des WHODAS 2.0	53
Abbildung 21: Ergebnisse der Bland-Altman-Analyse bei den verschiedenen Gruppen der Fragstellung 3 in der Gesamtpunktzahl des WHODAS 2.0	54

Abkürzungsverzeichnis

ALDS	Academic Medical Center Linear Disability Score
ARDS	Acute Respiratory Distress Syndrome
EQ5D	European Quality of Life 5 Dimensions
HrQoL	Health-related quality of life
ICC	Intraklassenkorrelation
ICF	International Classification of Functioning, Disability and Health
IQA	Interquartilsabstand
KI	Konfidenzintervall
Kw	Gewichtetes Kappa
LOA	Limits of Agreement
LSA	Life Space Assessment
MCID	Minimal Clinically Important Differenz
MOCA	Montreal Cognitive Assessment
OARS	Older American Resources Services
PABAK	Prevalence adjusted and Bias adjusted Kappa
PCS	Physical Component Score des SF-36
QoL	Quality of Life
R ²	Bestimmtheitsmaß
SD	Standarddeviation
SF-12	Short Form 12
SF36	Short Form 36
SPSS	Statistik Package for Sozial Sciences
VES-13	Vulnerable Elder Survey
WHO	World Health Organization
WHODAS	World Health Organization Disability Assessment. Schedule

1 Abstrakt

Hintergrund: Um im Rahmen der intensivmedizinischen Behandlung von Patient*innen eine Therapieplanung und Erfassung von Langzeitfolgen zu ermöglichen, ist eine akkurate Einschätzung der prämorbidem gesundheitsbezogenen Lebensqualität (HrQoL) nötig. Hierzu werden häufig Angehörige befragt, da die Patient*innen aufgrund der schwerer der Erkrankung meistens nicht selbst antworten können.

Ziel: In dieser Dissertation soll die Übereinstimmung zwischen Patient*innen und deren Stellvertreter*innen bei der Beurteilung der prämorbidem HrQoL in der 12-Fragen Version des WHODAS 2.0 ermittelt werden. Weiterhin soll der Einfluss einer intensivmedizinischen Behandlung der Patient*innen auf die Übereinstimmung, sowie weiterer Einflussfaktoren ermittelt werden.

Methodik: Wir führten eine prospektive Observationsstudie durch, in der Patienten selbst sowie ihre Angehörigen stellvertretend mit Hilfe des WHODAS 2.0 Fragebogen zur gesundheitsbezogenen Lebensqualität befragt wurden. Die Patienten wurden vor elektiven Operationen befragt, ihre Angehörigen postoperativ auf der Normal- bzw. Intensivstation. Übereinstimmungen in den einzelnen Items des WHODAS 2.0 wurde mittels des gewichteten Kappa- sowie eines Prävalenz und Bias adjustiertem Kappa-Wertes (PABAK) ermittelt. Die Übereinstimmung in der Gesamtpunktzahl wurde mittels der „Limits of Agreement“ (LOA) nach Bland-Altman und der Intraklassenkorrelation (ICC) ermittelt. Dabei wurden untere Konfidenzintervalle der Kappa-Werte über 0,4, untere Konfidenzintervalle der ICC über 0,5 sowie LOA unter 10 Punkten als ausreichend gewertet.

Ergebnisse: Insgesamt wurden 204 Patient*innen rekrutiert, davon wurden 102 Patient*innen postoperativ auf einer Intensivstation behandelt. Bei den meisten Frage des WHODAS 2.0 liegt anhand der PABAK-Werte eine gute Übereinstimmung vor. Die ICC zeigt ebenfalls eine ausreichende Übereinstimmung in der Gesamtpunktzahl (ICC: 0,724, 95%KI: 0,627 bis 0,794). Die LOA liegen deutlich außerhalb von 10 Punkten (LOA -40,5 bis 30,9). Bei Patient*innen die intensivmedizinisch behandelt wurden (n=102), liegt ein Bias zur Unterschätzung der prämorbidem Lebensqualität durch die Angehörigen vor (p=0,009). Bei Stellvertreter*innen, welchen die Patient*innen pflegten (n=52), liegt eine niedrigere Übereinstimmung bei der Gesamtpunktzahl des WHODAS (ICC: 0,498 vs. 0,754) und ein negativer Bias (p=0,009) vor.

Diskussion: In dieser Dissertation konnte erste Evidenz für die Validität des WHODAS 2.0 bei der Befragung von Stellvertretern zur prämorbidem HrQoL von intensivmedizinisch behandelten Patient*innen erbracht werden. Obwohl der ICC eine moderate bis gute Übereinstimmung zeigte, bedeutet die Weite der LOA das ein relevanter Anteil der Patient*innen durch die Stellvertreter*innen falsch klassifiziert wurde. Es ist weitere Forschung nötig, um besser bewerten zu können, ob durch die Befragung von Stellvertreter*innen es ermöglicht wird ein patientenadaptiertes Therapieziel zu erfassen.

Abstrakt in englischer Sprache

Background: Accurate information about the premorbid health related quality of life (HrQoL) is needed to plan the therapy of patients undergoing intensive care treatment and to evaluate the success of interventions. The patients are often not able to provide these information's themselves, so proxies are used.

Aim: The goal of this dissertation is to evaluate the agreement between patients and their respective proxies pertaining to the premorbid HrQoL of the patients undergoing intensive care treatment via the 12-question version of the WHODAS 2.0. The second goal is to evaluate the influence the treatment of patient in intensive care units and chosen characteristics of the patients and proxies have on the afore mentioned agreement.

Method: 204 Patient-Proxy-Pairs were recruited, 102 of which were treated on an intensive care unit. The agreement in the individual items of the WHODAS 2.0 was measured via the weighted Kappa and PABAK. The agreement in the total score was measured with the LOA according to Bland-Altman and the ICC. Agreement was defined as sufficient if the lower confidence interval of the weighted Kappa and PABAK was greater than 0,4, the ICC was greater than 0,5 and the LOA were smaller than 10 Points.

Results: According to the PABAK, the agreement between patients and their proxies was sufficient in most of the items. The ICC also indicates sufficient agreement (ICC: 0,724, 95%KI: 0,627 to 0,794) in the overall score. The LOA were outside of the limit of 10 Points (LOA -40,5 to 30,9). Proxies of patients that underwent intensive care treatment significantly underestimated the HrQoL of patients relative to proxies of patients not undergoing intensive care ($p=0,009$). Proxies who support the patient with their daily activities ($n=52$) showed a lower agreement than proxies who did not (ICC 0,498 vs. 0,754) and showed a significant underestimation the HrQoL of the patients ($p=0,009$). Age and gender of the patient or proxy,

as well as the living together and the relationship between the proxy and patient had no influence on the agreement.

Discussion: In this dissertation, the first evidence for the validity of the WHODAS 2.0 for questioning proxies about the premorbid HrQoL of patients could be shown. Although statistically there is sufficient agreement in most items and the overall score of the WHODAS 2.0, the wide LOA show that some patients were falsely classified as healthy or impaired by their respective proxies. Further research is needed to better assess whether proxy interviewing allows for the formulation of better patient-adapted therapeutic goals.

2 Einleitung

2.1 Hinführung zum Thema

Trotz großer Fortschritte in der intensivmedizinischen Versorgung entwickeln viele Patient*innen nach intensivmedizinischer Behandlung tiefgreifende psychische und physische Störungen. Diese können zu langfristigen Einschränkungen der Lebensqualität führen. Diese Folgeerscheinungen werden unter dem Begriff „Post intensive care Syndrom“ zusammengefasst (1). Hierzu gehören Störungen der Muskelfunktion, Gedächtnis- und Konzentrationsstörungen, Depressionen und posttraumatische Belastungsstörungen. Aktuelle Studie hierzu haben gezeigt, dass 64% der intensivmedizinisch behandelten Patient*innen, welche vorher berufstätig waren, auch 2 Jahre nach Entlassung aus einer Intensivstation keiner Arbeit nachgehen konnten. Bis zu 50% der Patient*innen welche vorher selbstständig lebten, benötigten auch noch nach 2 Jahren substantielle Unterstützung im Alltag (2). Diese Einschränkungen der Lebensqualität und körperlichen Leistungsfähigkeit wurden auch mehr als 5 Jahre nach intensivmedizinischer Behandlung beobachtet (3). Dies zeigt, dass im Rahmen der intensivmedizinischen Versorgung kritisch kranker Patient*innen, nicht nur das akute Überleben zu bedenken ist, sondern auch langfristige Folgen erfasst und therapiert werden müssen.

Um die langfristigen Einschränkungen der Patient*innen erfassen und Fortschritte im Verlauf der Therapie beurteilen zu können, werden klinische Test und standardisierte Fragebögen, wie zum Beispiel der Short Form 36 (SF-36), Barthel-Index oder der WHODAS 2.0, genutzt (4-6). Je nach Fragebogen und klinischer Fragestellung werden hierbei verschieden Gesundheitsaspekte abgefragt. Hierzu gehören unter anderem die Mobilität, die Fähigkeit zur Bewältigung alltäglicher Aufgaben, wie dem selbstständige Anziehen oder Waschen, sowie Einschränkungen des sozialen Lebens und dem Schweregrade emotionaler Probleme. Für die behandelnden Ärzt*innen und medizinischen Hilfsberufe ist es hierbei kritisch zu wissen, wie stark Patient*innen initial eingeschränkt waren. Durch diese Einschätzung der prämorbidem Lebensqualität und Fähigkeiten der Patient*innen, können individuell angepasste Therapieziele gesetzt werden und Folgeerscheinungen von vorbestehenden Einschränkungen differenziert werden. Dies ist besonders relevant, da aufgrund des steigenden Alters der Bevölkerung und ausgedehnterer intensivmedizinischer Behandlungsmöglichkeiten mehr ältere Patient*innen mit relevanten Vorerkrankungen auf Intensivstationen behandelt werden (7-9).

Kenntnisse über den prämorbidem Status der Patient*innen sind auch für die akute Therapieplanung von Bedeutung. In verschiedenen Studien konnte gezeigt werden, dass Patient*innen mit vorbestehender Einschränkung der HrQoL eine erhöhte Mortalität aufweisen und häufiger Komplikationen wie z.B. Multiorganversagen, nosokomialen Pneumonien oder Komplikationen bei der Beatmungsentwöhnung entwickeln (10-13).

Die Erhebung der prämorbidem HrQoL ist im Rahmen der intensivmedizinischen Behandlungssituation oft erschwert oder nur indirekt möglich, da Patient*innen oft komatös, delirant oder sediert sind. Um diese Informationen trotzdem zu erlangen, werden Stellvertreter*innen befragt. Die Verlässlichkeit von Stellvertreter*innen bei der Beurteilung der prämorbidem gesundheitsbezogenen Lebensqualität von Patient*innen ist Thema dieser Dissertation.

Verlässlichkeit von Fremdbewertungen ist auch in anderen Bereichen der Medizin relevant, wie zum Beispiel bei der Beurteilung der Lebensqualität bei Demenzerkrankungen, nach einem Schlaganfall oder bei Kindern. Fremdbewertungen der Gesundheit werden dann eingesetzt, wenn Patient*innen selbst nicht in der Lage sind, glaubhafte Antworten zu geben (14). Bei der Befragung von Stellvertreter*innen auf Intensivstationen soll die Fremdeinschätzung die Eigeneinschätzung der Patient*innen ersetzen. Dies bedeutet, dass eine Fremdeinschätzung als verlässlich eingestuft wird, wenn diese der Eigeneinschätzung entspricht. Dies wird auch als „Proxy-Patient“-Perspektive beschrieben (15). Um die Verlässlichkeit von „Proxy-Patient“-Fremdeinschätzungen zu beurteilen, wird diese in Studien typischerweise mit der Eigenbewertung der Patient*innen verglichen (14).

Es ist zu beachten, dass verschiedene Faktoren einen Einfluss auf die Übereinstimmung haben. Hierzu gehört unter anderem die Art des abgefragten Gesundheitsaspektes. Bei Gesundheitsaspekten, die von außen sichtbar sind, wie z. B. der Mobilität, liegt häufig eine höhere Übereinstimmung vor, als bei inneren Prozessen, wie z. B. Angst, Depression oder die emotionalere Belastung (14). In der Literatur werden weitere Einflussfaktoren genannt, wie z. B. das Verhältnis zwischen den Patient*innen und Stellvertreter*innen, die spezifische Erkrankung, die Erkrankungsschwere, die kognitiven Fähigkeiten, der Bildungsstand und auch das verwendete Erhebungsinstrument (15).

Bei der Befragung von Stellvertreter*innen im Rahmen der intensivmedizinischen Behandlung sind auch weitere Einflussfaktoren möglich, dessen Einfluss in dieser Dissertation untersucht werden soll. Zum einen wird nicht nach der aktuellen Lebensqualität gefragt, sondern nach der

Lebensqualität, wie sie vor der Erkrankung war. Zum anderen kann durch die schwere und oft akute Erkrankung, sowie die Umgebung der Intensivstation eine emotionale Belastung vorliegen.

Neben der Behandlungsplanung und der Überprüfung des Therapieerfolges ist die Verlässlichkeit von Fremdeinschätzung der prämorbidem HrQoL auch im Bezug auf Entscheidungen zum Lebensende von großer Bedeutung. Hier kann es dazu kommen, dass, wenn die prämorbidem Lebensqualität als zu gering eingeschätzt wird, Entscheidungen getroffen werden, welche nicht mit dem Patientenwillen übereinstimmen (16)

2.2 ICF und WHODAS 2.0

In dieser Dissertation wurde zur Erhebung der gesundheitsbezogenen Lebensqualität der WHODAS 2.0 (World Health Organisation Disability Assessment Schedule) Fragebogen verwendet. Dieser Fragebogen basiert auf der Gesundheitsklassifizierung der „International Classification of Functioning, Disability and Health“ (ICF) und bildet verschiedene Aspekte der gesundheitsbezogenen Lebensqualität ab. In dem WHODAS 2.0 werden neben körperlichen Einschränkungen, auch Einschränkungen der gesellschaftlichen Partizipation berücksichtigt. Der ICF ist dazu konzipiert, möglichst universal anwendbar zu sein, sodass diese Klassifikation krankheitsübergreifend und länderübergreifend anwendbar ist (17). Der genaue Aufbau des WHODAS 2.0 Fragebogens wird im Methodik-Teil dieser Dissertation näher erläutert.

2.3 Stand der Literatur

Um relevante Studien zu diesem Thema zu finden, wurde eine Literaturrecherche in der medizinischen Bibliothek des „National Institutes of Health“ (PubMed) durchgeführt (18). Hierbei wurden nach Studien gesucht, in welchen die Verlässlichkeit von Stellvertretern bei der Beurteilung der prämorbidem Lebensqualität von intensivmedizinisch behandelten Patient*innen untersucht wurde.

Eine Suche am 24.10.2019, mit folgendem Suchbegriff ergab 196 Treffer: (*“proxy-reported”*[MAJR] OR *“Proxy-informants”*[MAJR] OR *“proxy”*[MAJR] OR *“Inter Rater Reliability”*) AND (*“Health Status”* OR *“HrQoL”* OR *“QoL”* OR *“Quality of Life”*[MAJR] OR *“Quality of Life/psychology”* OR *“activities of daily living”* OR *“adl”* OR *“WHODAS”* OR *“IADL”* OR *“Disability Evaluation”* OR *“Barthel”*). Nach Screening der Titel, Abstracts und Lesen der Volltexte, konnten 7 relevante Studien identifiziert werden (19-25). Weiterhin wurden die Literaturverweise der gefundenen Studien durchsucht und so 6 weitere Studien ermittelt

(26-31). Bei keiner der erfassten Studien, wurde die Verlässlichkeit der Fremdbewertung der Gesundheit mittels des WHODAS 2.0 überprüft. Um mögliche Studien zu identifizieren welche die die Verlässlichkeit vom Fremdeinschätzungen durch den WHODAS 2.0 in anderen klinischen Situationen beurteilen wurde am 30.10.2019, eine Suche mit folgendem Suchbegriff durchgeführt: (*"interraterreliability" OR "Proxy" OR "proxy-reported" OR "Proxy-informants"*) AND (*"WHODAS"*). Diese Suche ergab 25 Treffer. Nach Screening der Titel, Abstracts und lesen desrVolltexte, konnten 2 relevante Studien identifiziert werden (32, 33). Eine Übersicht zu allen erfassten Studien, mit Beschreibung der untersuchten Population und den verwendeten Fragebögen ist in der Tabelle 1 dargestellt.

Bei allen Studien wurden die Patient*innen retrospektiv zu ihrer prämorbidem Lebensqualität befragt. Dies hat den Nachteil, dass Patient*innen dadurch potentiell dazu neigen, ihre Lebensqualität, wie sie vor einer Erkrankung war, zu überschätzen (13).

Tabelle 1: Stand der Literatur

	Autor	Jahr	Anzahl	Population	Fragebögen
Prämorbide Lebensqualität	Maxwell et al.	2015	77	Patient*innen über 65 Jahre	VES-13 Barthel-Index LSA
	Dinglas et al.	2013	140	ARDS Überlebende	EQ-5D
	Hofhuis et al.	2011	140	Intensivmedizinische und kardiologische Patient*innen	SF-12 (PCS) ALDS PCS
	Gifford et al.	2010	136	ARDS Überlebende	SF-36
	Elliott et al.	2006	96	Herzchirurgische Patient*innen	SF-36
	Sacles et al.	2006	46	ARDS-Überlebende	SF-36
	Hofhuis et al.	2003	112	55 Elektive Operationen 52 Notfall Operationen	SF-36
	Chaboyer et al.	2002	20	Patient*innen auf Intensivstation	Eigener Fragebogen
	Capuzzo et al.	1999	172	Patient*innen auf Intensivstation	QoL-Italien QoL-Spanien
	Diaz-Prieto et al.	1998	523	Patient*innen mit Verletzungen oder Operationen	EQ-5D Kanofsky Index
	Rogers et al.	1997	99	Patient*innen auf Intensivstation	SF-36
	Bidia et al.	1996	103	Patient*innen auf Intensivstation	EQ-5D
	Magaziner et al.	1988	386	Patient*innen mit Beckenfraktur	OARS
Aktuelle Lebensqualität	Rames et al.	2014	312	Rheumatischen Erkrankungen	WHODAS 2.0
	Schlote et al.	2008	84	Schlaganfall	WHODAS 2.0

ARDS = Acute Respiratory Distress Syndrome, VES-13 = Vulnerable Elder Survey, LSA = Life Space Assessment. EQ5D = European Quality of Life 5 Dimensions, SF-12 = Short Form 12, ALDS = Academic Medical Center Linear Disability Score, PCS = Physical Component Score des SF-36, SF36 = Short Form 36, QoL = Quality of Life; OARS = Older American Resources Services

In 7 der Studien wurde neben der Übereinstimmung zwischen den Patient*innen und deren Stellvertreter*innen, der Einfluss von verschiedenen Faktoren auf die Übereinstimmung untersucht (21, 22, 24, 25, 28, 31) (vgl. Tabelle 2). In allen dieser Studien wurden die Stellvertreter*innen anhand der festgelegten Charakteristika in dichotome Gruppen eingeteilt.

Tabelle 2: In anderen Studien untersuchte Einflussfaktoren auf die Übereinstimmung zwischen Patient*innen und deren Stellvertreter*innen

	Patientenabhängige Faktoren	Stellvertreterabhängige Faktoren	Umgebungsabhängige Faktoren
Hofhuis et al. 2011	-	-	Einlieferungsgrund: elektive Operation oder Notfall
Scales et al. 2006	Alter, Geschlecht, Krankheitsstaus	-	-
Hofhuis et al. 2003	-	-	Einlieferungsgrund: elektive Operation oder Notfall
Capuzzo et al. 1999	-	Geschlecht, Verhältnis	-
Diaz-Prieto et al. 1998	-	Geschlecht, Bildungsstand	-
Magaziner et al. 1988	Kognitive Funktion, Depressivität	Alter, Geschlecht, Zusammenleben, Verhältnis, Selbsteinschätzung zu Wissen über die Gesundheit der Patient*innen	-
Schlote et al. 2014	Art der Rheumatischen Erkrankung	-	-

Im Rahmen der Literaturrecherche zeigte sich, dass es keine Studie gibt, welche den Einsatz des WHODAS 2.0 Fragebogens bei der Fremdbewertung der präorbiden HrQoL von Patient*innen untersucht, welche intensivmedizinisch behandelt werden. Es gibt weiterhin keine Studie, in welcher der Einfluss einer intensivmedizinischen Behandlung auf diese Übereinstimmung untersucht wird. Weitere Forschung zu diesem Thema ist somit dringend notwendig.

2.4 Fragestellung

Ziel dieser Dissertation ist es, das Ausmaß der Übereinstimmung zwischen Patient*innen und deren Stellvertreter*innen bei der Einschätzung der prämorbidem Lebensqualität, mittels der 12 Fragen Version des WHODAS 2.0 zu ermitteln. Weiterhin soll der Einfluss einer intensivmedizinischen Behandlung der Patient*innen auf die genannte Übereinstimmung untersucht werden. Ein weiteres Ziel dieser Dissertation ist es Gruppen von Stellvertreter*innen und Patient*innen zu identifizieren, bei welchen eine besonders hohe Übereinstimmung vorliegt.

Hierzu wurden folgende Fragestellungen und Hypothesen aufgestellt. Bei allen Fragestellungen bezieht sich die erwähnte „Übereinstimmung“, auf die Übereinstimmung zwischen Patient*innen und deren Stellvertreter*innen in der Beurteilung der prämorbidem Lebensqualität mittels der 12 Fragen Version des WHODAS 2.0 Fragebogens. Die verwendeten Begriffe „klinisch ausreichend“ und „relevant“ sowie die genauen Charakteristika der Fragestellung 3 werden im Methodik-Teil dieser Dissertation erläutert.

A: Gibt es eine klinisch ausreichende Übereinstimmung bei der Beurteilung der prämorbidem gesundheitsbezogenen Lebensqualität der Patient*innen?

H0: Es gibt keine klinisch ausreichende Übereinstimmung.

H1: Es gibt eine klinisch ausreichende Übereinstimmung.

B: Gibt es einen relevanten Unterschied in der Übereinstimmung zwischen Patient*innen die intensivmedizinisch behandelt wurden und Patient*innen, die nicht intensivmedizinisch behandelt wurden?

H0: Es gibt keinen relevanten Unterschied.

H1: Es gibt einen relevanten Unterschied.

C: Haben die untersuchten Charakteristika einen relevanten Einfluss auf die Übereinstimmung?

H0: Die Charakteristika haben keinen relevanten Einfluss.

H1: Die Charakteristika haben einen relevanten Einfluss.

3 Methodik

3.1 Studiendesign

Die Daten dieser Dissertation wurden im Rahmen der ASTON-Studie unter der Leitung von Prof. Dr. Weber Carstens und Prof. Dr. Stefan Schaller erhoben. Die ASTON Studie wurde als prospektive multizentrische Observationsstudie geplant und durchgeführt. Die Datenerhebung hat an den Standorten Berlin, im Charité Campus Virchow Klinikum (Leitung: Prof. Weber Carstens) und München, im Klinikum der technischen Universität München (Leitung: Prof. Dr. Stefan Schaller) stattgefunden.

3.2 Ethikvotum

Die ASTON Studie wurde in Übereinstimmung mit der Deklaration von Helsinki durchgeführt. Vor Beginn der Datenerhebung wurde ein Ethikvotum der verantwortlichen Ethikkommission eingeholt (Aktenzeichen 542/18 S-AS).

3.3 Rekrutierung

Um Patient*innen vor ihrer medizinischen Behandlung befragen zu können, wurden Patient*innen rekrutiert bei welchen eine Operation geplant war. Die Belastung und Einschränkungen der Gesundheit nach einer Operation wurden als ausreichend angesehen, um die Befragungssituation bei einer akuten Erkrankung oder Trauma widerzuspiegeln.

Die Rekrutierung und Befragung der Patient*innen geschah in den jeweiligen Krankenhäusern vor der Operation in der präoperativen Anästhesiesprechstunde. Potenzielle Studienteilnehmer wurden durch die in der Anästhesiesprechstunde vorliegenden Akten, der Klinik internen Anästhesieplanungssoftware (COBRA 6) und der elektronischen Patientenakte (SAP/Orbis) identifiziert. Die Patient*innen wurden nach Befragung dazu aufgefordert, ihre Antworten nicht mit ihren Angehörigen oder Bekannten zu teilen, um eine Verfälschung der Ergebnisse zu verhindern. Die Stellvertreter*innen wurden bei Besuch der Patient*innen im Krankenhaus nach der Operation rekrutiert und separat von den Patient*innen befragt.

Vor Aufnahme in die Studie erfolgte eine mündlich und schriftlich Aufklärung über die Zielsetzung und den Ablauf der Studie sowie die Freiwilligkeit der Teilnahme. Vor der Befragung wurde ebenfalls eine schriftliche Einwilligung zur Teilnahme an der Studie und zur pseudonymisierten Speicherung und Verarbeitung der personenbezogenen Daten eingeholt.

3.4 Einschluss- und Ausschlusskriterien

Einschlusskriterien in die Studie war eine geplante Operation an den Studienstandorten, die Volljährigkeit der Proband*innen und eine schriftliche Einwilligung in die Studie. Proband*innen wurden nicht eingeschlossen, wenn diese aufgrund von kognitiven Einschränkungen oder esprachlicher Barriere nicht in der Lage waren die Fragebögen eigenständig auszufüllen. Proband*innen wurden aus der Studie ausgeschlossen, wenn keine Stellvertreter*innen befragt werden konnten, diese vor Befragung der Stellvertreter*innen verstorben sind oder die Einwilligung zur Teilnahme and der Studie widerrufen wurde.

3.5 Erhebungsinstrumente

Zur Erfassung der prämorbidem Lebensqualität, wurde in dieser Dissertation die deutsche 12 Fragen Version des WHODAS 2.0(34). Dabei wurde jeweils eine individuelle Version zur Eigeneinschätzung und Fremdeinschätzung genutzt. Die kognitiven Fähigkeiten bzw. Einschränkungen der Patient*innen wurden mittels des MOCA-Blind erfasst (35). Die Stellvertreter*innen wurden weiterhin zu ihrem Geschlecht und Alter sowie zu ihrer Beziehung zu den Patient*innen befragt. Es wurde weiterhin erhoben, ob sie mit den Patient*innen zusammenleben und ob sie die Patient*innen pflegen.. Die in dieser Studie genutzten Version des WHOODAS 2.0 und MOCA-Blind sind im Anhang hinterlegt.

Im Rahmen der ASTON-Studie wurden die Patient*innen und deren Stellvertreter*innen, mittels 4 weiterer Fragebögen zu ihrer Lebensqualität, Alltagskompetenz befragt. Diese sind nicht Teil dieser Dissertation und werden deshalb nicht weiter erläutert.

3.5.1 WHODAS 2.0

Der WHODAS 2.0 umfasst die Gesundheitsdomänen „Kognition“, „Mobilität“, „Selbstversorgung“, „soziale Interaktion“, „Lebensaktivitäten“ und „gesellschaftliche Teilnahme“. In dieser Studie wurde die 12 Fragen Version des WHODAS 2.0 verwendet. Bei dieser Version werden die Einschränkungen in den einzelnen Gesundheitsdomänen mit jeweils 2 Fragen abgefragt. Hierbei werden nur Einschränkungen der Lebensqualität aufgrund von gesundheitlichen Problemen berücksichtigt.

Die ersten 12-Fragen des WHODAS 2.0 sind lickert-skaliert mit jeweils 5 Antwortmöglichkeiten, welche von „keine“ bis „sehr starke“ Schwierigkeiten, mit den Zwischenstufen „leicht“, „mäßig“ und „stark“ reichen. Im Rahmen der Auswertung wurden die verschiedenen Einschränkungsgrade mit 0 Punkten für die Antwort „keine

Einschränkungen“ und bis zu 4 Punkten für die Antwort „sehr starke Schwierigkeiten“ beziffert. Bei den einzelnen Gesundheitsdomänen können somit jeweils 0 bis 8 Punkte erreicht werden, wobei eine höhere Punktzahl eine größere Einschränkung bedeutet.

In der 12-Fragen Version gibt es 3 zusätzliche Fragen, welche die Häufigkeit und Schwere der gesundheitlichen Probleme in den letzten 30 Tagen erfragen. Die Punktwerte dieser 3 Fragen gehen nicht in die Gesamtpunktzahl oder die Punktwerte der Gesundheitsdomänen ein. Die Übereinstimmung in diesen Fragen wurde deswegen in dieser Dissertation nicht analysiert.

Bei der 12-Fragen-Version des WHODAS 2.0 können maximal 48 Punkte bzw. 100 Prozent erreicht werden. In dieser Studie wurde im Weiteren mit 100 Prozent gearbeitet, welche auch in der offiziellen Berechnung der WHO verwendet werden (17). Ein höherer Prozentwert ist gleichbedeutend mit einer größeren Einschränkung. Im Rahmen der weiteren Auswertung wurde ein Prozentwert von 0 als keine, 1-9 als leichte, 10 bis 20 als moderate und über 20 als schwere Einschränkung klassifiziert (36).

3.5.2 MOCA-Blind

Der MOCA-Blind ist eine abgewandelte Form des MOCA Testes, bei welchem die visuellen Aufgaben herausgenommen sind. Im MOCA-Blind werden das Gedächtnis, die Aufmerksamkeit, das Sprachverständnis, die Abstraktionsfähigkeit und die Orientierung überprüft. Beim MOCA-Blind können maximal 22 Punkte, anstatt der 30 Punkte des MOCA Testes mit visuellen Aufgaben erreicht werden. Ein Punktwert von mindestens 18 Punkten ist dabei als normal zu werten (37).

3.6 Statistische Analyse

Zur Speicherung der Daten wurde Excel 2016 (Microsoft Corp. Redmond, WA) verwendet. Für die statistische Analyse der Daten wurde ebenfalls Excel 2016 sowie das „Statistik Package for Sozial Sciences“ (SPSS Version 27.0, DM Corp. Armonk, NY) verwendet. Das Signifikanzniveau wurde auf $p < 0,05$ festgelegt.

Fallzahlberechnung

Die Fallzahl dieser Dissertation wurde anhand der Fragestellung 2 dieser Dissertation gewählt. Die Fallzahl wurde so gewählt, um einen mindestens moderaten Unterschied (Cohens $d = 0,4$) in der Verteilung bzw. im Mittelwert zwischen den Gruppen nachweisen zu können (38). Hierzu

waren bei einem Signifikanzniveau von 0,05 und einem β -Fehler von 0,20 mindestens 82 Patient*innen pro Gruppe notwendig. Zur Power-Berechnung wurde das Programm „G*Power“ Version 3.1.9.7 der Heinrich Heine Universität Düsseldorf verwendet (39).

Skalenniveaus der Daten

Die einzelnen Fragen WHODAS 2.0 sind ordinal skaliert. Die Punktzahlen der Gesundheitsdomänen und der Summenpunktzahl ergeben sich aus den einzelnen Fragen und sind demzufolge auch ordinal skaliert. Im Bereich der sozialmedizinischen Forschung wird das Konzept von „quasimetrischen“ bzw. „pseudometrischen“ Variablen verwendet. Dieser Begriff ist vor allem bei Lickert-Skalen, wie sie hier vorliegen, anwendbar (40). Im Rahmen dieser Dissertation wurde die Gesamtpunktzahl des WHODAS 2.0 als semi-metrisch betrachtet und wurde mittels der im weiteren beschriebenen Analyse nach Bland-Altman und der Intraklassenkorrelation (ICC) ausgewertet.

Testung auf Normalverteilung

Die Daten wurden mittels des Kolmogorov-Smirnov-Test auf eine Normalverteilung geprüft (41). Hierbei zeigte sich, dass keiner der Parameter Normalverteilt ist ($p < 0,001$). Deshalb wurde als Lagemaß der Median und als Verteilungsmaß der Interquartilsabstand (IQA) verwendet.

Populationscharakteristika

Bei der Bestimmung der Populationscharakteristika wurde das durchschnittliche Alter und die Geschlechtsverteilung der Patient*innen und Stellvertreter*innen angegeben. Bei der Beschreibung der Patient*innen wurde zusätzlich angegeben, ob diese postoperativ auf einer Intensivstation oder Normalstation behandelt wurden, die Dauer der intensivmedizinischen Behandlung und die Dauer des Krankenhausaufenthaltes.

Analyse zur Fragestellung 1

Die Übereinstimmung zwischen den Patient*innen und deren Stellvertreter*innen wurde jeweils für die Einzelfragen, die Gesundheitsdomänen und die Gesamtpunktzahl des WHODAS 2.0 ermittelt.

Im Rahmen der deskriptiven Analyse wurde die Häufigkeit der jeweiligen Antworten in den einzelnen Fragen des WHODAS 2.0 angegeben, sowie die Häufigkeit der jeweiligen Punktwerte in den Gesundheitsdomänen. Weiterhin wurde der Prozentsatz gleicher Antworten bzw. Punktwerte berechnet, sowie der Prozentsatz von Über- und Unterschätzungen der prämorbidem Lebensqualität durch die Stellvertreter*innen.

Im Zuge dessen, wurde ebenfalls der Prozentsatz an ähnlichen Punktwerten angegeben. Ähnliche Punktwerte, wurden bei den Einzelfragen als maximal 1 Kategorie Abweichung, bei den Domänen als maximal 2 Punkte Abweichung und bei der Gesamtpunktzahl als maximal 10 Punkte Abweichung definiert. Da bei dem WHODAS 2.0 Fragebogen kein Grenzwert für einen minimal klinisch relevanten Unterschied (Minimal Clinically Important Differenz; MCID) bei intensivmedizinisch behandelte Patient*innen etabliert ist, wurden diese Werte anhand klinischer Überlegungen und dem Vergleich zu der genutzten MCID des WHODAS 2.0 in anderen Studien festgelegt (36, 42-45). Bei der Gesamtpunktzahl wurde ebenfalls der Prozentsatz von Antworten, die um maximal 25 Punkte abweichen berechnet. Dieser Wert entspricht einer Abweichung von je einer Kategorie in jeder der einzelnen Fragen.

Um zu testen, ob die Stellvertreter*innen die Einschränkungen insgesamt über- oder unterschätzen, wurde die Differenz zwischen den jeweiligen Punktwerten berechnet. Bei der Berechnung der Differenz, wurde der von den Stellvertreter*innen angegebene Punktwert von dem der Patient*innen abgezogen. Eine negative Differenz entspricht somit einer Überschätzung der Einschränkungen durch die Stellvertreter*innen und umgekehrt. Diese Differenz wurde als klinisch relevant erachtet, wenn diese größer als die MCID war.

Um die Übereinstimmung zwischen den Patient*innen und deren Stellvertreter*innen in der Gesamtpunktzahl des WHODAS 2.0 zu ermitteln, wurde die Bland-Altman-Analyse und die ICC verwendet. Bei der Bland-Altman-Analyse wird der Mittelwert der Differenz zwischen den Punktwerten der Patient*innen und Stellvertreter*innen, sowie die oberen und unteren Grenzen der Übereinstimmung berechnet (Limits of Agreement = LOA). Die LOA geben den Bereich an, in welcher 95 % der Differenzen liegen und entsprechen somit dem 95 % Konfidenzintervall der Differenz. Die LOA werden berechnet, indem zu dem Mittelwert der

Differenz die 1,96-fache der Standarddeviation der Differenz addiert bzw. subtrahiert wird (46). Wenn die LOA kleiner als die MCID (hier 10 Punkte) sind, ist die Übereinstimmung als ausreichend bewertet worden.

Die ICC ist ein Maß für die Übereinstimmung zwischen verschiedenen Beurteilern und wird für metrische Daten verwendet (47). Die ICC kann von 0 bis 1 reichen, wobei verschiedene Vorschläge für die Interpretation dieses Wertes existieren. In dieser Dissertation wurden die Werte nach von Terry K. Koo et al. etablierten Grenzwerten interpretiert. Dabei wurden Werte von 0 bis $< 0,5$ als eine schlechte, $0,5$ bis $< 0,75$ als eine moderate, $0,75$ bis $< 0,9$ als eine gute und Werte über $0,9$ als eine exzellente Übereinstimmung gewertet (48). Die Übereinstimmung wurde als ausreichend gewertet, wenn das untere Konfidenzintervall der ICC größer als $0,5$ war. Bei der Berechnung der ICC wurde das „One-Way-Random“ (ICC (1,1)) Modell benutzt (48).

Um die Übereinstimmung in den Einzelfragen und Gesundheitsdomänen zu ermitteln, wurde das gewichtete Kappa nach Jacob Cohen bestimmt (49). Das gewichtete Kappa beschreibt die nicht-zufällige Übereinstimmung zwischen zwei verschiedenen Beurteilern. Zur Berechnung des gewichteten Kappa wird aus den Antworten der Patient*innen und deren Stellvertreter*innen eine Kreuztabelle gebildet und die tatsächliche Übereinstimmung mit der Übereinstimmung verglichen, wie sie zufällig zustande kommen würde. Anders als bei dem nicht-gewichteten Kappa werden hier, je nach Höhe der Abweichung, Gewichtungen verteilt. Diese gehen in die Berechnung des gewichteten Kappa-Wertes mit ein.

Als etablierte Methoden der Gewichtung werden die lineare und quadratische Gewichtung von Abweichungen beschrieben (49). In dieser Dissertation wurde diese lineare Gewichtung verwendet. Dabei wird eine genaue Übereinstimmung maximal gewichtet mit einem linearen Abfall der Höhe der Gewichtung, je weiter die Antwort der Stellvertreter*innen von der Antwort der Patient*innen entfernt ist. In der Abbildung 1 ist die Gewichte-Tabelle dargestellt, welche für die Auswertung der einzelnen Fragen des WHODAS 2.0 benutzt wurde. Für die Berechnung der gewichteten Kappa-Werte in den Gesundheitsdomänen des WHODAS 2.0 wurde ebenfalls eine lineare Gewichtung genutzt.

		Antwort der Patient*innen				
		Keine Probleme	Leichte Probleme	Mäßige Probleme	Starke Probleme	Sehr starke Probleme
Antwort der Stellvertreter*innen	Keine Probleme	1	0,75	0,5	0,25	0
	Leichte Probleme	0,75	1	0,75	0,5	0,25
	Mäßige Probleme	0,5	0,75	1	0,75	0,5
	Starke Probleme	0,25	0,5	0,75	1	0,75
	Sehr starke Probleme	0	0,25	0,5	0,75	1

Abbildung 1: Lineare Gewichtung für die Fragen 1 bis 12 des WHODAS 2.0

Für den gewichteten Kappa-Wert ist eine Neigung zu potentiell falsch niedrigen Werten beschrieben, wenn bei den Antworten eine Kategorie übermäßig repräsentiert ist, also z. B. ein Großteil der Befragten antworten, dass sie „keine Probleme“ oder „sehr starke Probleme“ haben. Durch diese Eigenschaft des gewichteten Kappa-Wertes kann trotz eines hohen Prozentsatzes gleicher Antworten, ein niedriger Kappa-Werte vorliegen (50, 51).

Aus diesem Grund wurde neben dem gewichteten Kappa-Wert auch der PABAK-Wert (Prevalence adjusted and Bias adjusted Kappa) berechnet. Zur Berechnung dieses Wertes wurde der Web basierte Rechner der Single Case Research Gruppe verwendet (52). Bei der Berechnung des PABAK-Wertes wurde ebenfalls eine lineare Gewichtung verwendet.

Die Werte des gewichteten Kappa-Wertes und PABAK-Wertes entsprechen nicht dem Prozentsatz gleicher Antworten, sondern sind eine dimensionslose Kennzahl von -1 bis 1. Die Interpretation dieses Wertes wurde nach Einteilung von Landis und Koch vorgenommen. Hiernach wird ein Kappa-Wert von 0 bis 0,2 als geringe Übereinstimmung gewertet, ein Wert von 0,21 bis 0,4 als mäßige, 0,41 bis 0,6 als moderate, 0,61 bis 0,8 als gute und ein Wert von 0,81 bis 1 als fast perfekte Übereinstimmung gewertet. Negative Werte des Kappa-Wertes bedeuten, dass die gemessene Übereinstimmung geringer als eine rein zufällige Übereinstimmung ist (53). Im Rahmen dieser Dissertation wurden nur Werte, dessen unteres

95 % Konfidenzintervall über 0,4 liegt als ausreichend gewertet. Die Konfidenzintervalle der gewichteten Kappa-Werte wurden nach den Empfehlungen von Fleiß und Cohen berechnet (54).

Außerdem wurde getestet, ob der Anteil an Patient*innen, welche angegeben haben, „keine Einschränkungen“ zu haben, einen Einfluss auf die Übereinstimmung hat. Hierzu wurde eine lineare Regressionsanalyse durchgeführt.

Analyse zu Fragestellung 2

Zur Beantwortung der Frage, welchen Einfluss die Behandlung des Patient*innen auf einer Intensivstation auf die Übereinstimmung hat, wurde die Population anhand der postoperativen Station der Patient*innen (Intensivstation oder Normalstation) in 2 Gruppen aufgeteilt. Bei diesem Vergleich wurde nur die Übereinstimmung in der Gesamtpunktzahl und der Gesundheitsdomäne betrachtet. Für die jeweiligen Gruppen wurde die mediane Punktzahl in den Gesundheitsdomänen und der Gesamtpunktzahl des WHODAS 2.0, sowie das Alter und das Geschlecht der Befragten bestimmt. Diese Charakteristika wurden, mittels Mann-Whitney-U-Test bzw. Chi-Quadrat-Test, auf signifikante Unterschiede getestet.

Bei dem Gruppenvergleich wurde mittels des Mann-Whitney-U-Test getestet, ob ein systematischer Bias zur Über- oder Unterschätzung der prämorbid Lebensqualität zwischen den Gruppen vorliegt. Um weiteren Unterschiede in der interindividuellen Übereinstimmung zu ermitteln, wurden für die jeweiligen Gruppen, die ICC und LOA sowie die gewichteten Kappa-Werte und PABAK-Werte ermittelt und zwischen den Gruppen verglichen. Dabei wurde ein Unterschied von 0,25 bzw. 0,2 in den ICC bzw. Kappa-Werten als relevant festgelegt. Diese Werte entsprechen der Breite der jeweiligen Kategorien in der Interpretation der ICC bzw. Kappa-Werte (48, 53). Die LOA wurden deskriptiv verglichen.

Analyse zur Fragestellung 3

Um zu überprüfen, ob ausgewählte Charakteristika einen Einfluss auf die Übereinstimmung haben, wurde die Population anhand der zu untersuchenden Charakteristika in dichotome Gruppen eingeteilt. Aufgrund des erhöhten Anteils an Gruppen wurde bei diesem Vergleich nur die Übereinstimmung in der Gesamtpunktzahl verglichen. Dabei wurde getestet, ob ein signifikanter Bias zwischen den Gruppen bzw. relevante Unterschiede in den ICC und LOA vorliegen. Dabei wurden die gleichen Grenzwerte für relevante Unterschiede wie in der Fragestellung 2 genutzt. Aufgrund der erhöhten Anzahl an Gruppen wurde nicht für jeden

Vergleich ein komplettes Bland-Altman-Diagramm erstellt, sondern nur der Mittelwert der Differenz und die unteren und oberen LOA angegeben. Die genaue Gruppeneinteilung ist in Tabelle 3 beschrieben.

Für die jeweiligen Gruppen wurde die Anzahl der Proband*innen angegeben sowie die mediane Gesamtpunktzahl der Patient*innen im WHODAS 2.0 berechnet und mittels Mann-Whitney-U-Test auf signifikante Unterschiede getestet. Bei der Gruppe „Pflege“ wurde weiterhin ein Fallkontrollabgleich anhand der prämorbidem Einschränkung der Patient*innen durchgeführt, da signifikante Unterschiede zwischen den beiden Gruppen vorlagen.

Tabelle 3: In Fragestellung 3 untersuchte Charakteristika und Gruppeneinteilung

Einflussfaktor	Gruppe 1	Gruppe 2
Leben in häuslicher Gemeinschaft	Ja	Nein
Verhältnis zwischen dem Patient*innen und Stellvertreter*innen	Ehepartner	Andere
Pflege des Patient*innen durch den Stellvertreter*innen	Ja	Nein
Geschlecht der Stellvertreter*innen	Weiblich	Männlich
Geschlecht der Patient*innen	Weiblich	Männlich
Alter der Patient*innen	Maximal 65 Jahre	Mindestens 66 Jahre
Alter der Stellvertreter*innen	Maximal 65 Jahre	Mindestens 66 Jahre

4 Ergebnisse

4.1 Populationscharakteristika

Charakteristika der Patient*innen

Von den 204 eingeschlossenen Patient*innen, waren 83 (40,7 %) weiblich und 121 (59,3 %) männlich. Das mittlere Alter der Patient*innen betrug 67 Jahre (Min: 20, Max: 90, SD 13,5). 102 (50 %) Patient*innen wurden postoperativ auf einer Intensivstation behandelt, mit einer durchschnittlichen Liegezeit auf der Intensivstation von 3,84 Tagen (Min: 1, Max: 25, SD: 3,6). 102 (50 %) Patient*innen wurden postoperativ auf einer Normalstation behandelt. Die durchschnittliche Behandlungsdauer aller Patient*innen lag bei 18,6 Tagen (Min: 2, Max: 192, SD: 20,4).

Charakteristika der Stellvertreter*innen

Insgesamt wurden 204 Stellvertreter*innen, davon 138 (67,6 %) weibliche und 65 (32,0 %) männliche Stellvertreter*innen befragt, ein Stellvertreter*innen hat das Geschlecht nicht angegeben. Das mittlere Alter der befragten Stellvertreter*innen betrug 61 Jahre (Min: 24, Max: 87, SD: 13,7). Von den Stellvertreter*innen der 102 Patient*innen, welche postoperativ auf einer Intensivstation behandelt wurden, wurden 38 (37,3 %) der Stellvertreter*innen bei Besuch des Patient*innen auf der Intensivstation befragt, die restlichen 64 (62,7 %) wurden erst befragt, nachdem der Patient*innen bereits auf eine Normalstation verlegt worden war.

Fehlende Werte

Mit fehlenden Werten in dem WHODAS 2.0 wurde nach den Richtlinien der WHO verfahren. Dabei wurde, falls maximal eine Antwort bei einer Frage fehlte, bei dieser Frage der Mittelwert der Punktwerte in den restlichen Fragen eingesetzt. Wenn mehr als eine Frage nicht beantwortet wurde, wurde die Gesamtpunktzahl nicht berechnet (55).

4.2 Ergebnisse zur Fragestellung 1

4.2.1 Verteilung

Verteilung der Antworten in den einzelnen Fragen des WHODAS 2.0

In der Abbildung 2 sind die Häufigkeiten der Antworten in den einzelnen Fragen des WHODAS 2.0 dargestellt. Bei allen Fragen, außer Frage 5, wurde von den Patient*innen und Stellvertreter*innen am häufigsten angegeben, dass „keine Probleme“ vorliegen. Besonders hoch ist dieser Prozentsatz auch bei den Fragen 3 und 8 bis 11.

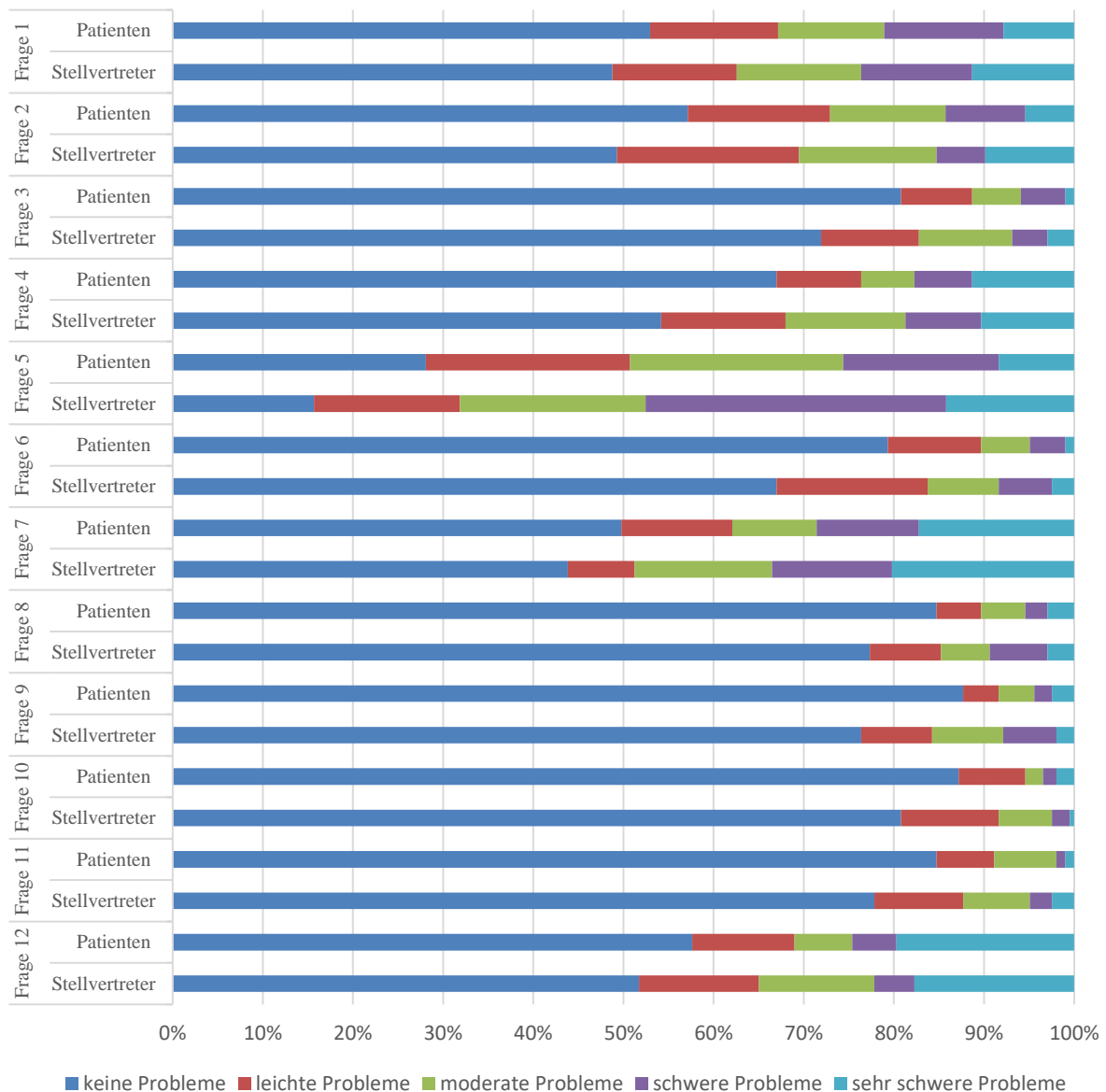


Abbildung 2: Verteilung der Antworten in den Fragen 1 bis 12 des WHODAS 2.0; n = 204.

Verteilung der Punktwerte in den Gesundheitsdomänen des WHODAS 2.0

In Abbildung 3 ist der Prozentsatz der erreichten Punktwerte in den Gesundheitsdomänen angegeben. Der Prozentsatz an 0 Punkten, ist in den Gesundheitsdomänen „Kognition“, „Selbstversorgung“ und „mit anderen zurechtkommen“ jeweils über 50 %.

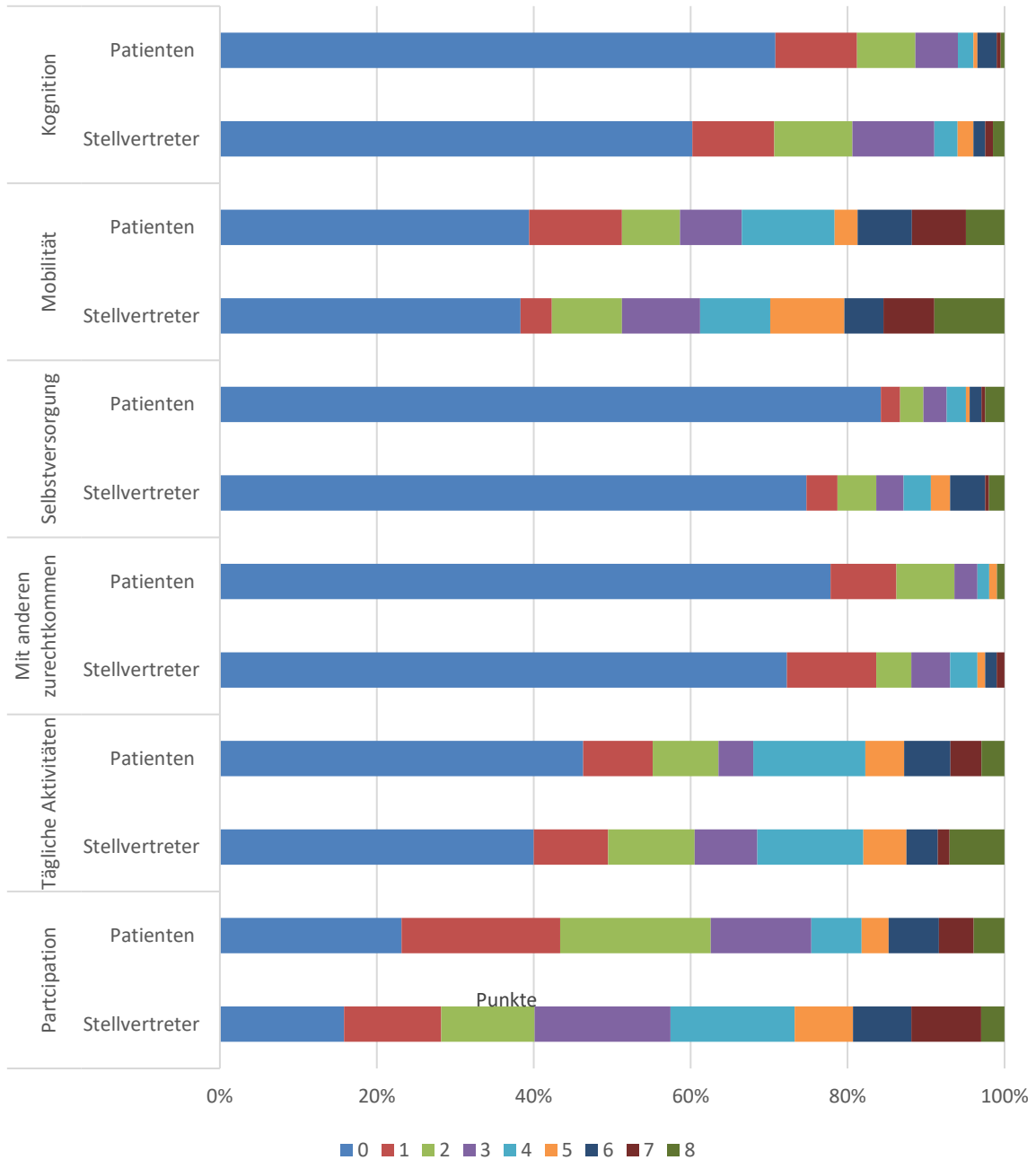


Abbildung 3: Verteilung der Punktwerte in den Gesundheitsdomänen des WHODAS 2.0; n = 204.

Mediane

In der Tabelle 4 sind die Mediane in den Einzelfragen, Gesundheitsdomänen und der Gesamtpunktzahl des WHODAS 2.0 angegeben. Die geringsten Einschränkungen liegen in den Gesundheitsdomänen „Kognition“, „Selbstversorgung“ und „soziale Interaktionen“ vor.

Tabelle 4: Mediane Punktwerte und Differenzen in den Fragen, Gesundheitsdomänen und der Gesamtpunktzahl des WHODAS 2.0

	Mediane Punktzahl der Patient*innen (IQA)	Mediane Punktzahl der Stellvertreter*innen (IQA)	Mediane Differenz (IQA)
Frage 1	0 (0 - 2)	1 (0 - 2)	0 (-1 - 0)
Frage 2	0 (0 - 2)	1 (0 - 2)	0 (-1 - 0)
Frage 3	0 (0 - 0)	0 (0 - 1)	0 (0 - 0)
Frage 4	0 (0 - 1)	0 (0 - 2)	0 (-1 - 0)
Frage 5	1 (0 - 3)	2 (1 - 3)	0 (-1 - 0)
Frage 6	0 (0 - 0)	0 (0 - 1)	0 (-1 - 0)
Frage 7	1 (0 - 3)	1 (0 - 3)	0 (-1 - 0)
Frage 8	0 (0 - 0)	0 (0 - 0)	0 (0 - 0)
Frage 9	0 (0 - 0)	0 (0 - 0)	0 (0 - 0)
Frage 10	0 (0 - 0)	0 (0 - 0)	0 (0 - 0)
Frage 11	0 (0 - 0)	0 (0 - 0)	0 (0 - 0)
Frage 12	0 (0 - 2)	0 (0 - 2)	0 (-1 - 0)
(1) Kognition	0 (0 - 1)	0 (0 - 2)	0 (-1 - 0)
(2) Mobilität	1 (0 - 4)	2 (0 - 5)	0 (-1 - 1)
(3) Selbstversorgung	0 (0 - 0)	0 (0 - 1)	0 (0 - 0)
(4) Mit anderen Menschen zurechtkommen	0 (0 - 0)	0 (0 - 1)	0 (0 - 0)
(5) Tägliche Aktivitäten	1 (0 - 4)	2 (0 - 4)	0 (-1 - 1)
(6) Partizipation	2 (1 - 3)	3 (1 - 5)	-1 (-2 - 1)
Gesamtpunktzahl	13 (4 - 26)	19 (6 - 33)	-2 (-13 - 4)

4.2.2 Prozent gleicher und ähnlicher Werte

Einzelfragen

In der Abbildung 4 ist der Prozentsatz von gleichen bzw. um maximal einer Kategorie abweichenden Antworten in den Fragen dargestellt. Der niedrigste Prozentsatz gleicher Antworten liegt bei der Einschätzung der emotionalen Belastung (Frage 5) vor (40%). Die Einschränkungen in der prämorbidem Lebensqualität der Patient*innen wurden durch die Stellvertreter*innen bei jeder der Fragen häufiger überschätzt als unterschätzt. Die höchsten Prozentwerte gleicher Antworten liegen in den Fragen 3 sowie 8 bis 11 vor.

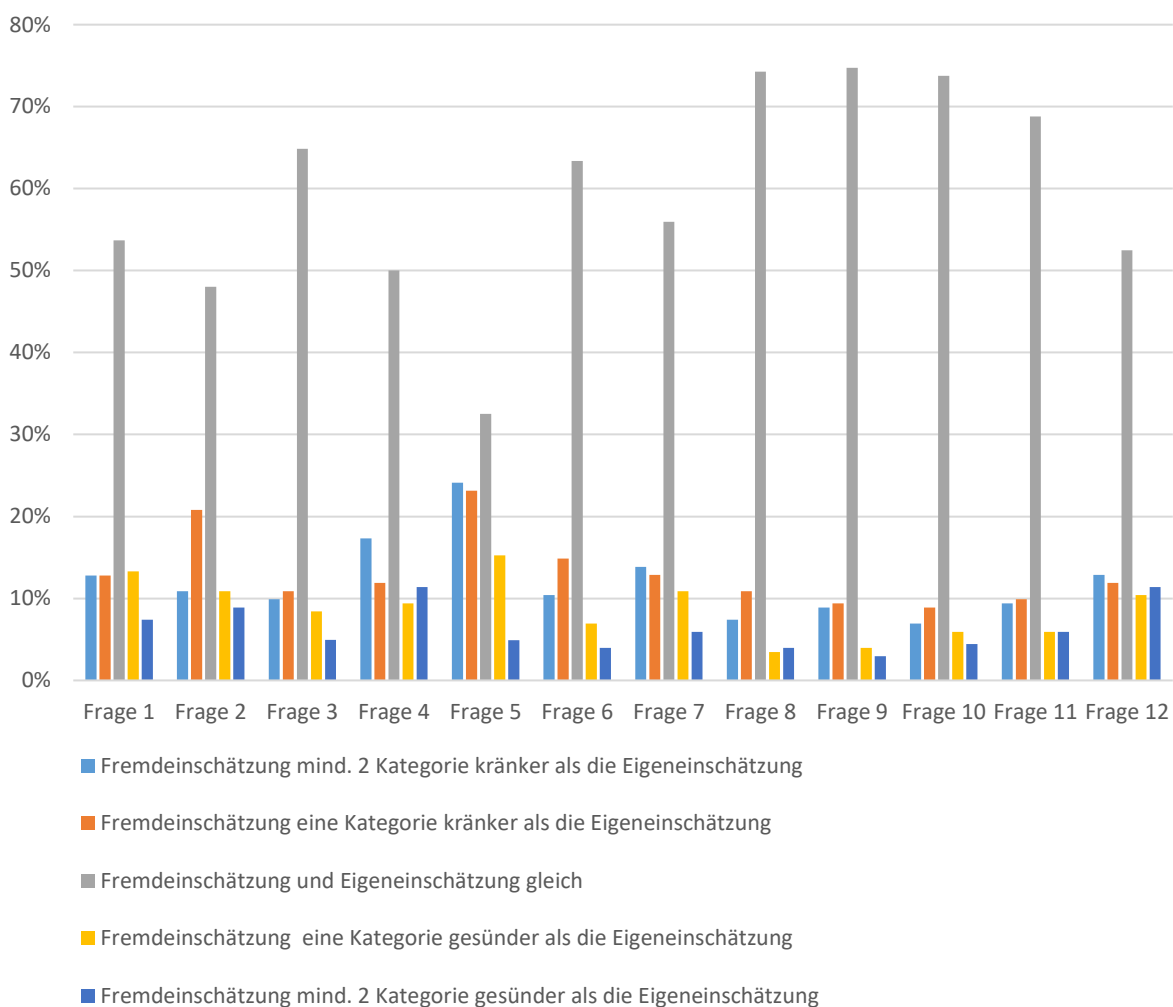


Abbildung 4: Anteil der gleichen, um 1 Kategorie und um mind. 2 Kategorien abweichenden Antworten in den einzelnen Fragen des WHODAS 2.0; n = 204.

Gesundheitsdomänen

In der Abbildung 6 ist der Prozentwert gleicher, um 1-2 Punkte und um maximal 3 Punkte abweichenden Punktwerte in den Gesundheitsdomänen dargestellt. Bei den Gesundheitsdomänen liegen ähnliche Werte wie bei den einzelnen Fragen vor. Der Prozentsatz an ähnlichen Punktwerten (maximal 2 Punkte Abweichung) liegt zwischen 70 % und 90 %. Die Einschränkungen wurden in der Fremdbewertung häufiger überschätzt als unterschätzt. Besonders hohe Prozentwerte gleicher Punktzahlen liegen in den Gesundheitsdomänen „Kognition“ (54%), „Selbstversorgung“ (69%) und „soziale Interaktionen“ (60%) vor.

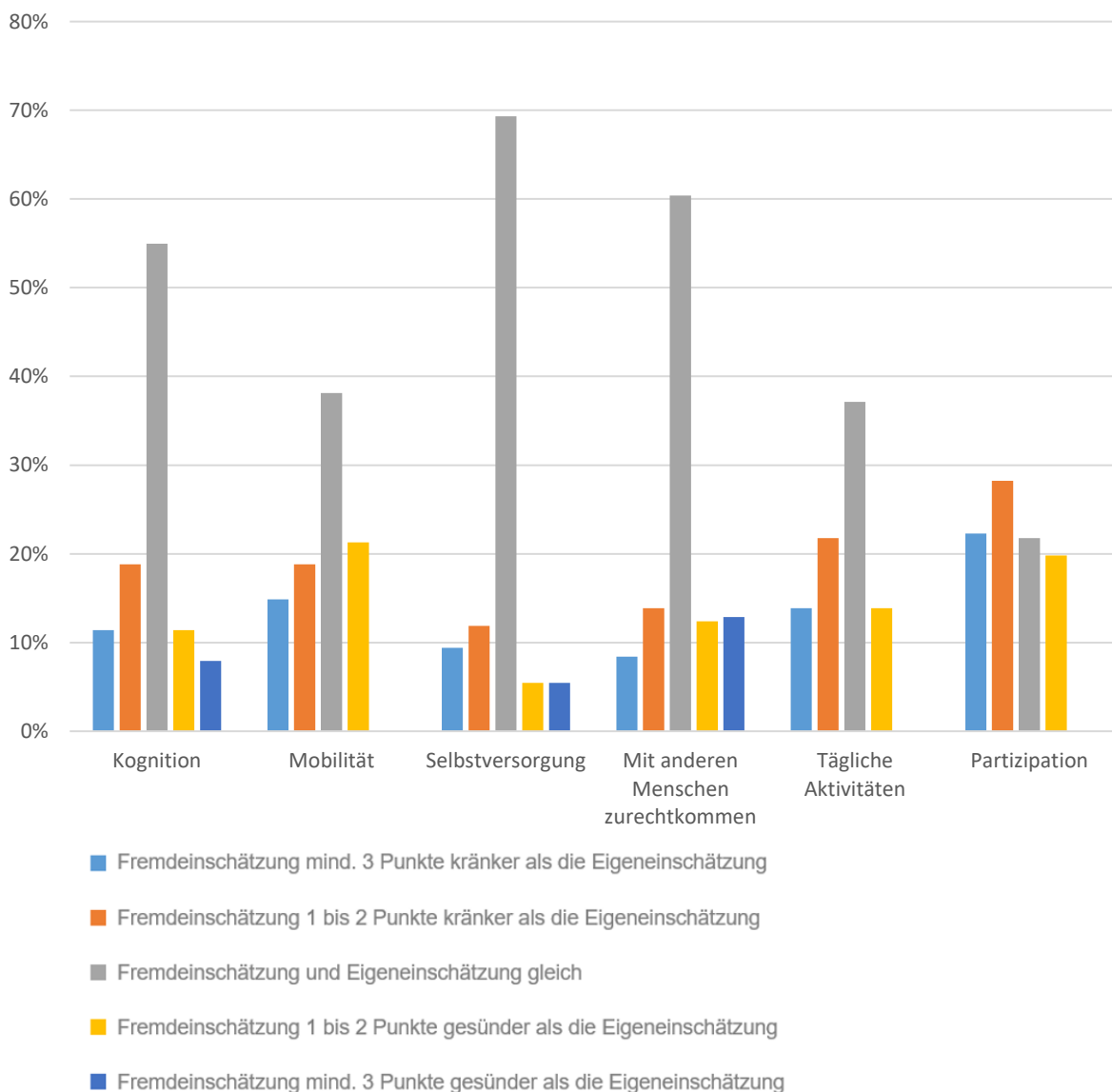


Abbildung 5: Anteil gleichen, um 1 bis 2 und um mind. 3 Punkte abweichender Punktwerte der Patient*innen und deren Stellvertreter*innen in den Gesundheitsdomänen des WHODAS 2.0; n = 204.

Gesamtpunktzahl

In der Abbildung 6 ist der Prozentsatz gleicher sowie um maximal 10 bzw. 25 Punkte abweichender Punktwerte in der Gesamtpunktzahl dargestellt. Wie auch bei den Einzelfragen wurden die Einschränkungen häufiger durch die Stellvertreter*innen überschätzt als unterschätzt.

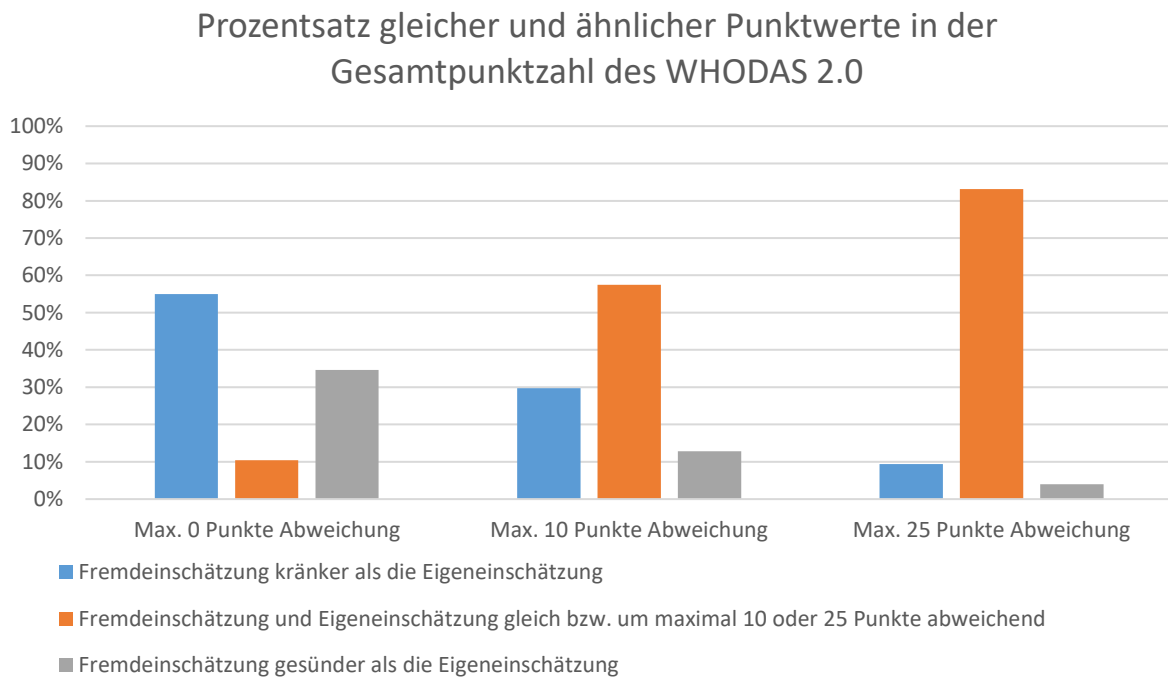


Abbildung 6: Anteil um max. 0, 10 und 25 Punkte abweichender Punktwerte in der Gesamtpunktzahl des WHODAS 2.0; n = 204. Um als kränker oder gesünder gewertet zu werden, ist die Fremdeinschätzung in diesen Kategorien jeweils um 1 bzw. 11 oder 26 Punkte abweichend.

4.2.3 Übereinstimmung in der Gesamtpunktzahl

Limits of Agreement

Die LOA liegen bei -40,54 und 30,94 und sind somit deutlich größer als die MCID von 10 Punkte (vgl. Abbildung 7).

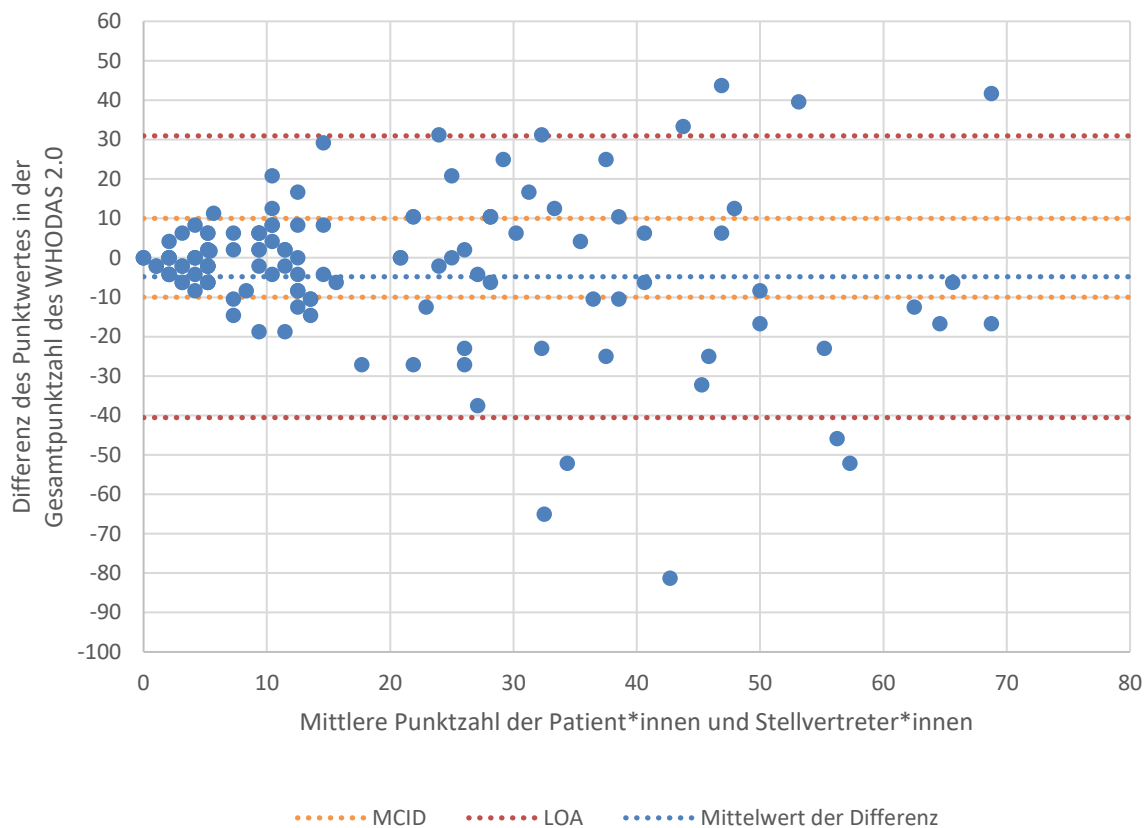


Abbildung 7: Bland-Altman-Diagramm für die Übereinstimmung zwischen den Patient*innen und derer Stellvertreter*innen in der Gesamtpunktzahl des WHODAS 2.0; $n = 204$. LOA = Limit of Agreement, MCID = Minimal clinically important difference. Die Differenz wurde als Punktwert der Patient*innen minus Punktwert der Stellvertreter*innen berechnet, sodass eine negative Differenz eine Überschätzung der Einschränkungen bedeutet und umgekehrt.

ICC

Die ICC für die Übereinstimmung in der Gesamtpunktzahl des WHODAS 2.0 beträgt 0,724 (95 % KI: 0,627 bis 0,794). Hiernach ist die Übereinstimmung als moderat bis gut zu Werten.

4.2.4 Übereinstimmung in den Einzelfragen und Gesundheitsdomänen

Kappa-Werte

In den Abbildungen 8 und 9 sind die gewichteten Kappa-Werte der einzelnen Fragen und den Gesundheitsdomänen dargestellt. Nur in Frage 7 liegt eine, nach den in dieser Dissertation festgelegten Kriterien (siehe Abschnitt 3.6), ausreichende Übereinstimmung vor. Die höchsten gewichteten Kappa-Werte kommen in den Fragen 1 und 7 vor, welche beide zur Gesundheitsdomäne „Mobilität“ gehören. Bei Frage 10 und 11 wurden die geringsten gewichteten Kappa-Werte erreicht. Nur in der Gesundheitsdomäne „Mobilität“ liegt anhand der Kappa-Werte eine ausreichende Übereinstimmung ($K_w > 0,4$) vor.

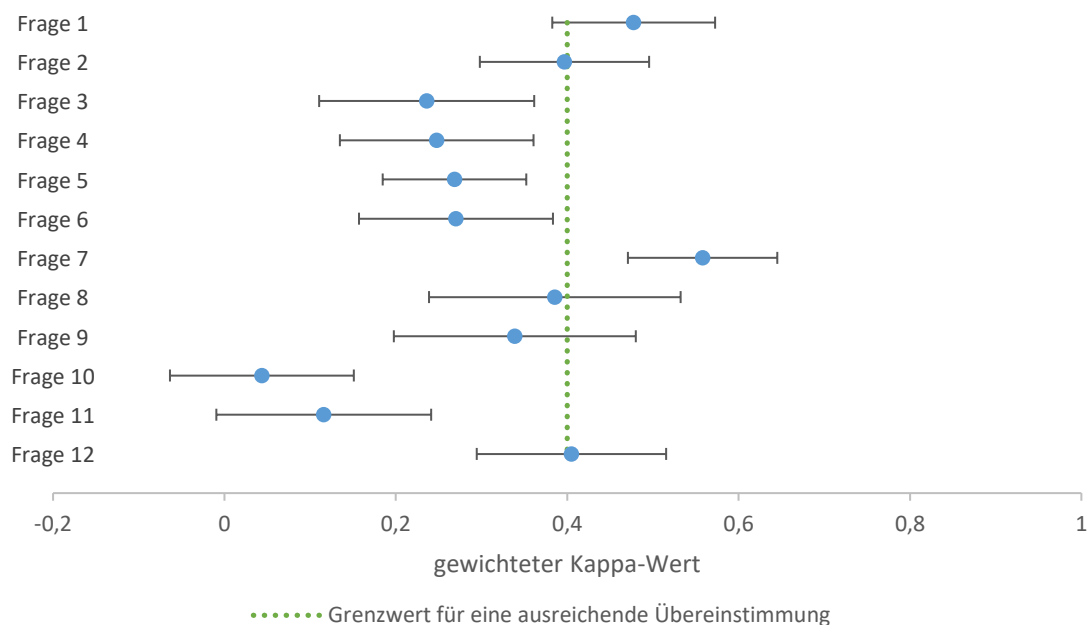
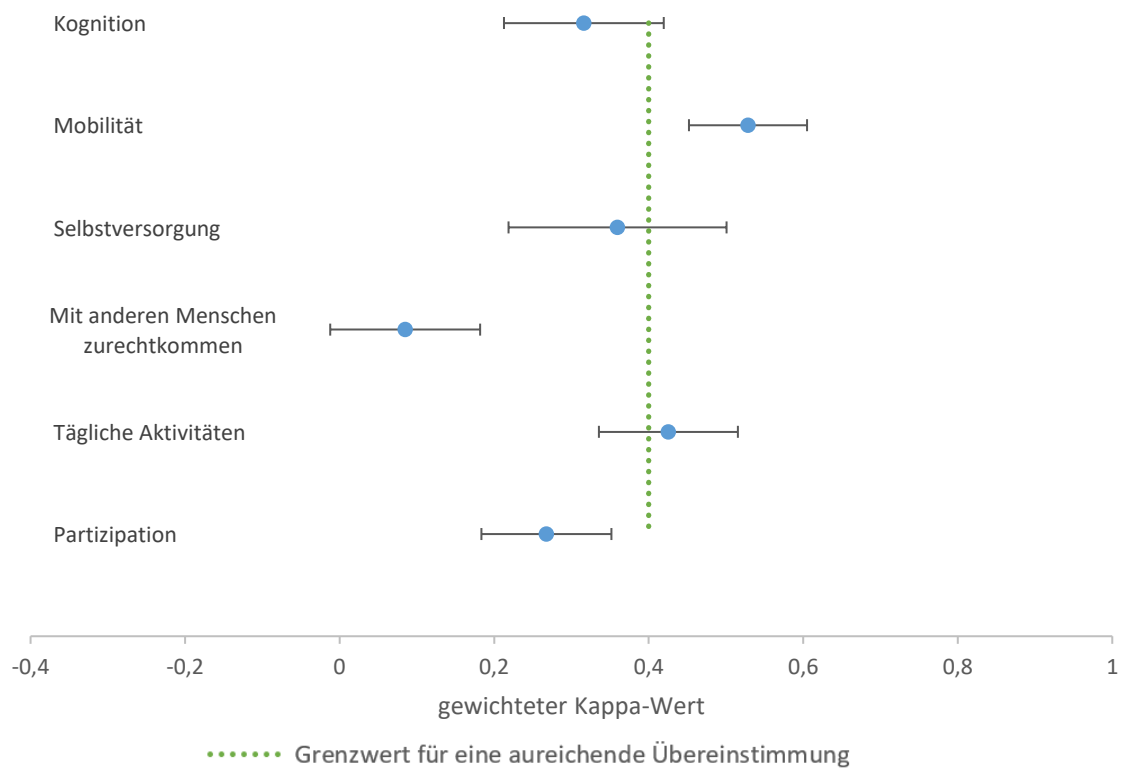


Abbildung 8: Gewichtete Kappa-Werte in den Fragen 1 bis 12 Fragen des WHODAS 2.0; $n = 204$. Die gewichteten Kappa-Werte sind mit ihrem 95 % Konfidenzintervall (Fehlerindikatoren) dargestellt.



z

Abbildung 9: Gewichteten Kappa-Werte in den Gesundheitsdomänen des WHODAS 2.0; n = 204. Die gewichteten Kappa-Werte sind mit ihren 95% Konfidenzintervall (Fehlerindikatoren) dargestellt.

PABAK-Werte

In den Abbildungen 10 und 11 sind die PABAK-Werte der einzelnen Fragen und Gesundheitsdomänen dargestellt. Bei allen Fragen, außer den Fragen 4 und 5, liegt der PABAK-Wert über 0,4 und ist somit als ausreichend zu werten. Bei den Fragen 8 bis 10 liegt eine gute Übereinstimmung ($\text{PABAK} > 0,6$) vor. Außer in der Gesundheitsdomäne Partizipation, welche aus den Fragen 4 und 5 gebildet wird, liegt in allen restlichen Gesundheitsdomänen eine ausreichende ($\text{PABAK} > 0,4$) vor. In 3 der 6 Gesundheitsdomänen ist die Übereinstimmung anhand des PABAK-Wertes als „gut“ zu werten.

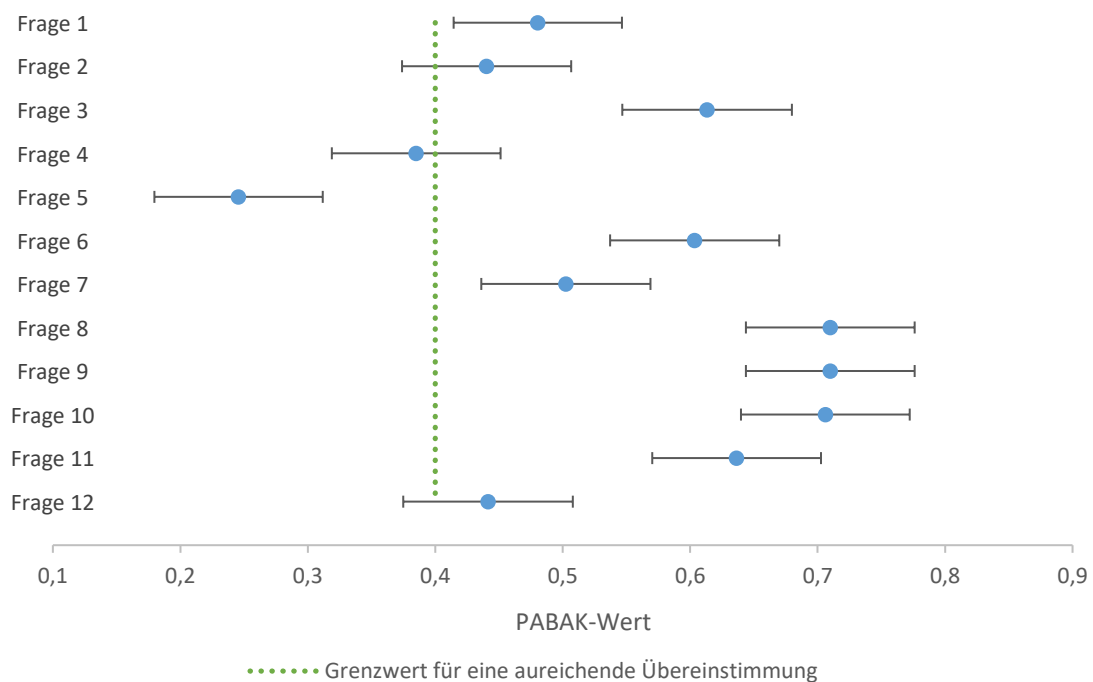


Abbildung 10: PABAK-Werte in den Fragen 1 bis 12 des WHODAS 2.0; $n = 204$. Die PABAK-Werte sind mit ihrem 95% Konfidenzintervall (Fehlerindikatoren) dargestellt.

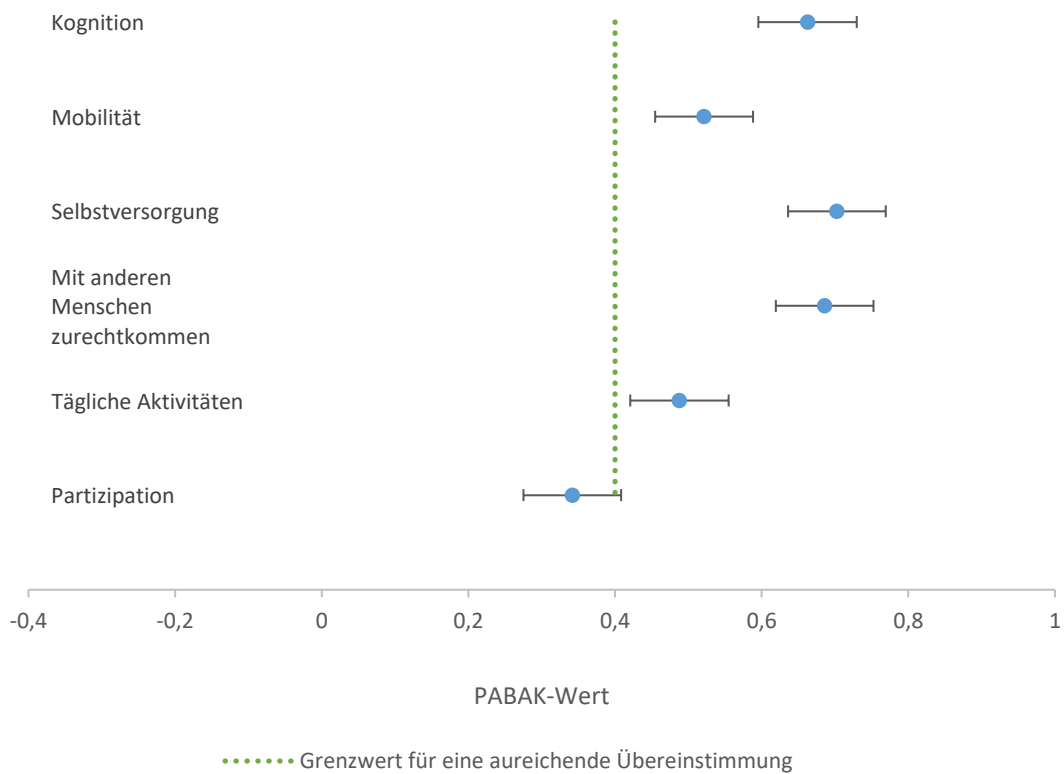


Abbildung 11: PABAK-Werte in den Gesundheitsdomänen des WHODAS 2.0; n = 204. Die PABAK-Werte sind mit ihrem 95% Konfidenzintervall (Fehlerindikatoren) dargestellt.

4.2.5 Zusammenhang zwischen dem Einschränkungsggrad der Patient*innen und der Übereinstimmung

Gewichtete Kappa-Werte

In der Abbildung 13 ist die Höhe der gewichteten Kappa-Werte im Zusammenhang zu dem Prozentsatz in Patient*innen ohne Einschränkungen in den entsprechenden Fragen und Gesundheitsdomänen dargestellt. Das Bestimmtheitsmaß (R^2) beträgt 0,202 ($p = 0,06$). Wenn die Frage 5 und die Gesundheitsdomäne Partizipation nicht mitbetrachtet werden, beträgt das Bestimmtheitsmaß (R^2) 0,538 ($p < 0,001$). Dies bedeutet, dass bei den Kategorien in denen mehr Patient*innen „keine Einschränkungen“ hatten, anhand der Kappa-Werte eine schlechtere Übereinstimmung vorlag.

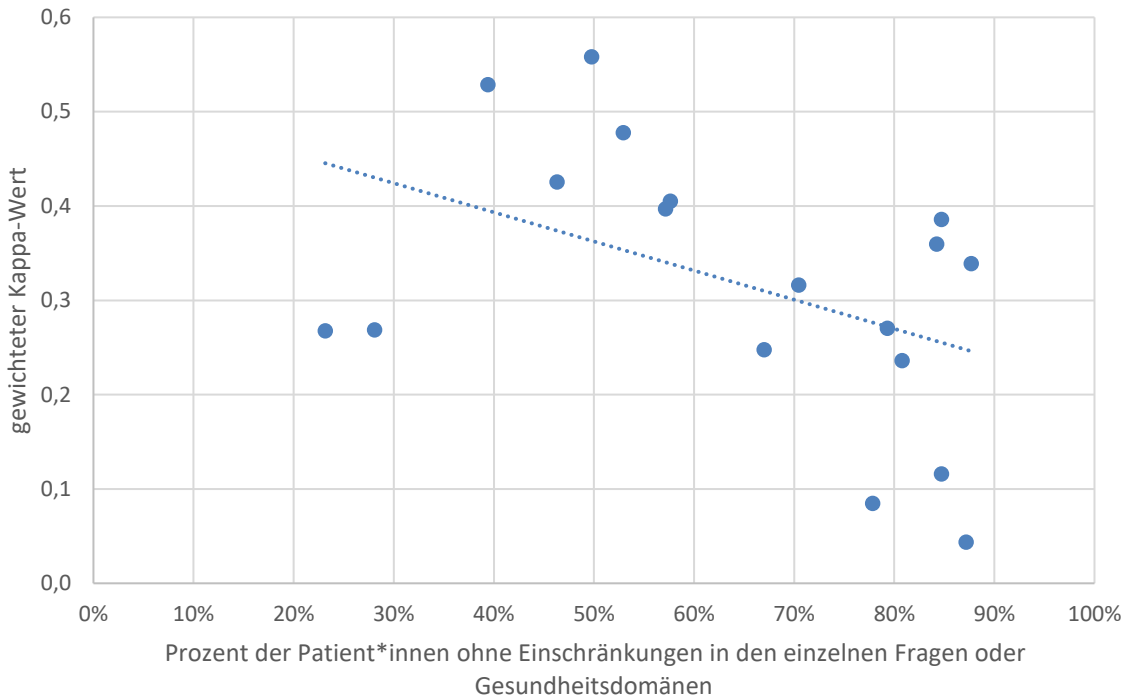


Abbildung 12: Zusammenhang zwischen dem Prozentsatz an Patient*innen die angegeben haben „keine Einschränkungen“ zu haben und den gewichteten Kappa-Werten. Die Punkte entsprechen den einzelnen Fragen und Gesundheitsdomänen.

PABAK-Werte

In Abbildung 14 ist die Höhe der gewichteten PABAK-Werte im Zusammenhang zu dem Prozentsatz der Patient*innen ohne Einschränkungen in den entsprechenden Fragen und Gesundheitsdomänen dargestellt. Das Bestimmtheitsmaß (R^2) beträgt 0,887 ($p < 0,001$). Dies bedeutet, dass bei den Kategorien in denen mehr Patient*innen „keine Einschränkungen“ hatten, anhand der PABAK-Werte eine signifikant bessere Übereinstimmung vorlag.

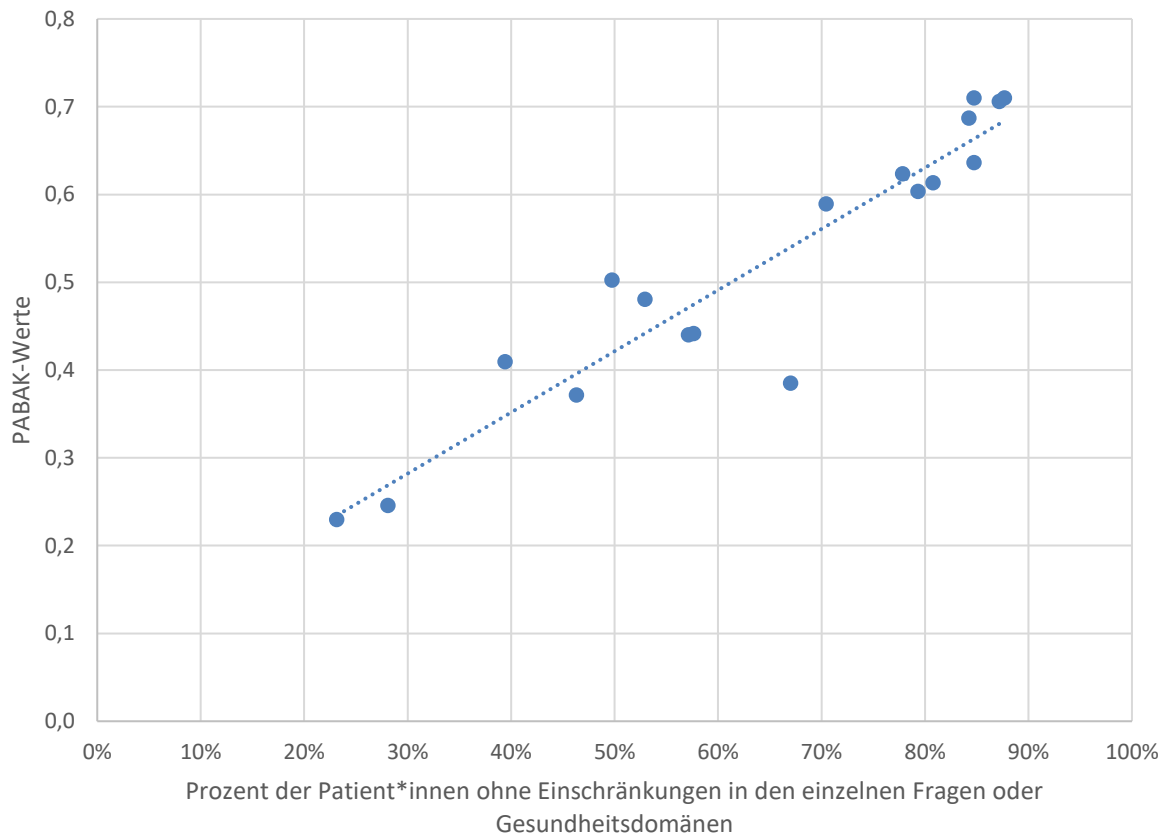


Abbildung 13: Zusammenhang zwischen dem Prozentsatz an Patient*innen die angegeben haben „keine Einschränkungen“ zu haben und den PABAK-Werten. Die Punkte entsprechen den einzelnen Fragen und Gesundheitsdomänen.

4.3 Ergebnisse zur Fragestellung 2

4.3.1 Populationscharakteristika

In der folgenden Tabelle sind die Gruppencharakteristika beider Populationen dargestellt. Patient*innen welche intensivmedizinisch behandelt wurden, waren signifikant älter als Patient*innen, die nicht intensivmedizinisch behandelt wurden ($p = 0.002$). Zwischen den Gruppen liegen keine signifikanten Unterschiede im Alter der Stellvertreter*innen oder im Geschlecht vor.

Tabelle 5: Populationscharakteristika der postoperativ auf Normal- und Intensivstation behandelten Patient*innen und deren Stellvertreter*innen

		Normalstation	Intensivstation	p-Wert
Anzahl		102	102	
Alter der Patient*innen (MW, SD)		63,7 (12,9)	69,4 (13,5)	0,002
Alter der Stellvertreter*innen (MW, SD)		60,4 (14)	63 (13,4)	0,181
Geschlecht der Patient*innen (N, %)	männlich	60 (58,8%)	61 (59,8%)	0,887
	weiblich	42 (41,2%)	41 (40,2%)	
Geschlecht der Stellvertreter*innen	männlich	29 (28,4%)	36 (35,3%)	0,315
	weiblich	72 (70,6%)	66 (64,7%)	
Befragungsort der Stellvertreter*innen	Normalstation	102 (100%)	64 (62,7%)	
	Intensivstation	0 (0%)	38 (37,3%)	

Patient*innen, die auf einer Intensivstation behandelt wurden, haben insgesamt mehr Einschränkungen angegeben ($p < 0,001$) (Vgl. Tabelle 6). Der größte Unterschied zwischen den beiden Gruppen liegt in der Gesundheitsdomäne „Mobilität“, wobei Patient*innen, die intensivmedizinisch behandelt wurden, größeren Einschränkungen angegeben haben ($p < 0,001$). In den anderen Gesundheitsdomänen lagen keine signifikanten Unterschiede vor. Eine ähnliche Tendenz zeigt sich auch bei der Fremdeinschätzung der Gesundheit durch die Stellvertreter*innen (vgl. Tabelle 6).

Tabelle 6: Mediane Punktwerte der Patient*innen und derer Stellvertreter*innen in den Gesundheitsdomänen und der Gesamtpunktzahl des WHODAS 2.0

		Mediane Punktzahl der Patient*innen (IQA)	-Wert*	Mediane Punktzahl der Stellvertreter*innen (IQA)	p-Wert*
Kognition	Normalstation	0 (0 - 1)	0,753	0 (0 - 1,8)	0,075
	Intensivstation	0 (0 - 1)		0 (0 - 2)	
Mobilität	Normalstation	0 (0 - 3)	<0,001	0 (0 - 3,8)	<0,001
	Intensivstation	3 (0,3 - 5)		4 (2 - 6)	
Selbstversorgung	Normalstation	0 (0 - 0)	0,743	0 (0 - 0)	0,051
	Intensivstation	0 (0 - 0)		0 (0 - 2)	
Mit anderen Menschen	Normalstation	0 (0 - 0)	0,618	0 (0 - 0)	0,01
	Intensivstation	0 (0 - 0)		0 (0 - 1)	
Tägliche	Normalstation	0 (0 - 4)	0,201	0,5 (0 - 3,7)	0,006
	Intensivstation	1 (0 - 4)		2 (0 - 4)	
Partizipation	Normalstation	2 (0 - 3)	0,192	3 (1 - 4)	0,012
	Intensivstation	2 (1 - 4)		3 (2 - 5)	
Gesamtpunktzahl	Normalstation	8,3 (2,1 - 25)	0,001	8,3 (4,2 - 27,1)	<0,001
	Intensivstation	16,7 (8,3 - 28,6)		22,9 (10,4 - 37,5)	

* Die Unterschiede im Median wurden mittels Mann-Whitney-U-Test auf Signifikanz

4.3.2 Systematischer Bias

In der Tabelle 7 ist die mediane Differenz der Punktzahlen der Patient*innen und derer Stellvertreter*innen in den Gesundheitsdomänen und der Gesamtpunktzahl des WHODAS 2.0 dargestellt. Zwischen den beiden Gruppen liegen im Mann-Whitney-U-Test signifikante Unterschiede bei den Gesundheitsdomänen „Kognition“, „mit anderen Menschen zurechtkommen“ und „tägliche Aktivitäten“ sowie der Gesamtpunktzahl vor. Das Ausmaß dieser Unterschätzung ist allerdings kleiner als die festgelegte MCID von 2 Punkten in den Gesundheitsdomänen und 10 Punkten in der Gesamtpunktzahl.

Tabelle 7: Vergleich der medianen Differenz der Punktwerten der Patient*innen und deren Stellvertreter*innen in den Gesundheitsdomänen und der Gesamtpunktzahl des WHODAS 2.0

		Median der gepaarten Differenz (IQA)	p-Wert*
Kognition	Normalstation	0 (0 - 0)	0,044
	Intensivstation	0 (-1 - 0)	
Mobilität	Normalstation	0 (0 - 1)	0,177
	Intensivstation	0 (-2 - 1)	
Selbstversorgung	Normalstation	0 (0 - 0)	0,092
	Intensivstation	0 (-1 - 0)	
Mit anderen Menschen zurechtkommen	Normalstation	0 (0 - 0)	0,008
	Intensivstation	0 (-1 - 0)	
Tägliche Aktivitäten	Normalstation	0 (-1 - 1)	0,041
	Intensivstation	0 (-2 - 0)	
Partizipation	Normalstation	0 (-2 - 1)	0,155
	Intensivstation	-1 (-3 - 1)	
Gesamtpunktzahl	Normalstation	0 (-6 - 6)	0,009
	Intensivstation	-6 (-17 - 4)	

* Die Unterschiede im Median wurden mittels Mann-Whitney-U-Test auf signifikant getestet. Die Differenz wurde als Punktwert der Patient*innen minus Punktwert der Stellvertreter*innen berechnet.

4.3.3 Vergleich der Übereinstimmung in der Gesamtpunktzahl

LOA

Der Mittelwert der gepaarten Differenz der Gesamtpunktzahl ist in der Gruppe der intensivmedizinisch behandelten Patient*innen negativer als bei nicht intensivmedizinisch behandelten Patient*innen (-8,3 vs. -1,3) (Vgl. Abbildung 14 und 15). Bei beiden liegt anhand der LOA keine ausreichende Übereinstimmung vor, da bei beiden Gruppen die LOA außerhalb der MCID liegen. Die Standarddeviation der Differenz in der Gruppe 1 beträgt 17,7 und in der Gruppe 2, 17,9. Somit umfassen die LOA in beiden Gruppen einen ähnlich großen Bereich, der allerdings in der Gruppe der intensivmedizinisch behandelten Patient*innen um 7 Punkte nach unten verschoben ist (Vgl. Abbildung 14 und 15).

ICC

Die Intraklassenkorrelation beträgt bei intensivmedizinisch behandelten Patient*innen und deren Stellvertreter*innen 0,732 (95%KI: 0,603 bis 0,819) und bei nicht-intensivmedizinisch behandelten Patient*innen und deren Stellvertreter*innen 0,684 (95%KI: 0,531 bis 0,787).

Hiernach ist die Übereinstimmung in beiden Gruppen als moderat bis gut zu werten und der Unterschied, nach den in dieser Arbeit festgelegten Kriterien, nicht relevant, da dieser kleiner als 0,25 ist (Siehe Abschnitt 3.6).

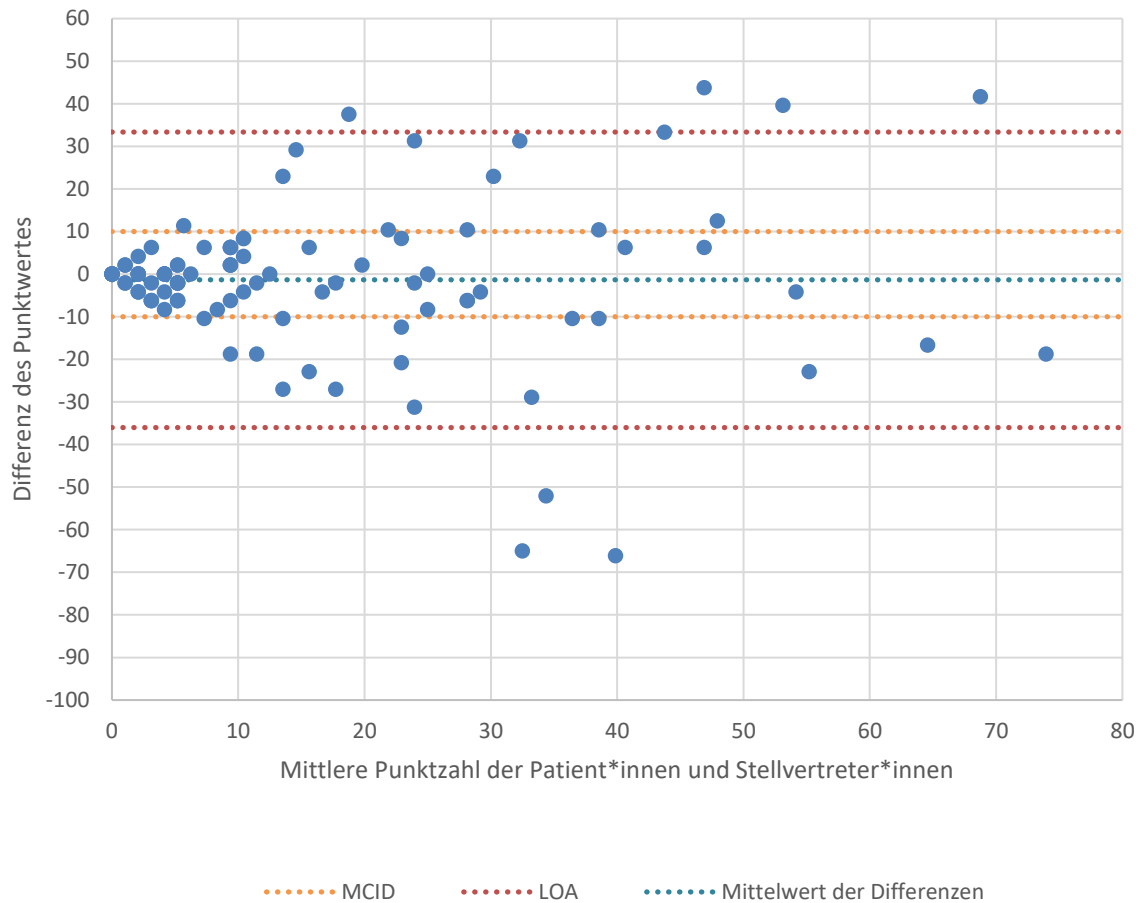


Abbildung 14: Bland-Altman-Diagramm für die Übereinstimmung in der Gesamtpunktzahl des WHODAS 2.0, zwischen den Patient*innen, die postoperativ nicht intensivmedizinisch behandelt wurden und deren Stellvertretern; n =204. LOA = Limit of Agreement, MCID = Minimal clinically important difference. Die Differenz wurde als Punktwert der Patient*innen minus Punktwert der Stellvertreter*innen berechnet, sodass eine negative Differenz eine Überschätzung der Einschränkungen bedeutet und umgekehrt.

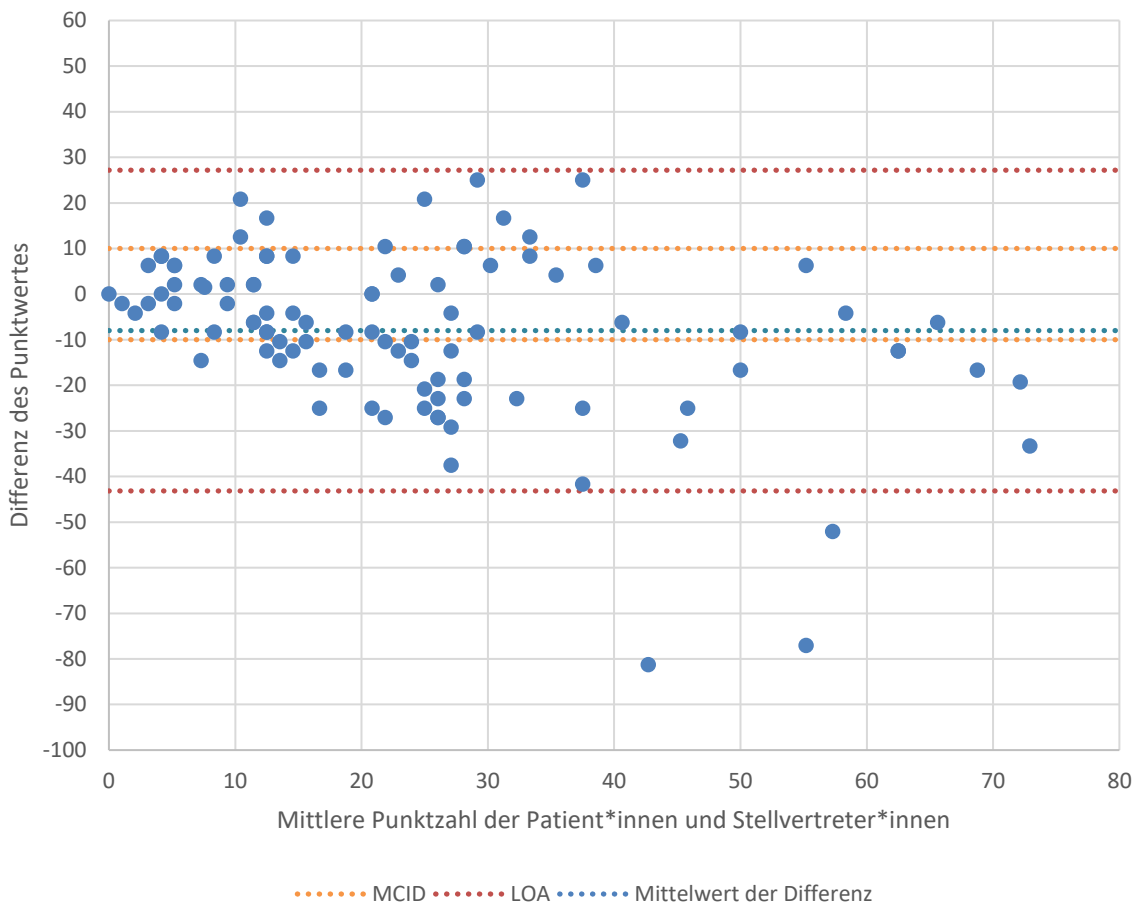


Abbildung 15: Bland-Altman-Diagramm für die Übereinstimmung in der Gesamtpunktzahl des WHODAS 2.0, zwischen den Patient*innen, die postoperativ intensivmedizinisch behandelt wurden und deren Stellvertretern; $n = 204$. LOA = Limit of Agreement, MCID = Minimal clinically important difference. Die Differenz wurde als Punktwert der Patient*innen minus Punktwert der Stellvertreter*innen berechnet, sodass eine negative Differenz eine Überschätzung der Einschränkungen bedeutet und umgekehrt.

4.3.4 Vergleich der Übereinstimmung in den Gesundheitsdomänen

Gewichtete Kappa-Werte

In der Abbildung 16 sind die gewichteten Kappa-Werte für beide Gruppen in den Gesundheitsdomänen dargestellt. Bei keiner der Gruppen oder Gesundheitsdomänen liegt der gewichtete Kappa-Wert über 0,4. Die größten Unterschiede in den gewichteten Kappa-Werten liegen in den Gesundheitsdomänen „Mobilität“ und „Mit anderen Menschen zurechtkommen“, mit einer Differenz von jeweils 0,18 und 0,23.

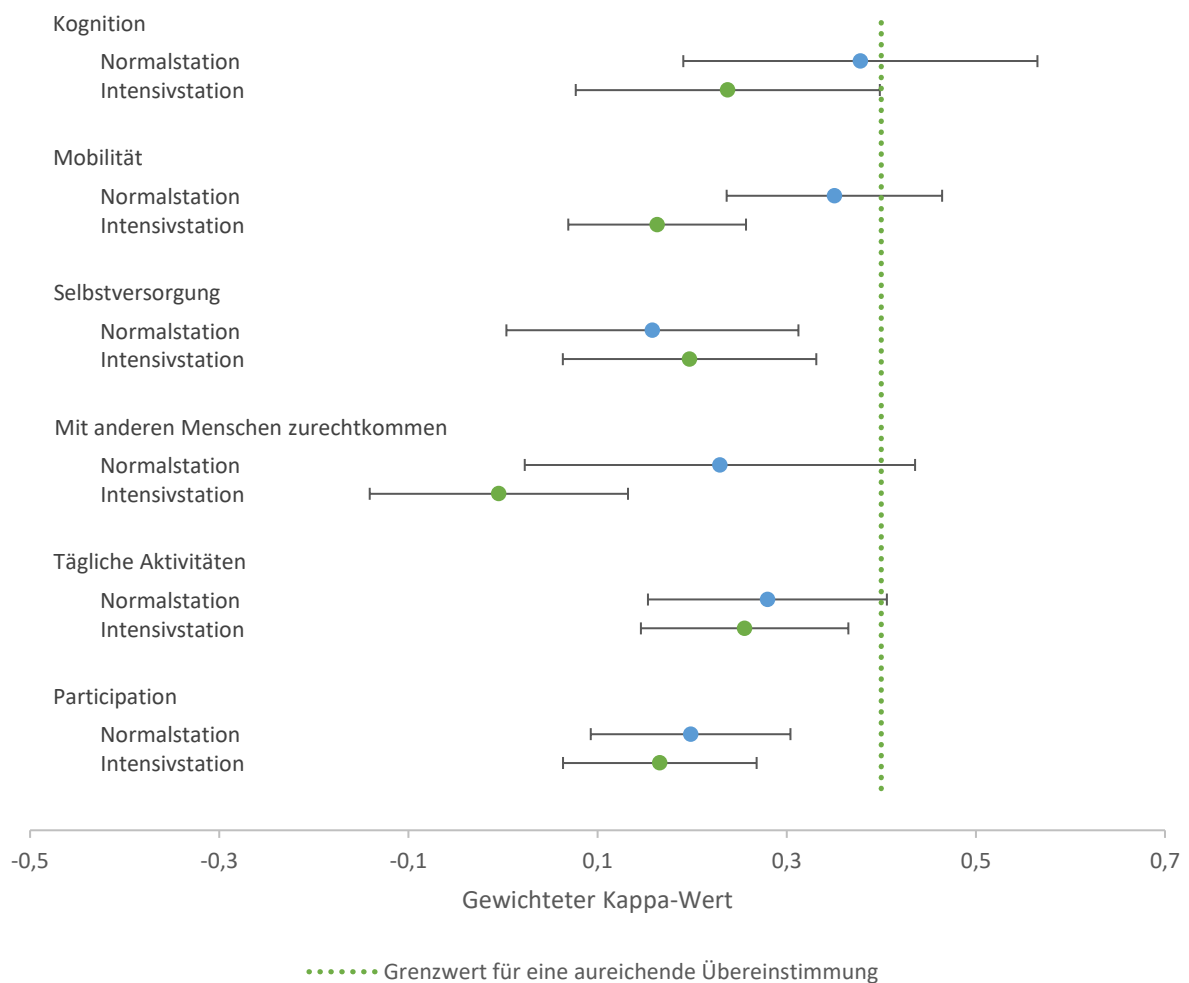


Abbildung 16: Vergleich der gewichteten Kappa-Werte in den Gesundheitsdomänen des WHODAS 2.0 zwischen den verschiedenen Gruppen. Die gewichtete Kappa-Werte sind mit ihrem 95% Konfidenzintervall (Fehlerindikatoren) dargestellt.

PABAK-Werte

In der folgenden Grafik sind die PABAK-Werte für beide Gruppen in den Gesundheitsdomänen dargestellt. In der Gruppe der Patient*innen, die nicht intensivmedizinisch behandelt wurden, sind die PABAK-Werte bei allen Gesundheitsdomänen höher als bei Patient*innen, die intensivmedizinisch behandelt wurden. Der größte Unterschied zwischen den Gruppen liegt in der Gesundheitsdomäne „Mobilität“ vor, mit einer Differenz von 0,23. Bei allen anderen Gesundheitsdomänen liegt der Unterschied unter 0,2.

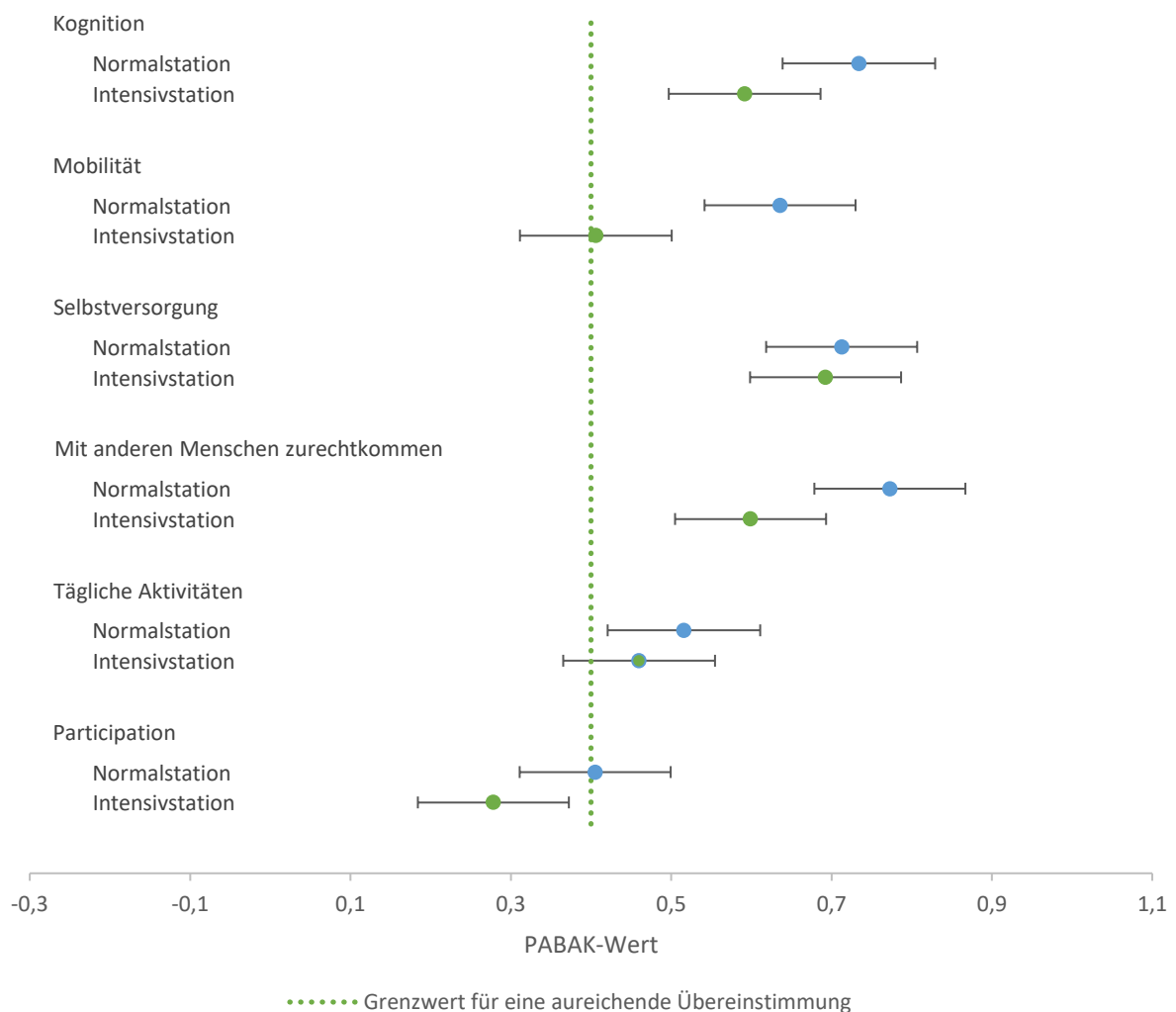


Abbildung 17: Vergleich der PABAK-Werte in den Gesundheitsdomänen des WHODAS 2.0 zwischen den verschiedenen Gruppen. Die PABAK -Werte sind mit ihren 95% Konfidenzintervall (Fehlerindikatoren) dargestellt.

4.3.5 Fallkontrollabgleich zwischen Normalstation und Intensivstation

Gruppencharakteristika

In Tabelle 8 sind die Gruppencharakteristika der beiden Gruppen des Fallkontrollabgleiches dargestellt. Dieser wurde, wie in Abschnitt 3.6 beschrieben, anhand der Gesamtpunktzahl der Patient*innen im WHODAS 2.0 durchgeführt.

Tabelle 8: Gruppencharakteristika der beiden Gruppen im Fallkontrollabgleich

	Normalstation	Intensivstation	p-Wert*
Anzahl	64	64	
Alter der Patient*innen (MW, SD)	63,5 (13)	69,2 (14,7)	0,021
Alter der Stellvertreter*innen (MW, SD)	58,9 (13,7)	64,9 (12,9)	0,012
Geschlecht der Patient*innen (N, %)	männlich 34 (53,1%) weiblich 30 (46,9%)	37 (57,8%) 27 (42,2%)	0,594
Geschlecht der Stellvertreter*innen (N, %)	männlich 21 (32,8%) weiblich 43 (67,2%)	24 (37,5%) 40 (62,5%)	0,315

MW = Mittelwert, SD = Standarddeviation *Die Signifikanz des Unterschiedes im Alter wurden mittels T-Test für unverbundene Stichproben berechnet. Die Signifikanz des Unterschiedes der Geschlechtsverteilung wurde mittels Chi-Quadrat-Test berechnet.

Mediane Gesamtpunktzahl der Patient*innen

Zwischen den beiden Gruppen besteht in keiner der Gesundheitsdomänen ein signifikanter Unterschied (Vgl. Tabelle 9).

Systematischer Bias

Im Median haben die Stellvertreter*innen von Patient*innen, die intensivmedizinisch behandelt wurden, die Einschränkungen der Patient*innen um 4 Punkte überschätzt, im Vergleich zu 0 Punkten bei Patient*innen, die nicht intensivmedizinisch behandelt wurden ($p = 0,093$). Bei der Differenz in den Gesundheitsdomänen liegen ebenfalls keine signifikanten Unterschiede vor (vgl. Tabelle 10).

Tabelle 9: Mediane Punktzahl der Patient*innen in der Gesamtpunktzahl des WHODAS 2.0 bei den verschiedenen Gruppen beim Fallkontrollabgleich zur Fragestellung 2

		Mediane Punktzahl der Patient*innen (IQA)	p – Wert*
Kognition	Normalstation	0 (0 - 1)	0,572
	Intensivstation	0 (0 - 1)	
Mobilität	Normalstation	1 (0 - 4)	0,086
	Intensivstation	2 (0 - 4)	
Selbstversorgung	Normalstation	0 (0 - 0)	0,436
	Intensivstation	0 (0 - 0)	
Mit anderen Menschen zurechtkommen	Normalstation	0 (0 - 1)	0,300
	Intensivstation	0 (0 - 0)	
Tägliche Aktivitäten	Normalstation	2 (0 - 4)	0,192
	Intensivstation	0 (0 - 4)	
Partizipation	Normalstation	2 (1 - 4)	0,851
	Intensivstation	2 (1 - 4)	
Gesamtpunktzahl	Normalstation	12,5 (6,7 - 27,1)	0,894
	Intensivstation	12,5 (7,8 - 27,1)	

*Die Unterschiede im Median wurden mittels Mann-Whitney-U-Test auf Signifikanz getestet.

Tabelle 10: Mediane Differenzen zwischen den Patient*innen und deren Stellvertreter*innen in den Gesundheitsdomänen und der Gesamtpunktzahl des WHODAS 2.0 beim Fallkontrollabgleich zur Fragestellung 2

		Median der gepaarten Differenz (IQA)	p – Wert*
Kognition	Normalstation	0 (-0,3 - 0)	0,423
	Intensivstation	0 (-1 - 0)	
Mobilität	Normalstation	0 (-1 - 1)	0,513
	Intensivstation	0 (-2 - 1)	
Selbstversorgung	Normalstation	0 (0 - 0)	0,311
	Intensivstation	0 (-1 - 0)	
Mit anderen Menschen zurechtkommen	Normalstation	0 (0 - 0)	0,106
	Intensivstation	0 (-1 - 0)	
Tägliche Aktivitäten	Normalstation	0 (-1 - 2)	0,113
	Intensivstation	0 (-1 - 0,5)	
Partizipation	Normalstation	0 (-2 - 1)	0,124
	Intensivstation	0 (-3 - 1)	
Gesamtpunktzahl	Normalstation	0 (-8 - 6)	0,093
	Intensivstation	-4 (-13 - 4)	

* Die Unterschiede im Median wurden mittels Mann-Whitney-U-Test auf Signifikanz getestet. Die Differenz wurde als Punktwert der Patient*innen minus Punktwert der Stellvertreter*innen berechnet.

LOA

Bei der Gruppe 1 (Normalstation) liegt der Mittelwert der Differenz zwischen den Patient*innen und deren Stellvertreter*innen in der Gesamtpunktzahl des WHODAS 2.0 bei -1,24 Punkten, die untere LOA bei -36,01 und die obere LOA bei 33,53. Bei der Gruppe 2 (Intensivstation) liegt der Mittelwert der Differenz bei -6,67 Punkten, die untere LOA bei -41,95 und die obere LOA bei 28,62. Die Standarddeviation der Differenz beträgt in den Gruppen 17,7 (Gruppe 1) und 18,0 (Gruppe 2). Die LOA liegen somit bei beiden Gruppen außerhalb der MCID von 10 Punkten.

ICC

Die ICC bei der Gesamtpunktzahl des WHODAS 2.0 liegt in der Gruppe 1 (Normalstation) bei 0,693 (95 % KI 0,494 bis 0,813) und in der Gruppe 2 (Intensivstation) bei 0,710 (95 % KI 0,507 bis 0,827). Der Unterschied in der ICC beträgt somit 0,017 zwischen den Gruppen.

Gewichtete Kappa-Werte

Auch im Fallkontrollabgleich liegt keiner der gewichteten Kappa-Werte über 0,4. Die größte Differenz liegt in der Gesundheitsdomäne „Mit anderen Menschen zurechtkommen“ mit einem Unterschied von 0,17 vor (Vgl. Abbildung 18).

PABAK-Werte

In dem Fallkontrollabgleich sind die PABAK-Werte in 4 der 6 Gesundheitsdomänen in der Gruppe 1 (Normalstation) größer als in der Gruppe 2 (Intensivstation). Der größte Unterschied liegt in der Gesundheitsdomäne „Mit anderen Menschen zurechtkommen“, mit einem Unterschied von 0,149 vor (Vgl. Abbildung 19).

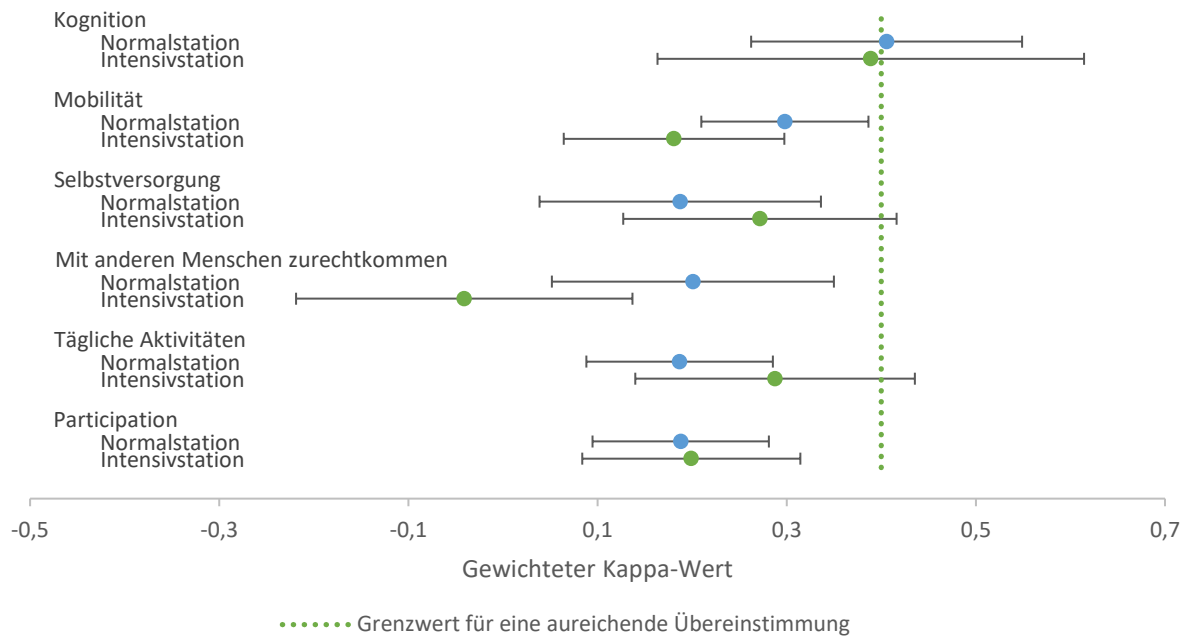


Abbildung 18: Vergleich der gewichteten Kappa-Werte in den Gesundheitsdomänen des WHODAS 2.0 zwischen den verschiedenen Gruppen im Fallkontrollabgleich. Die gewichteten Kappa-Werte sind mit ihrem 95% Konfidenzintervall (Fehlerindikatoren) dargestellt.

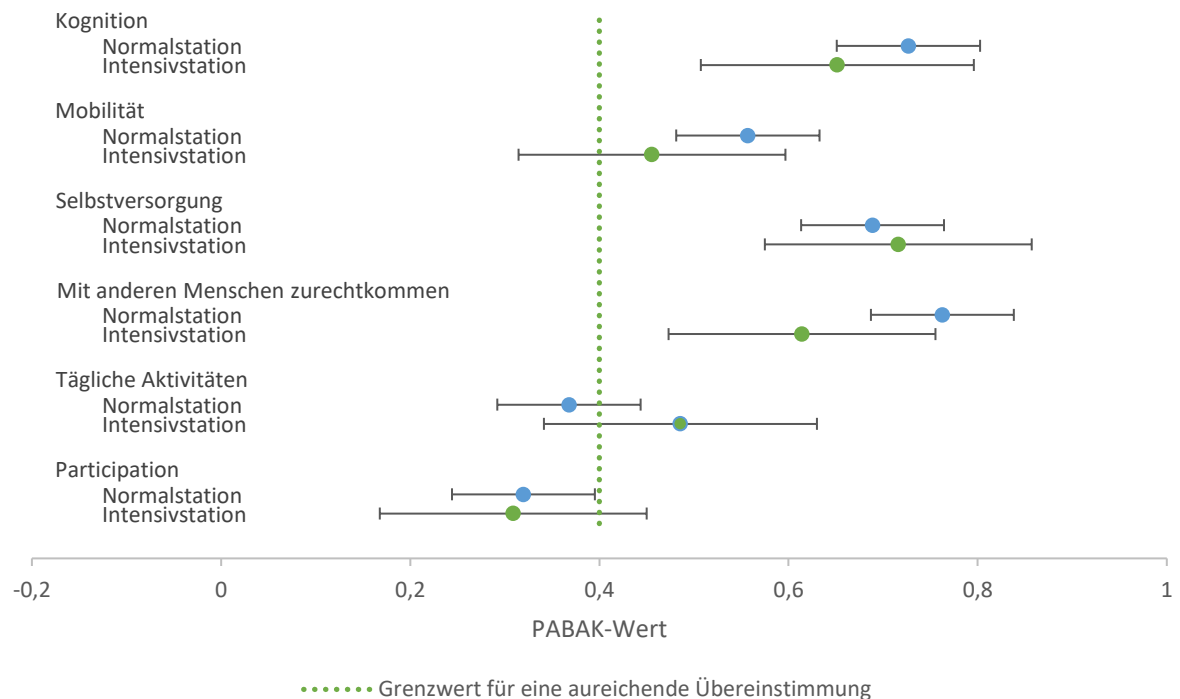


Abbildung 19: Vergleich der PABAK-Werte in den Gesundheitsdomänen des WHODAS 2.0 zwischen den verschiedenen Gruppen im Fallkontrollabgleich. Die PABAK-Werte sind mit ihren 95% Konfidenzintervall (Fehlerindikatoren) dargestellt.

4.4 Ergebnisse zur Fragestellung 3

4.4.1 Gruppencharakteristika

Anzahl der Patient*innen und Stellvertreter*innen je Gruppe

In Tabelle 11 ist die Anzahl der Patient*innen und Stellvertreter*innen in den jeweiligen Gruppen dargestellt. Bei den meisten Gruppen liegt eine ungleiche Verteilung vor, mit der kleinsten Population in der Gruppe der Patient*innen die pflegerisch unterstützt werden.

Tabelle 11: Anzahl in den jeweiligen Gruppen der Fragestellung 3

Charakteristika	Gruppe 1 (n)	Gruppe 2 (n)
Leben in häuslicher Gemeinschaft (Nein/Ja)	61	143
Verhältnis (Nicht-Ehepartner*innen/ Ehepartner*innen)	84	120
Pflege durch die Stellvertreter*innen (Nein/Ja)	159	45
Punktzahl im MOCA-Blind ($\leq 17/\geq 18$)	84	104
Geschlecht der Patient*innen (w/m)	121	83
Geschlecht des Stellvertreter*innen (w/m)	65	138
Alter der Patient*innen ($\leq 64/\geq 65$ Jahre)	80	124
Alter der Stellvertreter*innen ($\leq 64/\geq 65$ Jahre)	107	93

*MOCA-Blind = Montreal Cognitive Assessment - Blind, m = männlich, w = weiblich

Medianer Punktwert der Patient*innen in der Gesamtpunktzahl des WHODAS 2.0

In der Tabelle 12 ist der mediane Punktwert der Patient*innen in der Gesamtpunktzahl des WHODAS 2.0 dargestellt. Bei dem Vergleich der Verteilung mittels des Man-Whitney-U-Test, liegt der einzig signifikante Unterschied in der „Pflege“-Gruppe vor. Dies bedeutet, dass Patient*innen welche gepflegt wurden, mehr Einschränkungen als Patient*innen hatten die nicht gepflegt wurden.

Tabelle 12: Mediane Punktwerte der Patient*innen in der Gesamtpunktzahl des WHODAS 2.0 in den jeweiligen Gruppen der Fragestellung 3

Charakteristika	Medianer Punktwert der Patient*innen in der Gesamtpunktzahl des WHODAS 2.0 (IQA)		p – Wert*
	Gruppe 1	Gruppe 2	
Leben in häuslicher Gemeinschaft (Nein/ Ja)	12,5 (6,3 bis 25,0)	12,5 (4,2 bis 28,1)	0,911
Verhältnis (Nicht-Ehepartner*innen/ Ehepartner*innen)	14,6 (6,3 bis 27,1)	11,9 (4,2 bis 25,0)	0,296
Pflege durch die Stellvertreter*innen (Nein/ Ja)	10,4 (4,2 bis 20,8)	16,7 (8,3 bis 43,8)	0,003
Punktzahl im MOCA-Blind ($\leq 17/ \geq 18$)	12,5 (4,2 bis 28,1)	11,5 (6,3 bis 25,0)	0,758
Geschlecht der Patient*innen (w/ m)	10,4 (4,2 bis 27,1)	12,5 (6,3 bis 24,0)	0,623
Geschlecht der Stellvertreter*innen (w/ m)	16,7 (10,4 bis 33,3)	10,4 (4,2 bis 25,0)	0,120
Alter der Patient*innen ($\leq 64/ \geq 65$ Jahre)	14,6 (6,3 bis 30,2)	11,4 (4,2 bis 24,0)	0,770
Alter der Stellvertreter*innen ($\leq 64/ \geq 65$ Jahre)	12,5 (5,2 bis 27,1)	10,4 (4,2 bis 21,9)	0,154

*Die Unterschiede im Median wurden mittels Mann-Whitney-U-Test auf Signifikanz getestet.
MOCA

4.3.2 Systematischer Bias

In der Tabelle 13 ist die mediane Differenz zwischen der Punktzahl der Patient*innen und deren Stellvertreter*innen in der Gesamtpunktzahl dargestellt. Die Unterschiede sind in den Gruppen „Verhältnis“, „Pflege“, „Geschlecht der Stellvertreter*innen“ und „Alter des Patient*innen“ signifikant. Nur bei Stellvertreter*innen, welche die Patient*innen pflegten, lag die mediane Differenz oberhalb der MCID von 10 Punkten (Vgl. Abschnitt 3.6).

Tabelle 13: Differenz zwischen den Punktwerten der Patient*innen und deren Stellvertreter*innen in der Gesamtpunktzahl des WHODAS 2.0 bei den jeweiligen Gruppen der Fragestellung 3

Charakteristika	Mediane Differenz in der Gesamtpunktzahl (IQA)		p - Wert*
	Gruppe 1	Gruppe 2	
Leben in häuslicher Gemeinschaft (Nein/Ja)	-5,2 (-15,6 bis 2,1)	-2,1 (-10,4 bis 6,3)	0,078
Verhältnis (Nicht-Ehepartner*innen/ Ehepartner*innen)	0 (-8,3 bis 6,3)	-6,3 (-18,8 bis 2,1)	0,006
Pflege durch den Stellvertreter*innen (Nein/Ja)	0 (-8,3 bis 6,3)	-13,5 (-26 bis -1)	<0,001
Punktzahl im MOCA ($\leq 17/\geq 18$)	-2,1 (-10,4 bis 4,2)	-2,1 (-12,5 bis 6,3)	0,968
Geschlecht des Patient*innen (w/m)	-4,2 (-12,5 bis 2,1)	0 (-8,3 bis 6,3)	0,123
Geschlecht des Stellvertreter*innen (w/m)	0 (-10,4 bis 7,3)	-4,2 (-12,5 bis 2,1)	0,048
Alter des Patient*innen ($\leq 64/\geq 65$ Jahre)	0 (-8,3 bis 8,3)	-4,2 (-14,6 bis 2,1)	0,004
Alter des Stellvertreter*innen ($\leq 64/\geq 65$ Jahre)	-2,1 (-14,6 bis 6,3)	-2,1 (-12,5 bis 4,2)	0,765

*Die Unterschiede im Median wurden mittels Mann-Whitney-U-Test auf Signifikanz getestet. Die Differenz wurde als Punktwert der Patient*innen minus dem Punktwert der Stellvertreter*innen berechnet.

4.3.3 Vergleich der ICC

In der Abbildung 22 sind die ICC-Werte der Gesamtpunktzahl der verschiedenen Gruppen dargestellt. Der Unterschied zwischen den Gruppen in der ICC liegt nur bei dem Charakteristikum „Pflege“ über 0,25. Die ICC ist in der Gruppe von Stellvertreter*innen welche die Patient*innen pflegen, als nicht ausreichend zu bewerten (ICC < 0,5). Der zweitgrößte Unterschied in der ICC liegt bei der Gruppe „Alter der Stellvertreter*innen“, mit einem Unterschied von 0,16 vor.

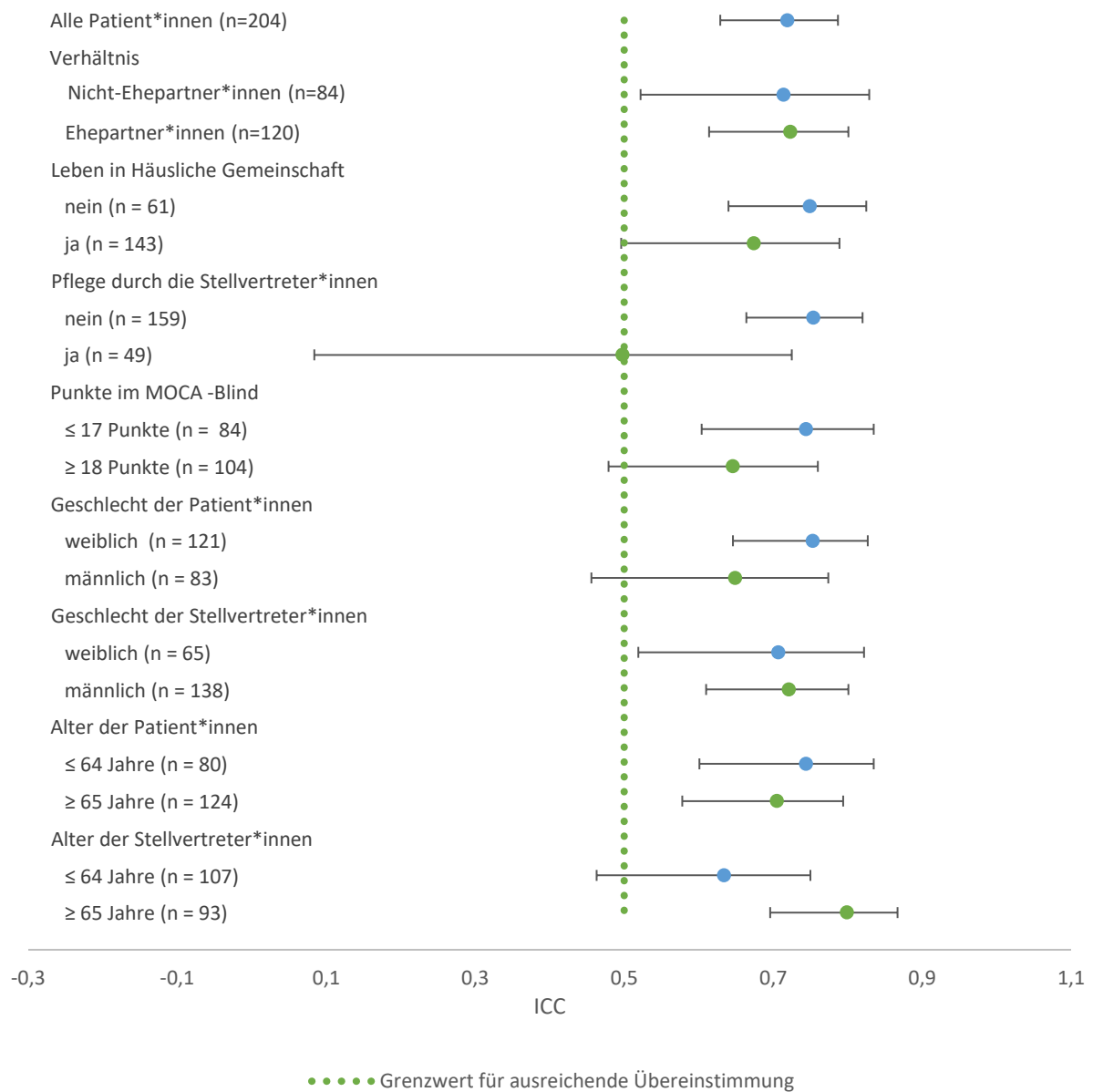


Abbildung 20: ICC der verschiedenen Gruppen der Fragstellung 3 in der Gesamtpunktzahl des WHODAS 2.0. Die ICC ist jeweils mit dem 95% Konfidenzintervall angegeben (Fehlerindikatoren).

4.4.4 Vergleich der LOA nach Bland-Altman

In der Abbildung 21 sind die in der Bland-Altman-Analyse ermittelten Mittelwerte der Differenz mit den oberen und unteren LOA dargestellt. Bei allen der in Fragestellung 3 beschriebenen Gruppen (Vgl. Abschnitt 3.6) liegen die LOA deutlich außerhalb der MCID. Bei der Gruppe der Stellvertreter*innen, welche die Patient*innen pflegten, liegt ebenfalls der Mittelwert der Differenz außerhalb der MCID.

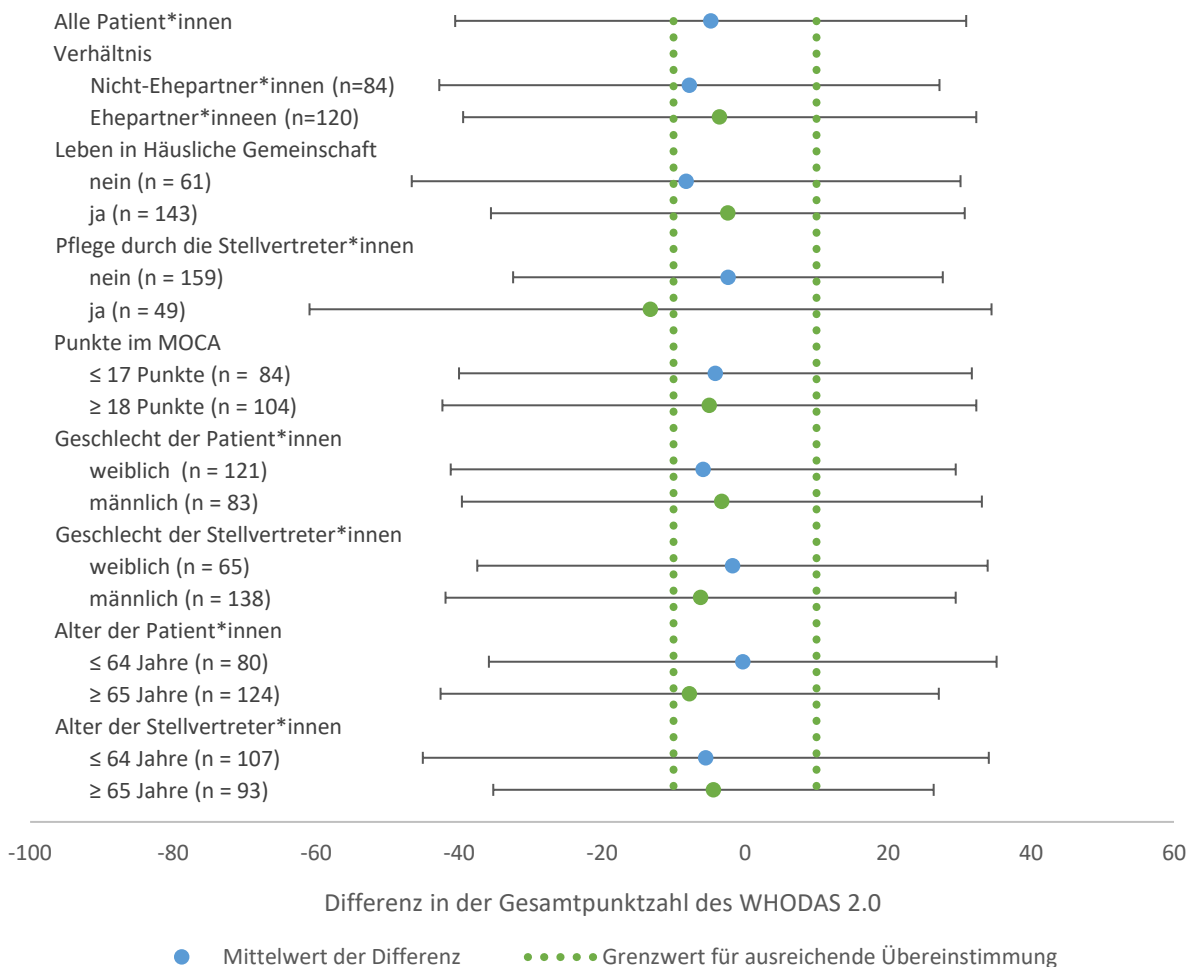


Abbildung 21: Ergebnisse der Bland-Altman-Analyse bei den verschiedenen Gruppen der Fragestellung 3. Die Enden der Fehlerindikatoren entsprechen den unteren und oberen LOA. Die Differenz wurde als Punktzahl der Patient*innen minus der Punktzahl der Stellvertreter*innen berechnet.

4.4.5 Fallkontrollabgleich für die Gruppen „keine Pflege“ und „Pflege“

Gruppencharakteristika

In Tabelle 14 sind die Gruppencharakteristika für den Fallkontrollabgleich anhand der Einschränkungen der Patient*innen dargestellt. Zwischen den Gruppen liegen keine signifikanten Unterschiede in den überprüften Eigenschaften vor.

Tabelle 14: Gruppencharakteristika der Patient*innen und derer Stellvertreter*innen im Fallkontrollabgleich der Fragstellung 3

		Gruppe 1 („keine Pflege“)	Gruppe 2 („Pflege“)	p-Wert*
Gruppengröße		41	41	
Mediane Punktzahl der Patient*innen (IQA)		16,7 (8,3 bis 37,5)	16,7 (8,3 bis 37,5)	0,996
Alter der Patient*innen (Mittelwert, SD)		69 (59 bis 75)	71 (61 bis 78)	0,551
Alter der Stellvertreter*innen (Mittelwert, SD)		66 (54,8 bis 77)	64 (53,8 bis 71)	0,404
Geschlecht der Patient*innen (N, %)	Weiblich	29 (70,7%)	23 (56,1%)	0,169
	Männlich	12 (29,3%)	18 (43,9%)	
Geschlecht der Stellvertreter*innen (N, %)	Weiblich	15 (36,6%)	14 (34,1%)	0,817
	Männlich	26 (63,4%)	27 (65,9%)	
Postoperative Station	Normalstation	18 (43,9%)	19 (46,3%)	0,391
	Intensivstation	23 (56,1%)	22 (53,7%)	
Station der Befragung	Normalstation	35 (85,4%)	32 (78%)	0,391
	Intensivstation	6 (14,6%)	9 (22%)	

*Die Signifikanz wurde für die Punktzahl mittels Mann-Whitney-U-Test, für das Alter mittels T-Test und für die restlichen Eigenschaften mittel Chi-Quadrat-Test berechnet.

Mediane Differenz der Gesamtpunktzahl

In der Gruppe 1 („keine Pflege“) liegt die mediane Differenz der Gesamtpunktzahl bei 4,17 (IQA: -4,17 bis 8,33). In der Gruppe 2 („Pflege“) liegt die mediane Differenz bei der Gesamtpunktzahl bei -12,5 (IQA: -27,08 bis -2,083). Die Signifikanz dieses Unterschiedes beträgt im Mann-Whitney-U-Test $p < 0,001$.

LOA nach Bland-Altman für die Übereinstimmung in der Gesamtpunktzahl

In der Gruppe 1 („keine Pflege“) liegt der Mittelwert der gepaarten Differenz bei 3,14, das untere LOA bei -12,9 und das obere LOA bei 25,29. In der Gruppe 2 („Pflege“) liegt der Mittelwert der gepaarten Differenz bei -15,32, das untere LOA bei -60,37 und das obere LOA bei 29,73. Die LOA liegen somit in beiden Gruppen außerhalb der MCID.

ICC der Gesamtpunktzahl

Die ICC liegt in der Gruppe 1 („keine Pflege“) bei 0,876 (95 % KI: 0,769 bis 0,934) und in der Gruppe 2 („Pflege“) bei 0,428 (95 % KI: -0,067 bis 0,694). Zwischen den Gruppen liegt somit ein, nach den in Abschnitt 3.6 festgelegten Grenzwerten, relevanter Unterschied ($>0,25$) vor.

5 Diskussion

5.1 Diskussion der Ergebnisse

5.1.1 Diskussion der Ergebnisse zur Fragestellung 1

Im Rahmen der ersten Fragestellung dieser Dissertation wurde die Übereinstimmung zwischen der Eigeneinschätzung von Patient*innen und einer Fremdeinschätzung, bei der Beurteilung der prämorbidem gesundheitsbezogenen Lebensqualität mittels des WHODAS 2.0 ermittelt. Hierbei wurde mittels verschiedener statistische Methoden die Übereinstimmung in den einzelnen Fragen, den Gesundheitsdomänen und der Gesamtpunktzahl des WHODAS 2.0 betrachtet.

Die Abweichung zwischen der Eigen- und Fremdeinschätzungen in der Gesamtpunktzahl des WHODAS 2.0 war bei 64 % der Proband*innen kleiner als die MCID (10 Punkte). Bei den einzelnen Fragen des WHODAS 2.0 lag der Prozentsatz der Stellvertreter*innen, die um maximal eine Kategorie von der Eigeneinschätzung abgewichen sind, bei 70,9 % bis 88,6 % und bei den Gesundheitsdomänen bei 86,6 % bis 89,8 %. Eine Abweichung von jeweils einer Kategorie in jeder der einzelnen Fragen würde einen Unterschied von 25 Punkten in der Gesamtpunktzahl bedeuten. Wenn dieser Grenzwert als MCID in der Gesamtpunktzahl benutzt wird, würde der Prozentsatz an ausreichend übereinstimmenden Paaren bei 83,2 % liegen. Bei Ramos et al. wurde eine Differenz bis 30 Punkten in der Gesamtpunktzahl des WHODAS 2.0 als ausreichend gewertet. (33). Eine MCID von 25 bzw. 30 Punkten würde allerdings keine Unterscheidung zwischen „nicht“, „leicht“ oder „moderat“ Eingeschränkten Patient*innen ermöglichen und liegt weit außerhalb der in anderen Studien ermittelten MCID (43-45).

Es konnte kein relevanter Bias zur Über- oder Unterschätzung der prämorbidem Lebensqualität durch die Stellvertreter*innen festgestellt werden. Die Differenz liegt bei den meisten der einzelnen Fragen und Gesundheitsdomänen im Median bei 0 und in der Gesamtpunktzahl bei -2 Punkten. Diese Differenz ist deutlich unterhalb der MCID. In der klinischen Anwendung ist allerdings vor allem die Übereinstimmung zwischen einzelnen Personen von Interesse.

Deswegen wurden die LOA, ICC, gewichteten Kappa-Werte und PABAK-Werte berechnet, welche eine Analyse der Übereinstimmung einzelner Paare ermöglichen.

In der ICC konnte eine moderate bis gute Übereinstimmung zwischen Patient*innen und deren Stellvertreter*innen in der Gesamtpunktzahl des WHODAS 2.0 nachgewiesen werden (ICC 0,72, 95% KI: 0,62 bis 0,79). Demnach wäre die Fremdbeurteilung eine verlässliche Informationsquelle. Ein ähnlicher ICC Wert wurde bei Schlote et al. in der Gesamtpunktzahl der 36 Fragen Version des WHODAS 2.0 gefunden (ICC: 0,84, 95% KI 0,70 bis 0,92) (32). Die LOA nach Bland-Altman liegen bei der Gesamtpunktzahl allerdings deutlich oberhalb der MCID (LOA: -40,5 bis 30,9). Diese LOA bedeuten, dass es keine 95% Sicherheit gibt, dass Fremdeinschätzung in einem ausreichenden Maß (i.e. geringer als die MCID) mit der Eigeneinschätzung übereinstimmt. Auch rein deskriptiv liegt bei 38% der Fremdbewertungen eine größere Abweichung als die MCID vor. Dies könnte zur Folge haben, dass eigentlich gesunden Patient*innen als schwer eingeschränkt angesehen werden und dadurch falsche Therapieziele gesetzt werden. Bei Hofhuis et al. und Capuzzo et al. wurden ähnliche großen LOA gefunden (24, 25). Eine 95% Übereinstimmung der Eigen- und Fremdeinschätzung wurde allerdings in keiner der Studien erreicht, welche in der Literaturrecherche gefunden wurden.

Bei den einzelnen Fragen und Gesundheitsdomänen lag, gemessen an den gewichteten Kappa-Werten, eine geringe bzw. nicht ausreichend Übereinstimmung vor. Nur in der Frage zu Einschränkungen beim Laufen (Frage 7) und der Gesundheitsdomäne „Mobilität“, lag eine mindestens moderate Übereinstimmung ($K_w > 0,5$) vor. Die Aussagekraft des gewichteten Kappas ist in dieser Studie allerdings zu hinterfragen. Besonders deutlich wird dies bei Frage 10. Bei dieser Frage liegt der höchste Prozentsatz gleicher und ähnlicher Antworten (88,6%) vor, trotzdem besteht hier der niedrigste Kappa Wert ($K = 0,04$, 95% KI: -0,06 bis 0,150). Dieser Unterschied ist, wie im Methodik Teil dieser Dissertation beschrieben, durch die ungleiche Verteilung der Antworten zu erklären, welche bei dieser Frage am stärksten ausgeprägt ist. Dieser Zusammenhang kann auch in anderen Studien beobachtet werden (25, 31). Die Eigenschaft des (gewichteten) Kappa-Wertes ist beabsichtigt und soll für die Wahrscheinlichkeit einer zufälligen Übereinstimmung bei einer ungleichen Verteilung korrigieren. Es wird allerdings von verschiedenen Autor*innen beschrieben, dass eine Überkorrektur stattfinden kann (50, 51, 56). Um diese Problematik zu lösen, wurde in dieser Dissertation ebenfalls der PABAK-Wert berechnet. Der PABAK-Wert wurde in keiner der anderen Studien genutzt, ist aber bei der ungleichen Verteilung eine gute Alternative zu dem (gewichteten) Kappa-Wert (50). Die PABAK-Werte zeigen bei den meisten der einzelnen

Fragen und Gesundheitsdomänen eine ausreichende Übereinstimmung ($PABAK > 0,5$). Die einzige Ausnahme bilden die Fragen 4 und 5, welche die emotionale Belastung und Einschränkungen der Teilhabe abfragen, sowie Frage 12, welche die Bewältigung des Alltags als Gegenstand hat (Vgl. Abschnitt 4.3).

In den meisten der anderen Studien wurde die Übereinstimmung zwischen den Patient*innen und Stellvertreter*innen mittels des gewichteten oder ungewichteten Kappa ermittelt, weiterhin wurde in einigen der Studien die Bland-Altman-Analyse (21, 25) oder die Intraklassen-Korrelation (19, 21, 22, 28, 32) genutzt. In 2 der Studien wurde nur der Prozentsatz ähnlicher bzw. gleicher Antworten ermittelt (20, 33). Die Übereinstimmung zwischen Patient*innen und deren Stellvertreter*innen unterscheidet sich in den verschiedenen Studien bzw. Fragebögen sowie auch innerhalb der jeweiligen Fragebögen deutlich. Die ermittelte Übereinstimmung reicht dabei von einer fast perfekten (definiert als $ICC > 0,9$) bis zu nicht vorhanden. In den Studien, bei welchen der WHODAS 2.0 genutzt wurde, lag eine höhere Übereinstimmung in den Gesundheitsdomänen „Mobilität“ und „tägliche Aktivitäten“, als in den Gesundheitsdomänen „Kognition“ und „Mit anderen Menschen zurechtkommen“ vor. Dieser Unterschied wird von den Autor*innen dadurch erklärt, dass von außen sichtbare Prozesse leichter zu beurteilen sind als innerliche Prozesse (32, 33). Diese Tendenz wurde auch in verschiedenen der anderen Studien beschrieben (20, 23, 25, 57). In dieser Dissertation gab es, gemessen an dem PABAK-Werten, ein umgekehrtes Verhältnis.

Eine mögliche Erklärung für diesen Unterschied ist, dass unterschiedliche Ausmaß von Einschränkungen in den verschiedenen Fragen und Gesundheitsdomänen. Bei den Gesundheitsdomänen „Mobilität“ und „Partizipation“ haben nur 39 % bzw. 23 % der Patient*innen angegeben „keine Einschränkungen“ zu haben. In den Gesundheitsdomänen „Selbstversorgung“ und „mit anderen Menschen zurechtkommen“ lag dieser Prozentsatz bei jeweils 84% und 78%. Dies deutet darauf hin, dass die höchste Übereinstimmung bei den Fragen mit den geringsten Einschränkungen vorlag. Diese Vermutung konnte in der Regressionsanalyse zu dem Zusammenhang zwischen dem Prozentsatz von Patient*innen, die angegeben haben „keine Einschränkungen“ zu haben und dem PABAK-Wert, bestätigt werden. Das Bestimmtheitsmaß lag in dieser Analyse bei 0,8 ($p < 0,001$). Dies bedeutet, dass fast der gesamte Unterschied in den PABAK-Werten durch den verschiedenen Anteil an Patient*innen ohne Einschränkungen erklärt werden kann (58). Mittels der vorliegenden Daten kann also nicht beantwortet werden, ob die Unterschiede in der Übereinstimmung durch eine unterschiedliche

Kenntnis der Stellvertreter*innen über die verschiedenen Gesundheitsaspekte entstanden sind. In den anderen Studien zu diesem Thema wurde diese Problematik nicht weiter diskutiert.

Zusammenfassend ist die Übereinstimmung zwischen den hier befragten Patient*innen und Stellvertreter*innen, gemessen an den ICC und PABAK-Werten, in den meisten der Fragen, den Gesundheitsdomänen sowie der Gesamtpunktzahl nach statistischen Definitionen als moderat bis gut zu bewerten. Für einzelne Patient*innen besteht allerdings keine 95% Sicherheit, dass die Fremdeinschätzung um weniger als die MCID von der Selbsteinschätzung abweicht. Dies ist besonders relevant, wenn die Angabe des Stellvertreters dazu führen würde, dass die Therapieziele oder Therapie begrenzt werden sollten. Es ist zu beachten, dass die Grenzwerte für eine „ausreichende“ Übereinstimmung in der ICC, der gewichteten Kappa- und PABAK-Werte sowie der MCID in dieser Dissertation nicht empirisch ermittelt wurden, sondern nach klinischen Überlegungen, in Referenz zu statistischen Empfehlungen sowie den Grenzwerten in anderen Studien festgelegt wurden. In der klinischen Anwendung stehen oft keine anderen Quellen zur prämorbidem Lebensqualität der Patient*innen zur Verfügung, sodass, wenn keine Stellvertreter*innen befragt werden können, keine direkten Aussagen zur prämorbidem Lebensqualität möglich sind. Somit ist die Frage zu beantworten, ob einige Fehleinschätzungen in Kauf genommen werden können um einen Großteil der Patient*innen potenziell besser behandeln zu können. Diese Frage kann nicht im Rahmen dieser Dissertation beantwortet werden und muss durch weitere Studien, welche z. B. die Behandlungserfolge der verschiedenen Ansätze vergleichen, geklärt werden. Hofhuis et al. Schlussfolgern weiterhin, dass die hohen interindividuellen Abweichungen der Eigen- und Fremdbeurteilung ein Grund dafür sind, dass in Studien, welche die prämorbidem Lebensqualität von Patient*innen untersuchen, Eigen- und Fremdbeurteilungen nicht vermischt werden sollten (24).

Da die Übereinstimmung in den Fragen mit der geringsten Häufigkeit von Einschränkungen am höchsten war, können diese Ergebnisse nicht für eine deutlich kränkere Population übernommen werden. In einer länderübergreifenden Studie der WHO zum WHODAS 2.0, lag der Gesamtpunktwert einer zufällig ausgewählten Referenzpopulation bei 0 Punkten und 11,1 Punkten in dem 82. Perzentil (36). Der Mediane Punktwert in der Eigeneinschätzung liegt bei den in dieser Dissertation befragten Patient*innen bei 13 Punkten. Die hier befragte Population weist also mehr Einschränkungen der Lebensqualität als die Referenzpopulation der WHO auf. Bei Ramos-Remus et al. und Schlote et al. lag der durchschnittliche Gesamtpunktwert je nach Vorerkrankung bei 17, 30 oder 27 Punkten und in einer Studie zu chronischen Schmerzpatient*innen bei 26 Punkten (32, 33, 59). Es sind also weitere Studien nötig, ob der

WHODAS 2.0 auch bei der Fremdbeurteilung der prämorbidem gesundheitsbezogenen Lebensqualität von Populationen geeignet ist, bei denen eine höhere Einschränkung der prämorbidem Lebensqualität vorliegt.

5.1.2 Diskussion zur Fragestellung 2

In Fragestellung 2 wurde untersucht, ob die Behandlung der Patient*innen auf einer Intensivstation einen relevanten Einfluss auf die Validität der Fremdbeurteilung hat. Zur Beantwortung dieser Frage wurde die Population in zwei Gruppen eingeteilt und getestet, ob ein relevanter systematischer Bias zur Über- oder Unterschätzung der prämorbidem Lebensqualität vorliegt. Um Unterschiede in der Übereinstimmung auf der interindividuellen Ebene zu ermitteln, wurden ebenfalls die LOA und ICC sowie die gewichteten Kappa- und PABAK-Werte zwischen den Gruppen verglichen.

Zwischen den beiden Gruppen lag ein statistisch signifikanter Bias vor. Stellvertreter*innen von Patient*innen, die intensivmedizinisch behandelt wurden, haben die prämorbidem Einschränkungen der Patient*innen, gemessen an der Gesamtpunktzahl des WHODAS 2.0, im Median um 6 Punkte überschätzt, im Vergleich zu 0 Punkten in der Referenzgruppe ($p = 0,009$). Dieser Bias ist allerdings kleiner als die MCID von 10 Punkten. Bei 3 der Gesundheitsdomänen, „Kognition“, „Mit anderen Menschen zurechtkommen“ und „Tägliche Aktivitäten“ konnte ebenfalls ein signifikanter Unterschied nachgewiesen werden. Da Patient*innen, welche auf einer Intensivstation behandelt wurden, im Median mehr Einschränkungen ihrer prämorbidem Lebensqualität angegeben haben, wurde ein Fallkontrollabgleich durchgeführt. Bei dieser Analyse wurde ein nicht-signifikanter Unterschied von nur 4 Punkten zwischen den Gruppen gefunden ($p = 0,09$). Bei dieser Analyse lag allerdings eine zu geringe Fallzahl für eine ausreichende Power vor.

Die ICC ist in der Gruppe der intensivmedizinisch behandelten Patient*innen niedriger (0,699 vs. 0,738). Die Übereinstimmung ist in beiden Gruppen anhand der ICC als moderat bis gut (Siehe Abschnitt 3.6) einzuschätzen und oberhalb des in dieser Dissertation festgelegten Grenzwertes von 0,5. Die LOA sind in der Gruppe der intensivmedizinisch behandelten Patient*innen aufgrund des vorliegenden Bias ins Negative verschoben. Die Standarddeviation der LOA ist in beiden Gruppen fast identisch (17,7 vs. 18,2), sodass in beiden Gruppen eine vergleichbare Präzision vorliegt. Bei dem Fallkontrollabgleich war die ICC in der Gruppe der intensivmedizinisch behandelten Patient*innen sogar leicht höher (0,700 vs. 0,694).

Die gewichteten Kappa-Werte sind bei beiden Gruppen, in allen der getesteten Gesundheitsdomänen, als nicht ausreichend zu werten. Bei den gewichteten Kappa-Werten liegen auch hier, die in der Methodik und Diskussion zur Fragestellung 1 beschriebenen Einschränkungen vor. Deswegen wurde keine weitere Betrachtung der Unterschiede in den verschiedenen Gesundheitsdomänen vorgenommen.

Die PABAK-Werte sind in der Gruppe der intensivmedizinisch behandelten Patient*innen bei allen Gesundheitsdomänen des WHODAS 2.0 niedriger als bei Patient*innen, die nicht intensivmedizinisch behandelt wurden. Der größte Unterschied zwischen den Gruppen, ist in der Gesundheitsdomäne „Mobilität“ vorhanden und ist in der Gruppe der intensivmedizinisch behandelten Patient*innen als nicht ausreichend zu werten. Bei allen anderen Gesundheitsdomänen lag kein relevanter Unterschied vor. Bei genauerer Betrachtung der Daten ist allerdings zu erkennen, dass Patient*innen welche intensivmedizinisch behandelt wurden, mehr Einschränkungen der Mobilität angegeben haben. Die mediane Punktzahl liegt bei intensivmedizinisch behandelten Patient*innen bei 3 (IQA 0-5) und bei nicht intensivmedizinisch behandelten Patient*innen bei 0 (IQA 0-3). Bei dem durchgeführten Fallkontrollabgleich lag der Unterschied der PABAK-Werte in der Gesundheitsdomäne Mobilität unter 0,2.

Zusammenfassend konnte gezeigt werden, dass bei intensivmedizinischer Behandlung zwar einen signifikanten Bias bei der Einschätzung der prämorbidem Lebensqualität vorliegt, dieser allerdings kleiner als die MCID ist. Bei der ICC, LOA und den PABAK-Werten lag nur bei der Gesundheitsdomäne „Mobilität“ ein relevanter Unterschied in der Übereinstimmung vor. Die Ergebnisse des Fallkontrollabgleichs deuten darauf hin, dass ein Teil des Unterschiedes durch die Unterschiede in den Einschränkungen der Patient*innen entstanden ist und weitere Studien mit adäquater Fallzahl nötig sind um diesen Einfluss näher zu erforschen.

5.1.3 Diskussion zur Fragestellung 3

In Fragestellung 3 wurde untersucht, ob ausgewählte Eigenschaften der Patient*innen oder der Stellvertreter*innen, einen Einfluss auf die Übereinstimmung bei der Beurteilung der prämorbidem Lebensqualität des Patient*innen mittels des WHODAS 2.0 haben.

Es konnte gezeigt werden, dass bei verschiedenen Gruppen ein signifikanter Bias zur Unterschätzung der prämorbidem Lebensqualität vorlag. Dabei handelte es sich um nicht mit den Patient*innen verheiratete Stellvertreter*innen, männliche Stellvertreter, Stellvertreter*innen, welche die Patient*innen pflegen und Patient*innen die älter als 64 Jahre

alt sind. Allerdings war dieser Bias nur bei der Gruppe der „Pflege“ größer als die MCID von 10 Punkten. Bei dem Vergleich der ICC konnte ebenfalls nur in der Gruppe „Pflege“ ein relevanter Unterschied nachgewiesen werden, mit einer deutlich geringeren ICC in der Gruppe der Stellvertreter*innen, welche die Patient*innen pflegerisch unterstützt haben. Bei keiner der Subgruppen liegen die LOA innerhalb der MCID. Zusammenfassend gibt es demnach keine Gruppe, bei der eine 95% Übereinstimmung zwischen der Fremd- und Eigeneinschätzung vorliegt und es konnte nur zwischen den Gruppen „Pflege“ und „keine Pflege“ ein klinisch relevanter Unterschied in der Validität der Fremdeinschätzung festgestellt werden.

Da Patient*innen, welche durch gepflegt wurden, deutlich höhere Einschränkungen ihrer prämorbidem Lebensqualität angegeben haben, wurde ein Fallkontrollabgleich anhand der medianen Punktzahl der Patient*innen im WHODAS 2.0 durchgeführt. Bei dieser Analyse blieben die Unterschiede zwischen den beiden Gruppen bestehen. Eine mögliche Erklärung für diese Ergebnisse ist, dass durch die Pflege ein Fokus auf die Einschränkungen in der Lebensqualität besteht. Weiterhin wird beschrieben, dass die mögliche Belastung durch eine pflegerische Aktivität, der sogenannte „Caregiver Burden“, einen Einfluss auf die Einschätzung der Lebensqualität hat (60). Die subjektive Belastung der Stellvertreter*innen durch die Pflege wurde in dieser Dissertation allerdings nicht überprüft. Es ist weiterhin auch möglich, dass Patient*innen, welche gepflegt wurden, ihre eigene Lebensqualität überschätzt haben. Die Einschätzung der Lebensqualität ist allerdings subjektiv und in dieser Studie wurde die Eigeneinschätzung der Patient*innen als Goldstandard verwendet. In dieser Studie gab es auch nur einen kleinen Anteil an Stellvertreter*innen, welche die Patient*innen pflegerisch unterstützten ($n = 54$), sodass keine weitere Graduierung in Abhängigkeit zum wöchentlichen Zeitaufwand der Pflege möglich war. Um die Gründe für diese Unterschiede zu ermitteln und um zu testen, ob diese Unterschiede durch eine Überschätzung der Lebensqualität in der Eigenbewertung oder Unterschätzung in der Fremdbewertung zu erklären sind, werden weitere Studien benötigt.

Bei allen der restlichen Subgruppen lag die ICC deutlich über 0,5 und der Unterschied zwischen den Gruppen bei weniger als 0,25. In Bezug auf diese Charakteristika können somit keine Empfehlungen gegeben werden, welche Stellvertreter*innen bevorzugt befragt werden sollten.

Nur in einer anderen Studie wurde der Einfluss des Alters und Geschlechts der Patient*innen untersucht. In dieser Studie wurde kein signifikanter Einfluss dieser Charakteristika nachgewiesen (22). In Studien, welche den Einfluss des Geschlechtes der Stellvertreter*innen untersucht haben, konnte ebenfalls kein signifikanter Einfluss nachgewiesen werden (25, 28,

31). Wie auch aus den Daten dieser Dissertation hervorgeht, beschreibt Magaziner et al. eine höhere Übereinstimmung bei älteren Stellvertreter*innen (> 65 Jahre) (31). Bei Magaziner et al. wurde die Übereinstimmung allerdings mittels der Pearson-Korrelation gemessen, welche nicht geeignet ist um die Übereinstimmung zweier Messmethoden zu bestimmen (46).

Zusammenfassend konnte nur für Stellvertreter*innen, welche den Patient*innen pflegen ein relevanter Bias gezeigt werden. Dies wurde in noch keiner anderen Studie untersucht und sollte aus den oben genannten Gründen weiter untersucht werden. Die Angaben von Stellvertreter*innen, welche die Patient*innen pflegen, sollten aufgrund der vorliegenden Ergebnisse mit besonderer Vorsicht genutzt werden und wenn möglich durch die Befragung weiterer Stellvertreter*innen gegengeprüft werden.

Eine Limitation ist, dass die Grenzen für einen relevanten Unterschied anhand der statistischen Methoden festgelegt wurden. Die Grenzen wurden so gewählt, dass ein relevanter Unterschied, einen Unterschied von mindestens einer Kategorie in der Interpretation der Werte bedeutet (48). Hierzu liegen keine empirischen Werte vor, sodass je nach Fragestellung auch schon geringere Unterschiede relevant sein könnten. Ein weiterer Kritikpunkt ist die Einteilung in der Subgruppe „Verhältnis“ in „Ehepartner*innen“ und „Nicht-Ehepartner*innen“, da hierbei viele verschiedene Gruppen von Stellvertreter*innen wie z. B. Kinder, Eltern oder nahe Bekannte vermischt wurden. Diese Einteilung ist allerdings durch den geringeren Anteil an „Nicht-Ehepartner*innen“ bedingt gewesen.

5.2 Stärken und Limitationen

Zu den Stärken dieser Dissertation gehört das prospektive Design der Studie. Patient*innen wurden bereits zur ihrer Gesundheit befragt wurden, bevor eine akute Einschränkung durch eine Operation vorlag. Somit wurde versucht einen „Recall Bias“ zu verhindern (13). Durch dieses Vorgehen sind aber auch Limitationen bedingt. Es ist möglich, dass aufgrund der bevorstehenden Operation ein Bias bei der Einschätzung der Lebensqualität entstanden ist. Weiterhin gibt es nur Evidenz aus einer Studie, dass durch die Situation einer geplanten Operation eine ähnliche Belastung für die Stellvertreter*innen erzeugt wird, wie bei einem akuten Trauma oder einer unerwarteten Erkrankung (21). Eine weitere Stärke dieser Dissertation ist das multizentrische Design und die im Vergleich zu ähnlichen Studien große Studienpopulation.

Durch die Befragung von intensivmedizinisch und nicht-intensivmedizinisch behandelter Patient*innen, ist anders als in anderen Studien zu diesem Thema, ein direkter Vergleich zwischen diesen beiden Gruppen möglich gewesen. Hierdurch konnte in dieser Studie der Einfluss der intensivmedizinischen Behandlung auf die Übereinstimmung zwischen den Patient*innen und Stellvertreter*innen möglichst direkt gemessen werden. Eine weitere Kritik an der Methodik ist, dass auch die Selbsteinschätzung der Patient*innen zu ihrer Lebensqualität subjektiv ist. Diese Problematik kommt vor allem bei den Ergebnissen der Fragestellung 3 zum Einfluss der Pflege des Patient*innen auf die Übereinstimmung zum Tragen.

In dieser Dissertation wurden keine Scoring-Systeme zur Erfassung der Erkrankungsschwere von Patient*innen wie z. B. der APACHE II-Score, die ASA-Klassifikation oder andere Komorbidität Indexes wie der „Charlson Comorbidity Index“ benutzt (61-63). Dies hätte einen Vergleich zwischen wenig und stark vorerkrankten Patient*innen, sowie eine umfassendere Beschreibung der der Population ermöglicht. Eine weitere Limitation ist, dass die Stellvertreter*innen der Patient*innen welche intensivmedizinisch behandelt wurden, die Patient*innen häufig erst besucht hatte, als diese bereits auf eine Normalstation verlegt wurden. Dadurch konnten die Stellvertreter*innen häufig nicht während der akuten Phase der intensivmedizinischen Behandlung befragt werden.

Wie bereits diskutiert, wurde von einem Großteil der Patient*innen angegeben, dass diese „keine Probleme“ haben. Die Ergebnisse dieser Studie sind somit nicht ohne Vorbehalte auf eine stärker eingeschränkte Population anwendbar, da gerade bei den Fragen mit den geringsten Einschränkungen die höchste Übereinstimmung vorlag. Eine weitere Limitation dieser

Ergebnisse ist, dass nur die deutsche Version des WHODAS 2.0 in dieser Studie verwendet wurde. Somit sind keine direkten Aussagen zu nicht Deutsch sprechenden Patient*innen und deren Stellvertreter*innen zu machen. Es wurden ebenfalls nur Patient*innen befragt, welche selbst in der Lage waren, die Fragebögen auszufüllen, wodurch ebenfalls keine Aussage zur Übereinstimmung von Patient*innen mit starken kognitiven Einschränkungen und deren Stellvertreter*innen möglich sind.

6 Anhang

WHODAS 2.0 - Patientenversion

WHODAS 2.0 12 Item

Wie viele Schwierigkeiten hatten Sie in den letzten **30 Tagen**:

		Keine	Geringe	Mäßige	Starke	Sehr starke/ nicht möglich
S1	Eine längere Zeit (ca. 30 Minuten) zu stehen?					
S2	Ihren Haushaltspflichten nachzukommen?					
S3	Neue Aufgaben zu lernen (z.B. lernen an einen neuen Ort zu gelangen)?					
S4	An gesellschaftlichen Aktivitäten (wie z.B. Festlichkeiten, religiöse oder andere Aktivitäten) in der gleichen Weise teilzunehmen, wie andere?					
S5	Wie sehr sind sie durch ihren gesundheitlichen Status emotional belastet?					
S6	Sich für 10 Minuten auf etwas zu konzentrieren (z.B. ein Buch lesen)?					
S7	Eine längere Strecke (ca. 1 km) zu laufen?					
S8	Ihren gesamten Körper zu waschen?					
S9	Sich anzuziehen?					
S10	Im Umgang mit Personen die sie nicht kennen?					
S11	Eine Freundschaft aufrechtzuerhalten?					
S12	Bei der bewältigung ihrer Arbeits-/Schulalltags?					

H1	An wie vielen Tagen traten diese Schwierigkeiten während der letzten 30 Tage auf?	
H2	An wie vielen Tagen in den letzten 30 Tagen waren sie aufgrund ihrer Gesundheitsprobleme absolut unfähig alltägliche Aktivitäten oder ihre Arbeit zu verrichten?	
H3	An wie vielen Tagen in den letzten 30 Tagen mussten Sie aufgrund Ihrer Gesundheitsprobleme alltägliche Aktivitäten oder Ihre Arbeit reduzieren?	

WHODAS 2.0 – Stellvertreter*innen-version

WHODAS 2.0 12 Item

Wie viele Schwierigkeiten hatte Ihr Angehöriger in den letzten **30 Tagen**:

		Keine	Geringe	Mäßige	Starke	Sehr starke/ nichtmöglich
S1	Eine längere Zeit (ca. 30 Minuten) zu stehen?					
S2	Seinen/Ihren Haushaltspflichten nachzukommen?					
S3	Neue Aufgaben zu lernen (z.B. lernen an einen neuen Ort zu gelangen)?					
S4	An gesellschaftlichen Aktivitäten (wie z.B. Festlichkeiten, religiöse oder andere Aktivitäten) in der gleichen Weise teilzunehmen, wie andere?					
S5	Wie sehr ist er/sie durch seinen/ihren gesundheitlichen Status emotional belastet?					
S6	Sich für 10 Minuten auf etwas zu konzentrieren (z.B. ein Buch lesen)?					
S7	Eine längere Strecke (ca. 1 km) zu laufen?					
S8	Seinen/Ihren gesamten Körper zu waschen?					
S9	Sich anzuziehen?					
S10	Im Umgang mit Personen die er/sie nicht kennt?					
S11	Eine Freundschaft aufrechtzuerhalten?					
S12	Bei der Bewältigung seines/ihrer Arbeits-/Schulalltags?					

H1	An wie vielen Tagen traten diese Schwierigkeiten während der letzten 30 Tage auf?	
H2	An wie vielen Tagen in den letzten 30 Tagen war Ihr/e Verwandte/r aufgrund seiner/ihrer Gesundheitsprobleme absolute unfähig alltägliche Aktivitäten oder seine/ihre Arbeit zu verrichten?	
H3	An wie vielen Tagen in den letzten 30 Tagen musste Ihr/e Verwandte/r aufgrund seiner/ihrer Gesundheitsprobleme alltägliche Aktivitäten oder seine/ihre Arbeit reduzieren?	

MOCA-Blind

Montreal Cognitive Assessment (MoCA)-Blind

Aufgabe	Punkte
Wiederholen von 5 Wörtern: Gesicht, Samt, Kirche, Tulpe, Rot - 1 Punkt pro richtigem Wort -	/5P
Zahlen wiederholen: [2 1 8 5 4] -1 Punkt, wenn alle richtig -	/2 P
Zahlen Rückwärts wiederholen: [7 4 2] - 1 Punkt, wenn alle richtig -	
Folgende Buchstaben vorlesen. Patient soll bei dem Buchstaben "A" mit der Hand klopfen F B A C M N A A J K L B A F A K D E A A A J A M O F A A B - 1 Punkt bei maximal 2 Fehlern -	/1 P
Serielle Subtraktion: 100 - 7 (93, 86,79,72,65) - 4 o. 5 Werte korrekt = 3 Punkte - - 2 o. 3 Werte korrekt = 2 Punkte - - 1 Wert korrekt = 1 Punkt -	/3 P
Wiederholen von folgenden Sätzen: " Ich weiß lediglich, dass Hans heute an der Reihe ist zu helfen" " Die Katze versteckte sich immer unter der Couch, wenn die Hunde im Zimmer waren" - 1 Punkt pro Satz -	/2 P
Möglichst viele Wörter in einer Minute nennen, die mit dem Buchstaben "F" beginnen -1 Punkt, wenn mindestens 11 Wörter genannt wurden -	/1 P
Gemeinsamkeiten von Wörtern beschreiben, z.B. Banane und Apfelsine = Frucht: -> Eisenbahn - Fahrrad [] -> Uhr - Lineal [] - 1 Punkt pro nachvollziehbaren Vergleich -	/2 P
Wiederholen der Anfangs genannten 5 Wörter: Gesicht [], Samt [], Kirche [], Tulpe [], Rot [] - 1 Punkt pro Wort, nur bei richtigem Nennen OHNE Hinweise -	/5 P
Korrektes nennen von: Datum, Monat, Jahr, Wochentag, Ort, Stadt - 1 Punkt pro richtigem Wort -	/6 P
Gesamt	

7 Literaturverzeichnis

1. Rawal G, Yadav S, Kumar R. Post-intensive Care Syndrome: an Overview. *J Transl Int Med.* 2017;5(2):90-2.
2. Quasim T, Brown J, Kinsella J. Employment, social dependency and return to work after intensive care. *J Intensive Care Soc.* 2015;16(1):31-6.
3. Herridge MS, Tansey CM, Matte A, Tomlinson G, Diaz-Granados N, Cooper A, Guest CB, Mazer CD, Mehta S, Stewart TE, Kudlow P, Cook D, Slutsky AS, Cheung AM, Canadian Critical Care Trials G. Functional disability 5 years after acute respiratory distress syndrome. *N Engl J Med.* 2011;364(14):1293-304.
4. Denehy L, Hough CL. Critical illness, disability, and the road home. *Intensive Care Med.* 2017;43(12):1881-3.
5. Turnbull AE, Sepulveda KA, Dinglas VD, Chessare CM, Bingham CO, 3rd, Needham DM. Core Domains for Clinical Research in Acute Respiratory Failure Survivors: An International Modified Delphi Consensus Study. *Crit Care Med.* 2017;45(6):1001-10.
6. Needham DM, Sepulveda KA, Dinglas VD, Chessare CM, Friedman LA, Bingham CO, 3rd, Turnbull AE. Core Outcome Measures for Clinical Research in Acute Respiratory Failure Survivors. An International Modified Delphi Consensus Study. *Am J Respir Crit Care Med.* 2017;196(9):1122-30.
7. Guidet B, Vallet H, Boddaert J, de Lange DW, Morandi A, Leblanc G, Artigas A, Flaatten H. Caring for the critically ill patients over 80: a narrative review. *Ann Intensive Care.* 2018;8(1):114.
8. Flaatten H. Intensive care in the very old: are we prepared? *Acta Anaesthesiol Scand.* 2007;51(5):519-21.
9. Laake JH, Dybwik K, Flaatten HK, Fonneland IL, Kvale R, Strand K. Impact of the post-World War II generation on intensive care needs in Norway. *Acta Anaesthesiol Scand.* 2010;54(4):479-84.
10. Hofhuis JG, Spronk PE, van Stel HF, Schrijvers AJ, Bakker J. Quality of life before intensive care unit admission is a predictor of survival. *Crit Care.* 2007;11(4):R78.
11. Langlet K, Van Der Linden T, Launois C, Fourdin C, Cabaret P, Kerkeni N, Barbe C, Lebarry F, Deslee G. Predictive value of daily living score in acute respiratory failure of COPD patients requiring invasive mechanical ventilation pilot study. *BMC Pulm Med.* 2012;12:66.
12. Pietilainen L, Hastbacka J, Backlund M, Parviainen I, Pettila V, Reinikainen M. Premorbid functional status as a predictor of 1-year mortality and functional status in intensive care patients aged 80 years or older. *Intensive Care Med.* 2018;44(8):1221-9.
13. Wehler M, Geise A, Hadzionerovic D, Aljukic E, Reulbach U, Hahn EG, Strauss R. Health-related quality of life of patients with multiple organ dysfunction: individual changes and comparison with normative population. *Crit Care Med.* 2003;31(4):1094-101.
14. Lynn Snow A, Cook KF, Lin PS, Morgan RO, Magaziner J. Proxies and other external raters: methodological considerations. *Health Serv Res.* 2005;40(5 Pt 2):1676-93.
15. Pickard AS, Knight SJ. Proxy evaluation of health-related quality of life: a conceptual framework for understanding multiple proxy perspectives. *Med Care.* 2005;43(5):493-9.
16. Tourtier JP, Moullec DL, Auroy Y. How surrogate decision makers can provide an accurate point of view to discuss limitation of treatment. *Am J Respir Crit Care Med.* 2010;182(11):1455-6.
17. Üstün TB, Chatterji S, Villanueva M, Bendib L, Sadana R, Valentine N, Ortiz J, Tandon A, Salomon J, Cao Y, Jun X, Özaltın E, Mathers C, editors. WHO Multi-country Survey Study on Health and Responsiveness 2000-20012001.

18. National Center for Biotechnology Information [Available from: <https://www.ncbi.nlm.nih.gov/>. Zugriffsdatum: 24.10.2019
19. Maxwell CA, Dietrich MS, Minnick AF, Mion LC. Preinjury Physical Function and Frailty in Injured Older Adults: Self- Versus Proxy Responses. *J Am Geriatr Soc.* 2015;63(7):1443-7.
20. Dinglas VD, Gifford JM, Husain N, Colantuoni E, Needham DM. Quality of life before intensive care using EQ-5D: patient versus proxy responses. *Crit Care Med.* 2013;41(1):9-14.
21. Hofhuis JG, Dijkgraaf MG, Hovingh A, Braam RL, van de Braak L, Spronk PE, Rommes JH. The Academic Medical Center Linear Disability Score for evaluation of physical reserve on admission to the ICU: can we query the relatives? *Crit Care.* 2011;15(5):R212.
22. Scales DC, Tansey CM, Matte A, Herridge MS. Difference in reported pre-morbid health-related quality of life between ARDS survivors and their substitute decision makers. *Intensive Care Med.* 2006;32(11):1826-31.
23. Elliott D, Lazarus R, Leeder SR. Proxy respondents reliably assessed the quality of life of elective cardiac surgery patients. *J Clin Epidemiol.* 2006;59(2):153-9.
24. Hofhuis J, Hautvast JLA, Schrijvers AJP, Bakker J. Quality of life on admission to the intensive care: can we query the relatives? *Intensive Care Med.* 2003;29(6):974-9.
25. Capuzzo M, Grasselli C, Carrer S, Gritti G, Alvisi R. Quality of life before intensive care admission: agreement between patient and relative assessment. *Intensive Care Med.* 2000;26(9):1288-95.
26. Gifford JM, Husain N, Dinglas VD, Colantuoni E, Needham DM. Baseline quality of life before intensive care: a comparison of patient versus proxy responses. *Crit Care Med.* 2010;38(3):855-60.
27. Chaboyer W, Foster M, Creamer J. Health status of ICU survivors: a pilot study. *Aust Crit Care.* 2002;15(1):21-6.
28. Diaz-Prieto A, Gorriz MT, Badia X, Torrado H, Farrero E, Amador J, Abos R. Proxy-perceived prior health status and hospital outcome among the critically ill: is there any relationship? *Intensive Care Med.* 1998;24(7):691-8.
29. Rogers J, Ridley S, Chrispin P, Scotton H, Lloyd D. Reliability of the next of kins' estimates of critically ill patients' quality of life. *Anaesthesia.* 1997;52(12):1137-43.
30. Badia X, Diaz-Prieto A, Rue M, Patrick DL. Measuring health and health state preferences among critically ill patients. *Intensive Care Med.* 1996;22(12):1379-84.
31. Magaziner J, Simonsick EM, Kashner TM, Hebel JR. Patient-proxy response comparability on measures of patient health and functional status. *J Clin Epidemiol.* 1988;41(11):1065-74.
32. Schlote A, Richter M, Wunderlich MT, Poppendick U, Moller C, Wallesch CW. [Use of the WHODAS II with stroke patients and their relatives: reliability and inter-rater-reliability]. *Rehabilitation (Stuttg).* 2008;47(1):31-8.
33. Ramos-Remus C, Castillo-Ortiz JD, Sandoval-Castro C, Paez-Agraz F, Sanchez-Ortiz A, Aceves-Avila FJ. Divergent perceptions in health-related quality of life between family members and patients with rheumatoid arthritis, systemic lupus erythematosus, and ankylosing spondylitis. *Rheumatol Int.* 2014;34(12):1743-9.
34. World Health Organisation. [Available from: https://www.who.int/classifications/icf/form_whodas_downloads/en/ Zugriffsdatum: 25.02.2019
35. MOCA Cognitive Assessment [Available from: <https://www.mocatest.org/>.
36. Tarvonen-Schroder S, Hurme S, Laimi K. The World Health Organization Disability Assessment Schedule (WHODAS 2.0) and the WHO Minimal Generic Set of Domains of

- Functioning and Health versus Conventional Instruments in subacute stroke. *J Rehabil Med*. 2019;51(9):675-82.
37. Wittich W, Phillips N, Nasreddine ZS, Chertkow H. Sensitivity and Specificity of the Montreal Cognitive Assessment Modified for Individuals who are Visually Impaired. *Journal of Visual Impairment & Blindness*. 2010;104(6):360-8.
 38. Sullivan GM, Feinn R. Using Effect Size-or Why the P Value Is Not Enough. *J Grad Med Educ*. 2012;4(3):279-82.
 39. Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*. 2007;39(2):175-91.
 40. Völkl K, Korb C. *Deskriptive Statistik* 1st ed: Springer VS; 2018.
 41. Ong LD, LeClare PC. The Kolmogorov-Smirnov test for the log-normality of sample cumulative frequency distributions. *Health Phys*. 1968;14(4):376.
 42. Federici S, Bracalenti M, Meloni F, Luciano JV. World Health Organization disability assessment schedule 2.0: An international systematic review. *Disabil Rehabil*. 2017;39(23):2347-80.
 43. Bejer A, Cwirlej-Sozanska A, Wisniowska-Szurlej A, Wilmowska-Pietruszynska A, Spalek R, de Sire A, Sozanski B. Psychometric properties of the Polish version of the 36-item WHODAS 2.0 in patients with hip and knee osteoarthritis. *Qual Life Res*. 2021.
 44. Saltychev M, Katajapuu N, Barlund E, Laimi K. Psychometric properties of 12-item self-administered World Health Organization disability assessment schedule 2.0 (WHODAS 2.0) among general population and people with non-acute physical causes of disability - systematic review. *Disabil Rehabil*. 2021;43(6):789-94.
 45. Katajapuu N, Heinonen A, Saltychev M. Minimal clinically important difference and minimal detectable change of the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) amongst patients with chronic musculoskeletal pain. *Clin Rehabil*. 2020;34(12):1506-11.
 46. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307-10.
 47. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: Measures of agreement. *Perspect Clin Res*. 2017;8(4):187-91.
 48. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;15(2):155-63.
 49. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull*. 1968;70(4):213-20.
 50. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol*. 1993;46(5):423-9.
 51. Byrt T. Problems with kappa. *J Clin Epidemiol*. 1992;45(12):1452.
 52. Vannest KJP, R.I.; Gonen, O. ; , Adiguzel T. Pabak-OS Calculator 2016 [Available from: <http://www.singlecaseresearch.org/calculators/pabak-os>. Zugriffdatum 16.06.2020
 53. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-74.
 54. Fleiss J. *Psychological Bulletin*. 1969;72(5):323-7.
 55. Ustun TB, Kostanjsek N, Chatterji S, Rehm J, World Health O. *Measuring health and disability : manual for WHO Disability Assessment Schedule (WHODAS 2.0) / edited by T.B. Üstün, N. Kostanjsek, S. Chatterji, J.Rehm*. Geneva: World Health Organization; 2010.
 56. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276-82.

57. Steel JL, Geller DA, Carr BI. Proxy ratings of health related quality of life in patients with hepatocellular carcinoma. *Qual Life Res.* 2005;14(4):1025-33.
58. Cohen J. *Statistical power analysis for the behavioral sciences.* Hillsdale, N.J.: L. Erlbaum Associates; 1988.
59. Wawrzyniak KM, Finkelman M, Schatman ME, Kulich RJ, Weed VF, Myrta E, DiBenedetto DJ. The World Health Organization Disability Assessment Schedule-2.0 (WHODAS 2.0) in a chronic pain population being considered for chronic opioid therapy. *J Pain Res.* 2019;12:1855-62.
60. Romhild J, Fleischer S, Meyer G, Stephan A, Zwakhalen S, Leino-Kilpi H, Zabalegui A, Saks K, Soto-Martin M, Sutcliffe C, Rahm Hallberg I, Berg A, RightTimePlaceCare C. Inter-rater agreement of the Quality of Life-Alzheimer's Disease (QoL-AD) self-rating and proxy rating scale: secondary analysis of RightTimePlaceCare data. *Health Qual Life Outcomes.* 2018;16(1):131.
61. Rapsang AG, Shyam DC. Scoring systems in the intensive care unit: A compendium. *Indian J Crit Care Med.* 2014;18(4):220-8.
62. Molto A, Dougados M. Comorbidity indices. *Clin Exp Rheumatol.* 2014;32(5 Suppl 85):S-131-4.
63. Doyle DJ, Goyal A, Bansal P, Garmon EH. American Society of Anesthesiologists Classification. *StatPearls.* Treasure Island (FL)2021.

Eidesstattliche Versicherung

Ich, Lukas Schönagel, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: „Verlässlichkeit von Angehörigen bei der Beurteilung der prä-morbiden Gesundheit kritisch kranker Patienten“ (Engl.: „Reliability of proxies in the evaluation of the pre-morbid health related quality of life of critically ill Patients“) selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren/innen beruhen, sind als solche in korrekter Zitierung kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) werden von mir verantwortet.

Ich versichere ferner, dass ich die in Zusammenarbeit mit anderen Personen generierten Daten, Datenauswertungen und Schlussfolgerungen korrekt gekennzeichnet und meinen eigenen Beitrag sowie die Beiträge anderer Personen korrekt kenntlich gemacht habe (siehe Anteilserklärung). Texte oder Textteile, die gemeinsam mit anderen erstellt oder verwendet wurden, habe ich korrekt kenntlich gemacht.

Meine Anteile an etwaigen Publikationen zu dieser Dissertation entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Erstbetreuer/in, angegeben sind. Für sämtliche im Rahmen der Dissertation entstandenen Publikationen wurden die Richtlinien des ICMJE (International Committee of Medical Journal Editors; www.icmje.org) zur Autorenschaft eingehalten. Ich erkläre ferner, dass ich mich zur Einhaltung der Satzung der Charité – Universitätsmedizin Berlin zur Sicherung Guter Wissenschaftlicher Praxis verpflichte.

Weiterhin versichere ich, dass ich diese Dissertation weder in gleicher noch in ähnlicher Form bereits an einer anderen Fakultät eingereicht habe.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§§156, 161 des Strafgesetzbuches) sind mir bekannt und bewusst.

Datum

Unterschrift

Lebenslauf

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht

Danksagung

An dieser Stelle möchte ich allen Personen danken, die mich bei der Vollendung dieser Dissertation unterstützt haben. Mein besonderer Dank gilt dabei den Teilnehmern dieser Studie.

An zweiter Stelle möchte ich mich bei Dr. med. Tobias Wollersheim und Prof. Dr. med. Steffen Weber-Carstens bedanken, auf dessen Unterstützung und Ratschläge ich mich während dieses Projektes stets verlassen konnte.

Weiterhin gilt mein Dank Prof. Dr. med. Claudia Spies als Direktorin der Klinik, durch welche dieses Projekt erst möglich wurde.

Als letztes möchte ich mich auch bei meiner Familie und Freunden bedanken, welche mich durch die guten und herausfordernden Zeiten begleitet haben.

Bescheinigung des akkreditierten Statistikers



CharitéCentrum für Human- und Gesundheitswissenschaften

Charité | Campus Charité Mitte | 10117 Berlin

Institut für Biometrie und klinische Epidemiologie (iBIKE)

Direktor: Prof. Dr. Geraldine Rauch

Postanschrift:
Charitéplatz 1 | 10117 Berlin
Besucheranschrift:
Reinhardtstr. 58 | 10117 Berlin

Tel. +49 (0)30 450 562171
geraldine.rauch@charite.de
<https://biometrie.charite.de/>



Name, Vorname: Schönagel, Lukas
Emailadresse: lukas.schoennagel@charite.de
Matrikelnummer: 219585
PromotionsbetreuerIn: Prof. Dr. Steffen Weber-Carstens
Promotionsinstitution/ Klinik: CC07 Klinik für Anästhesiologie
m.S. operative Intensivmedizin CVK

Bescheinigung

Hiermit bescheinige ich, dass Herr Lukas Schönagel innerhalb der Service Unit Biometrie des Instituts für Biometrie und klinische Epidemiologie (iBIKE) bei mir eine statistische Beratung zu einem Promotionsvorhaben wahrgenommen hat. Folgende Beratungstermine wurden wahrgenommen:

- Termin 1: 21.02.2019
- Termin 2: 24.10.2019
- Termin 3: 10.05.2021

Folgende wesentliche Ratschläge hinsichtlich einer sinnvolle Auswertung und Interpretation der Daten wurden während der Beratung erteilt:

- Hypothesenformulierung im Zusammenhang mit statistischen Auswertungsverfahren
- Auf Interpretierbarkeit der ordinalen Abstände eingehen und Vor- und Nachteile der statistischen Methoden knapp beschreiben.
- Fallzahlen zu (Unter-)Auswertungen angeben.

Diese Bescheinigung garantiert nicht die richtige Umsetzung der in der Beratung gemachten Vorschläge, die korrekte Durchführung der empfohlenen statistischen Verfahren und die richtige Darstellung und Interpretation der Ergebnisse. Die Verantwortung hierfür obliegt allein dem Promovierenden. Das Institut für Biometrie und klinische Epidemiologie übernimmt hierfür keine Haftung.

Datum: 10.05.2021

Name des Beraters/ der Beraterin: Carolin Herrmann



Unterschrift Beraterin, Institutsstempel

Campus Charité Mitte
Charitéplatz 1 | D-10117 Berlin
Site: Reinhardtstr. 58